SCHOLARLY PUBLICATIONS

A CURRENT AWARENESS BULLETIN
OF RESEARCH OUTPUT

@DTU

(148th Edition)

APRIL 2025

BY: CENTRAL LIBRARY

DELHI TECHNOLOGICAL UNIVERSITY
(FORMERLY *DELHI COLLEGE OF ENGINEERING*)
GOVT. OF N.C.T. OF DELHI
SHAHBAD DAULATPUR, MAIN BAWANA ROAD
DELHI 110042

PREFACE

This is the **One Hundred Forty Eight** Issue of Current Awareness Bulletin started by Delhi Technological University, Central Library. The aim of the bulletin is to compile, preserve and disseminate information published by the faculty, students and alumni for mutual benefits. The bulletin also aims to propagate the intellectual contribution of Delhi Technological University (DTU) as a whole to the academia.

The bulletin contains information resources available in the internet in the form of articles, reports, presentations published in international journals, websites, etc. by the faculty and students of DTU. The publications of faculty and student which are not covered in this bulletin may be because of the reason that the full text either was not accessible or could not be searched by the search engine used by the library for this purpose.

The learned faculty and students are requested to provide their uncovered publications to the library either through email or in CD, etc. to make the bulletin more comprehensive.

This issue contains the information published during **April, 2025**. The arrangement of the contents is alphabetical. The full text of the article which is either subscribed by the university or available in the web is provided in this bulletin.

Central Library

CONTENTS

1. A Review on Feature Selection Techniques for Anxiety Classification from Physiological Signals, ***6.Shikha Shikha, 3.Divyashikha Sethia, 3.S. Indu,*** CSE,SE, ECE, DTU

2. A relational disaster database to document and resolve multihazard interactions, Maximillian Van Wyk de Vries, Lorenzo Nava, Ye Chen, Lisa Augustina, Louie Bell1, Joshua Nicholas, Ben Clarke, Reka Ungar, Arthur Hill, ***6.Kamini Sharma***, Julie Morin and Ekbal Hussain, Civil, DTU

3. A Study on the Viscosity of Viscoelastic Magnetic Abrasive Gel under Different Volumetric Proportions, ***3.Krovvidi Srinivas***, Mechanical Engg., DTU

4. Advancing Solar Energy through Artificial Intelligence: A Focus on Optimization and Forecasting, ***6.Nikhil Kushwaha, 3.Vinod Kumar Yadav, 3.Radheshyam Saha***, Electrical Engg., DTU

5. An Interplay Between Hypothalamic Microstructure, Systemic Metabolism and Gut Microbiome Composition in Male Rats at Hyperacute Timepoint Post TBI, Palkin Arora, ***6. Megha Kumari***, Kavita Singh, M. Memita Devi, Poonam Rana, Rajat Sandhir and Richa Trivedi, Biotech, DTU

6. AI-Driven Mental Healthcare 5.0: A Survey of Opportunities and Challenges in Leveraging Large Language Models and Generative AI, Akshi Kumar, ***8.1 Karan Gupta, 8.1 Kartik Verma, and 3.Sanjay Kumar***, CSE, DTU

7. Analysis of a Digitizer Circuit with T-Network for Low Resistance Measurement, Shiraz Sohail, ***7.1 Gopal Singh***, Tarikul Islam, Control and instrumentation, DTU

8. Analyzing Digital Image Anti-Forensics Techniques and Countermeasures, Saurabh Agarwal, ***3.Rajeev Kumar***, Abha Devi, Ki-Hyun Jung, CSE, DTU

9. Anomaly Detection for Sensor Manipulation in Matter Enabled-IoT Devices with Faulty Data Injection, ***7. Jahnvi Patel, 3.Kapil Sharma,*** IT DTU

10. Assessment of shear bond strength of triplets using different joint thicknesses and mortar types***, 3.Shilpa PAL , 7.1 Anshoo YADAV,*** & ***6. Khyati SAGGU***, Civil, DTU

11. A wideband dual polarized in-band full duplex antenna with higher order modes suppression, ***7.1 Sandeep Rana***,***8.Amit Kumar*** and Akhilesh Mohan, ECE, DTU

12. Advances in Negative Capacitance FinFET Technology: Insights into Design Innovations and Performance Enhancements, ***7.Shubhda, 3.Rishu Chaujar***, Applied Physics, DTU

13. A Novel n+ Pocket Doped L-Shaped Channel TFET for Biosensing Applications, ***8. Danishk Barwa,8. Dhruv Khanna, 8.Harshit Parihar***, & ***3.Sumit Kale***, ECE, DTU

14. Advanced Chest X-Ray Analysis via Transformer-Based Image Descriptors and Cross-Model Attention Mechanism***, 6. Lakshita Agarwal, 3.Bindu Verma***, IT DTU

15. Aggregation of Distributed Energy Resources for Optimal Operationin Local Energy Communities With Grid Constraints,***6. ShubhamGupta*** ,Vinod Kumar Yadav and ***3.MadhusudanSingh***, Electrical Engg., DTU

16. Achieving the positivity of the secret key in a BB84 like quantum key distribution protocol, ***6.Rashi Jain*** and ***3.Satyabrata Adhikar***, Applied Mathematics, DTU

17. Alzheimer's Disease Detection Using Bidirectional Empirical Mode Decomposition, Neha Garg, Amit Garg, ***3.Mahipal Singh Choudhry***, ECE, DTU

18. Bernstein type semi-exponential operators***, 6. Kanita*** and ***3.Naokant Deo***, Applied Maths, DTU

19. Building Scalable and Quantum Attack Resistant Authenticated Message Delivery System for Internet of Vehicles With Blockchain Consensus Mechanism***, 3. Rahul Singh, 3.Laxminarayan Das***, Dharminder Chaudhary, Applied Mathematics, DTU

20. Carboxymethyl Guar Gum: Bridging the Gap Between Biopolymers and Advanced Functional Materials, Vratika Verma, ***6.1 Jagram Meena***, Teena Saini, Sakshi Saini & Rajdeep Malik, Applied Chemistry, DTU

21. Conventional and Intelligent Control Schemes for Handling A Helicopter's Pitch Axis, ***3.Abhishek Chaudhary***, Electrical Engg., DTU

22. Comparative analysis of in silico tools for Drug Side effect prediction, ***6.Harshita Thakur, 3.Asmita Das***, Biotechnology, DTU

23. Coordination Frameworks for IoT-Driven Swarm Robotics Communication, Ayush Kumar Agrawal, Jayendra Kumar, ***3. Avanish Kumar***, Anumeha, Arvind R. Yadav, ECE, DTU

24. Development of intelligent hybrid controller for torque ripple minimization in electric drive system with adaptive flux estimator: An experimental case study, Surya Kant***, 3.Mini Sreejeth, 3.Madhusudan Singh***, Ambrish Devanshu, Majed A. Alotaibi, Hasmat Malik, Fausto Pedro García Márquez, Electrical Engg., DTU

25. Does ESG impact corporate performance in environmentally sensitive and non-sensitive industries? A dynamic panel approach, ***6.Prachi Jain*** and ***3.Priya Malhotra***, USME, DTU

26. Diverse Machine Learning Approaches for Early Prediction of Heart Attack, ***3.Meena Rawat***, Dhruv Shakti Singh and Rahul Boadh, Applied Science, DTU

27. Decaying vacuum energy, matter creation and cosmic acceleration, ***6 Lokesh Chander and 3.C P Singh***, Applied Mathematics, DTU

28. Design and Analysis of Substrate Integrated Folded Waveguide and Filter, Sudarshan Kumar and ***3.AsokDe***, ECE, DTU

29. Diffractive optical dermal sensor for non invasive arterial pulse wave monitoring, ***3.1. Mayank Pandey***, Sang-Woo Seo, Teaching Assistant, DTU

30. Donepezil Improves PSD95 Expression, Mitigates Neuroinflammation via PI3K/Akt/NF- κB and Mitochondrial Dysfunction in a Rodent Model of Subarachnoid Haemorrhage, Ahmed Shaney Rehman,,Ammar Tasleem, Nemat Ali, Rehan Khan, Mohd. Salman, ***3.Pravir Kumar*** and Suhel Parvez, Biotechnology, DTU

31. Enhanced Security In Matter-Enabled Iot Networks Through Anomaly Detection***, 6. Manjit Kumar***, ***3.Kapil Sharma***, IT, DTU

32. Evaluating the intensity of Typhoons over the West Pacific basin, ***6. Monu Yadav*** and ***3. Laxminarayan Das,*** Applied Mathematics, DTU

33. Enhancing the thermoelectric performance of Janus MoSSe monolayer via pressure, Anisha, Ramesh Kumar, ***3.Mukhtiyar Singh***, Sunita Srivastava, Tankeshwar Kumar, Applied Physics, DTU

34. Energy Efficient 8T SRAM Cell for Enhancing Read/Write Assist Performances with Stability, ***3.Ram Murti Rawat, 3.Vinod Kumar,*** Computer Science, DTU

35. Effect of Alfvén wave dispersion on localization and current sheet formation in solar wind near reconnection region***, 6.Ritu, 6.Jyoti, 3.Suresh C Sharma*** and R P Sharma, Applied Physics, DTU

36. Enhancing ground stability with quarry dust columns: A comparative study of ordinary and encased configurations***, 6.1 Srijan Srijan***, Civil, DTU

37. Exploring multi-material structures: An overview of arc additive Manufacturing, ***6.Dhruv Kumar, 3.Vipin and 3.Reeta Wattal***, Mechanical, DTU

38. Evaluating the intensity of Typhoons overtheWestPacificbasin, ***6. MonuYadav and 3.LaxminarayanDas***, Applied Maths, DTU

39. Exploring the dermal safety of green-synthesized Ag–TiO2 nanocomposites for topical applications, Archana Rana, Shweta Parashar, Diksha Singh, Kavita Singh, Debabrata Chanda, Anirban Pal, Ritu Srivastava and ***3.Shailesh Narain Sharma,*** Applied Physics, DTU

40. Exploring the psychometric properties of the adult hope scale in an Indian university student sample***, 6.Suhans Bansal***, Naval Garg,and Jagvinder Singh, USME, DTU

41. GRU Based Autoencoder with Maximum Overlap Discrete Wavelet Transfrmers for brain tumor detection***, 6. Bhawana Sachdeva, 3.Jeebananda Panda***, ECE, DTU

42. GMM-based Imbalanced Fractional Whale Particle Filter for Multiple Object Tracking in Surveillance Videos, ***3.Avinash Ratre***, ECE

43. Integrated wearable PPG: a multi-vital sign monitoring based on group sparse mode decomposition framework in remote health care using PPG signal, Pankaj · Pratibha Maan,· ***3.Manjeet Kumar***· Ashish Kumar,· Rama Komaragiri, ECE,DTU

44. Hybrid Anomaly Detection Framework for Matter-Enabled IoT Devices, ***6.Vaibhav Krishan, 8. Shivam Dhar, 3.Kapil Sharma***, IT, DTU

45. iHDoc: Secure NFC Mobile-based Digital Document Signatures Prototype ***7.Himanshu Sharma, 3.Divyashikha Sethia***, Computer Science, DTU

46. Incurred Changes in Water and Fat-Soluble Vitamin Profile of Alternative Ingredient Spirodela Polyrhiza Based Partial Fishmeal Substituted Aquafeed: Effect of Long Duration Variable Temperature Storage, Parul Puri1, Neha Tiwari, Vaibhav Puri, ***3.Ram Singh***, Jai Gopal Sharma, Rina Chakrabarti , Applied Chemistry, DTU

47. Linear and Nonlinear Photoluminescence in Ho3+, Yb3+ Doped CaBi2Nb2O9 Phosphor for Photonic Application***, 6.Sheetal Kumari, 6.Pooja Rohilla***, Anu, Srinibas Barik, Bhanupriya Yadav, Kaushal Kumar2, Runit Isaac***, 3. A.S.Rao***, Aman Prasad, Applied Physics, DTU

48. Literature review on SRAM cells under Stability, Energy, and Power Constraints, ***3.Ram Murti Rawat, 3.Vinod Kumar***, Computer Science, DTU

49. Longitudinal analysis of Covid-19 vaccine-related tweets in india: linking sentiment fluctuations with topic modeling***, 3.Seba Susan, 6.Anmol Bansal, 8.1Arjun Choudhry*** and ***8.1 Anubhav Sharma***, IT, DTU

50. Latch and Flip-Flop Design Using Memristors and Transmission Gates with Timer Counter Implementation, ***7.Aditya Rayer, 3.Mavulluru Ganesh and 3.Lavi Tanwar***, ECE, DTU

51. MXene coated concave shaped microchannel PCF SPR biosensor for the detection of HIV and sickle cell anaemia, ***7.Preethi Krishnan, 6.Akash Khamaru and 3.Ajeet Kumar***, Applied Physics, DTU

52. Modified Sigma-Pi neural network for dynamical system identification, ***6.Kartik Saini, 3.Narendra Kumar*** and ***3.Rajesh Kumar***, Electrical, DTU

53. Navigating the green combustion landscape: Optimizing emissions and performance in CI engines fueled by biogas and nanoparticle- doped biodiesel, S. LALHRIATPUIA, H. LALHMINGSANGA, ***6.Gulam MUSTAFA, 3.Neeraj BUDHRAJA*** and Kiran PAL, Mechanical Engg., DTU

54. Numerical Simulation on Sensitivity Modulation in Nanowire MOSFEB Detector for Biomedical Applications Amit Dasa, Anjana Bhardwaj, Kaushik Das, Shivani Yadavc***, 6.Aapurva Kaul***, Preeti Goyal, Swati Sharma, ***3.Sonam Rewari***, Binod Kumar Kanaujia & R S Gupta, ECE, DTU

55. NoCiPUF: NoC-Based Intrinsic PUF for MPSoC Authentication, ***8.1 Deepank Grover,*** Tarun Sharma, Sneha Agarwal, , Sidhartha Sankar Rout , Anushka, Madhur Kumar and Sujay Deb, ECE, DTU

56. Performance study of a retrofitted electric two-wheeler using GT-suite, ***8.1 Alpha Omar JALLOW, 8.Alphaus NKANDU , 8.Towett SAMUEL,3.Ashish Rajeshwar KULKARNI*** and ***3.Amit Pal,*** Electrical, Mechanical, DTU

57. Production of Phytase by Penicillium oxalicum in Solid State Fermentation Exhibiting Hydrolysis of Phytates and Haloperoxidase Activity, ***6.Priya,*** Bijender Singh, Ravi Kumar Goswami***, 3.Jai Gopal Sharma*** and Bhoopander Giri, Biotechnology, DTU

58. Predicting Financial Distress Using Machine Learning Techniques ***3.Pallavi Sethi, 3.Archana Singh*** and ***3.Vikas Gupta,*** USME, DSM, DTU

59. Photocatalytic degradation of crystal violet dye using phytofabricated zinc oxide nanoparticles from anthocyanin extract, ***6.Sarla Yadav, 6.Pooja Singh, 6.Gunjan Varshney, 3.Saurabh, 3.Poonam Singh, 3.Raminder Kaur***, Applied Chemistry, DTU

60. P4.0B-FAHP: prioritizing industry 4.0 barriers using fuzzy AHP—case of an Indian construction company, ***6. Ankur Tayal, 3.Saurabh Agrawal*** and ***3.Rajan Yadav***, DSM, DTU

61. ParametricAnalysisofOrganicLEDandPhoto-DiodeBased Sensor for Heart Rate Detection***, 6. SugandhaYadav, 6.PoornimaMittal***,and ShubhamNegi, ECE, DTU

62. Pointing error influence on the performance of dual-hop radio frequency/wireless optical communication system over Nakagami-m/ Fisher-Snedecor turbulence channels, ***6.Rekha Rani,*** M. Lakshmanan ***3.N. Jayanthi and 3.Anup Kumar Mandpura*** ECE, EE, DTU

63. Proposal and evaluation of a novel solar thermal driven cooling-power cogeneration system, Abdul Khaliq, ***3.Rajesh Kumar***, Tawfiq Al-Mughanam amd Khaled Zeid, Mechanical, DTU

64. Renewable Energy Systems: Engineering Solutions for a Sustainable World, ***3.Dr. Rajesh Gupta***, Mechanical, DTU

65. RankingInfluentialNodesinSocialNetworksBasedonthe Game of Thieves Algorithm, ***3.MinniJain, 8.1AdityaGupta, 8.1 AaryanArora*** and ***8.AayushPatel***, Computer Engg., DTU

66. Revocable and Privacy-Preserving CP-ABE Scheme for Secure mHealth Data Access in Blockchain, ***6.Anita Thakur***, ***6.Virender Ranga & 3.RituAgarwal,*** IT, DTU

67. Rice straw as the potential raw material for biogas production at variable temperature and feed ratio, Shubham Kumar, ***8.Sawan KAPOOR*** and ***6..Neeraj BUDHRAJA,*** Mechanical, DTU

68. SobelMNet: A Lightweight CNN based Approach for Median Filtering Detection in compressed images, ***3.Anukriti Kaushal, 3.Rajeev Kumar 3.Sanjay Kumar***, CSE

69. Substantial enhancement of the electroactive phase and dielectric properties at low loading levels of SnO 2 /MXene nanocomposite as a novel nanofiller in the PVDF matrix, ***6.Gagan Sharma, 3.Deshraj Meena***, Applied Physics, DTU

70. Stuttering Detection Using LSTM and LSTM Attention based Convolutional Neural Network, ***8.Monalisa Maity, 8.Ishita Gupta and 3.Dinesh Kumar Vishwakarma***, Applied Maths and IT, DTU

71. Space Physics of the Universe and the Evolution of π Space Quantum, ***8.1 Nishant Sahdev***, Chinmoy Bhattacharya, Biotechnology, DTU

72. Sensitivity Analysis of Dielectric Modulated Dual Gate Gallium Nitride Junctionless Accumulation Mode Nanowire FET Based Biosensor, ***6.Anupama, 3.Sonam Rewari and 3.Neeta Pandey,*** ECE, DTU

73. Stabilization and Performance Analysis of Double Link-Rotary Inverted Pendulum using LQR-I Controller***, 6.Anjali Rai, 6.Bharat Bhushan*** and ***6.Bhavnesh Jaint***, Electrical Engg., DTU

74. SS-EMERGE - self-supervised enhancement for multidimension emotion recognition using GNNs for EEG, ***6.Chirag Ahuja*** & ***3.Divyashikha Sethia,*** Computer Science, Software science, DTU

75. SobelMNet: A Lightweight CNN based Approach for Median Filtering Detection in compressed images, ***3.Anukriti Kaushal, 3.Rajeev Kumar*** and ***6.Sanjay Kumar***, Computer Science, DTU

76. Silicon slab waveguide based XOR and XNORall optical logic gates, ***6.Yash Yadav*** and ***3.Yogita Kalra***, Applied Physics, DTU

77. SDHO-KGNN:AnEffectiveKnowledge-EnhancedOptimal Graph NeuralNetwork Approach for Fraudulent Call Detection, ***6.PoojaMithoo, 3.ManojKumar,*** Computer Science, DTU

78. Temperature-dependent coupling characteristics in a silicon photonic crystal waveguide coupler, ***6.Varnam Sherawat , 3.Yogita Kalra, 3.Renuka Bokolia, 3.Ravindra Kumar Sinha***, Applied Physics, DTU

79. Tri-FusionNet: Enhancing Image Description Generation with Transformer-based Fusion Network and Dual Attention Mechanism, ***6.Lakshita Agarwal, 3.Bindu Verma***, IT, DTU

80. Towards Explainable AI: Multi-Modal Transformer for Video-based Image Description Generation, ***6.Lakshita Agarwal*** and ***3.Bindu Verma***, IT DTU

81. Towards Creating an Ecosystem of Paperless Society: Unveiling Issues and Challenges in Adoption of DigiLocker, ***3.Rajan Yadav***, DTU

82. Zn/Cu Layered Double HydroxideasaCatalyst for the Synthesis of Highly Functional ized Pyrazole Derivatives ***6. Charu Gupta, 6.Bhamini Pandey, 3.Poonam Singh and 3.Ram Singh***, Applied Chemistry, DTU

83. Urban greening for climate resilient and sustainable cities: grand challenges and opportunities, Prashant Kumar, Jeetendra Sahani , Karina Corada Perez, Ajit Ahlawat , Maria de Fatima Andrade, Maria Athanassiadou, Shi Jie Cao ,Lisa Collins , Sagnik Dey , , Silvana Di Sabatino , Christos H. Halios , Felicity Harris , Colin Horton , Luis Inostroza , Laurence Jones , Thomas Rodding Kjeldsen , Ben McCallan , Aonghus McNabola , ***3. Rajeev Kumar Mishra***

84. Vortex flow plasma reforming for hydrogen production from atomized water-methanol mixture and parameter optimization using RSM and ANN-GA***, 6.Neeraj Budhraja , 3.Amit Pa***l and ***3.R.S. Mishra***, Mechanical, DTU

1. *Vice Chancellor*
2. *Pro Vice Chancellor*
3. *Faculty*
4. *Teaching-cum-Research Fellow*
5. *Asst. Librarian*
6. *Research Scholar*
7. *PG Scholar*
8. *Undergraduate Student*

1.1. *Ex Vice chancellor*
2.1. *Ex Pro Vice Chancellor*
3.1. *Ex Faculty*
4.1. *Alumni*
5.1 *Others*
6.1. *Ex Research Scholar*
7.1. *Ex PG Scholar*
8.1. *Ex Undergraduate Student*

EGU25-12887, updated on 07 May 2025
https://doi.org/10.5194/egusphere-egu25-12887
EGU General Assembly 2025
© Author(s) 2025. This work is distributed under
the Creative Commons Attribution 4.0 License.

A relational disaster database to document and resolve multihazard interactions

Maximillian Van Wyk de Vries[1,2], Lorenzo Nava[2,1], Ye Chen[2,3], Lisa Augustina[4,5], Louie Bell[1,2], Joshua Nicholas[1], Ben Clarke[6], Reka Ungar[1], Arthur Hill[1], Kamini Sharma[7], Julie Morin[1], and Ekbal Hussain[5]

[1]Department of Geography, University of Cambridge, Cambridge, UK
[2]Department of Earth Sciences, University of Cambridge, Cambridge, UK
[3]Department of Geotechnical Engineering, Tongji University, Shanghai, China
[4]Nottingham Geospatial Institute, University of Nottingham, Nottingham, UK
[5]British Geological Survey, Keyworth, UK
[6]Earth Observatory of Singapore, Nanyang Technological University, Singapore, Singapore
[7]Department of Civil Engineering, Delhi Technological University, Delhi, India

Natural hazard-related disasters result in tens of thousands of deaths and billions in economic losses annually, with their frequency and intensity projected to increase due to climate change. These hazards, such as floods, landslides, and volcanic eruptions, often interact in ways that amplify their impacts, creating cascading or compounding risks. Despite these complexities, most hazard databases focus on single-hazard events, failing to capture critical interactions. To address this gap, we are developing a relational disaster database designed to systematically document and analyse multihazard interactions.

The database is designed to capture both individual hazard events and their interrelations, including causal, temporal, spatial, and amplifying interactions. It consists of two modules: a hazard characteristics and impacts module, which records essential details such as location, magnitude, and consequences, and a hazard linkages module, which documents relationships between hazards with attributes such as time lags, interaction intensity, and confidence levels. This scalable design is interoperable with existing databases like DesInventar and EM-DAT, enabling integration of existing data and automated processing and data analysis. We aim for flexibility, open access, and good metadata to ensure utility for both academic and operational (e.g,. disaster risk managers) end-users.

Complementing the database, we are developing tools for the automated generation of process-linked and coinciding multihazard groups, cumulative impact analysis, and the creation of associated visualizations. Initial database entries include case studies such as the 2023 Lhonak glacial lake outburst flood (GLOF), involving rainfall, landslides, and dam failure, and Hurricane Helene, highlighting meteorological, marine, and geological interactions. The database is currently at a prototype stage, and we welcome community input and collaboration to refine its design, expand its coverage, and ensure its long-term relevance and usability.

Figure: Concept diagram for the structure of the relational multihazard database.

2025 First International Conference on Advances in Computer Science, Electrical, Electronics, and Communication Technologies (CE2CT)

A Review on Feature Selection Techniques for Anxiety Classification from Physiological Signals

1st Shikha Shikha
Department of Computer Science and Engineering
Delhi Technological University
New Delhi, India
shikha.9126@gmail.com

2nd Divyashikha Sethia
Department of Software Engineering
Delhi Technological University
New Delhi, India
divyashikha@dtu.ac.in

3rd S. Indu
Department of Electronics and Communication
Delhi Technological University
New Delhi, India s.indu@dce.ac.in

***Abstract*—Anxiety disorders significantly impact individuals' quality of life and are traditionally assessed using subjective methods, which often lack accuracy. Recent advancements have enabled the use of physiological signals for anxiety classification, providing a more objective approach. However, the multidomain nature of these signals results in high-dimensional feature sets, necessitating effective feature selection. Therefore, this paper reviews various feature selection techniques—filter, wrapper, embedded, and hybrid—used in physiological-based anxiety classification, analyzing the advantages and limitations. It also highlights key challenges and suggests future research directions. Furthermore, this review introduces a proposed study utilizing a novel Corr-FGA hybrid feature selection algorithm to develop an efficient anxiety classification framework. The proposed framework achieves the highest accuracy of 88.93% using 8 features. This review aims to guide researchers in developing a robust, efficient anxiety classification model while providing insights that can help streamline feature selection and optimize machine learning model performance in real-world applications.**

***Index Terms*—Feature Selection, Physiological signals, Anxiety, Machine Learning**

I. Introduction

Anxiety disorders are among the most common mental health conditions, affecting millions worldwide. Characterized by excessive worry and fear, these disorders significantly impair daily functioning and quality of life [1]. Common subtypes include generalized anxiety disorder, social anxiety disorder, and panic disorder [1], each presenting with psychological and physiological symptoms like restlessness, muscle tension, and fatigue.

Traditional anxiety assessments rely on psychological tools such as questionnaires and clinical interviews, including the State-Trait Anxiety Inventory (STAI) and the Generalized Anxiety Disorder 7-item (GAD-7) scale [1]. These methods are inherently subjective, depending on self-reported symptoms, introducing biases and inconsistencies. Moreover, they often fail to capture physiological indicators crucial for understanding anxiety's severity.

To overcome these limitations, researchers have adopted physiological monitoring as a more objective method [2]. Physiological signals, which are regulated by the autonomic nervous system (ANS), offer valuable insights into changes associated with anxiety. Signals like electrocardiography (ECG), electrodermal activity (EDA), and electroencephalography (EEG) reflect autonomic processes, making them essential for detecting anxiety.

Furthermore, Machine Learning (ML) algorithms have gained popularity in analyzing physiological signals to automate anxiety detection. These techniques can identify complex, non-linear patterns in physiological data that traditional methods might overlook. Figure 1 shows an automated anxiety detection system that uses physiological signals and involves several technical phases, each contributing to the overall effectiveness of the system:

1) *Data Acquisition:* Researchers use wearable sensors or devices to gather physiological signals, including ECG, EEG, EDA, EMG, and RESP. These sensors allow for non-invasive, continuous monitoring, which is vital for capturing anxiety-related physiological changes in natural settings.
2) *Preprocessing:* The collected raw signals often contain noise and artifacts that must be removed to maintain data quality. Preprocessing techniques, such as filtering, normalization, and artifact removal, are applied to ensure the signals are clean and suitable for subsequent analysis.
3) *Feature Engineering:* Feature engineering involves transforming raw physiological signals into meaningful metrics. Features are extracted from different domains, such as time, frequency, and time-frequency, to capture the physiological processes comprehensively.
4) *Classification:* Researchers train machine learning models with the selected features to classify anxiety states. Algorithms like Random Forest (RF), Support Vector Machines (SVM), and Convolutional Neural Networks (CNN) are commonly employed to effectively identify anxiety-related patterns in physiological data, offering improved accuracy and robustness over traditional methods.
5) *Evaluation:* Finally, the model is evaluated to assess its effectiveness in classifying anxiety states, using metrics

979-8-3315-1857-8/25/$31.00 ©2025 IEEE

2025 First International Conference on Advances in Computer Science, Electrical, Electronics, and Communication Technologies (CE2CT) | 979-8-3315-1857-8/25/$31.00 ©2025 IEEE | DOI: 10.1109/CE2CT64011.2025.10939412

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:15:41 UTC from IEEE Xplore. Restrictions apply.

such as accuracy, precision, recall, and F1-score.

Feature engineering plays a critical role in ML-based automated anxiety detection by converting raw physiological signals into meaningful features. These features are derived from domains like time and frequency to provide a comprehensive understanding of physiological changes. Nevertheless, high dimensionality can result in overfitting and increased computational complexity, making feature selection an essential step to improve performance and efficiency.

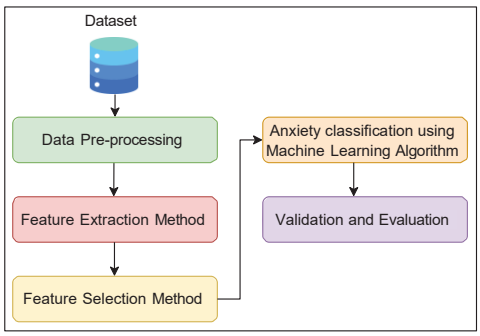


Fig. 1. Anxiety Classification Phases

Feature selection aims to identify the most informative features that contribute significantly to classification, improving model accuracy and reducing computational costs. By eliminating redundant and irrelevant features, the Feature Selection Algorithm (FSA) enhances generalizability and reduces overfitting. In recent years, numerous feature selection techniques, ranging from statistical methods like Pearson correlation and mutual information to advanced methods like wrapper-based and embedded approaches, have been explored to enhance the performance of anxiety classification models.

Despite the importance of feature selection in improving classification accuracy, to the best of our knowledge, no prior study has provided a comprehensive review of feature selection algorithms specifically for physiology-based anxiety classification. This review aims to bridge this gap by discussing the physiological signals commonly used in anxiety detection, the feature selection methods employed, and the challenges and future directions in this domain. Further, the study utilizes a spider-phobic anxiety dataset to extract ECG, EDA and RESP features. It proposes a novel Corr-FG hybrid FSA to remove irrelevant and redundant features for anxiety classification.

A. Search Process

This review focuses on advancements in feature selection techniques for anxiety classification using physiological signals, considering studies published between 2016 and 2024. Studies were gathered from several databases, including IEEE Explore, Elsevier, Springer, and Google Scholar. A total of 39 studies on physiology-based anxiety classification were identified, of which 20 specifically employed feature selection algorithms.

II. Feature Selection Techniques

Feature selection is one of the critical phases in anxiety classification methodology. It involves selecting the most informative subset of features that are crucial and valuable for classification algorithms. This process also effectively eliminates irrelevant data, reduces redundant information, and enhances learning accuracy.

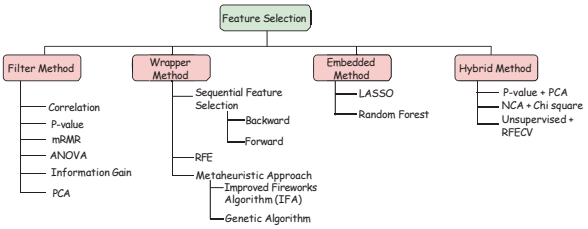


Fig. 2. Feature selection algorithms utilized in anxiety classification

This paper reviews the feature selection techniques applied in anxiety classification studies, categorizing them into three main types: filter, wrapper, and Embedded methods. Each technique serves a distinct purpose in selecting the most informative features. Table II summarizes various studies that utilized physiological signals and machine learning techniques for anxiety classification. Fig. 2 shows the list of distinct feature selection algorithms applied in anxiety classification. Further, Fig. 3 illustrates the distribution of major feature selection methods used across various physiological signals, such as EEG, ECG, EDA, EMG, and RESP, in the reviewed anxiety classification studies.

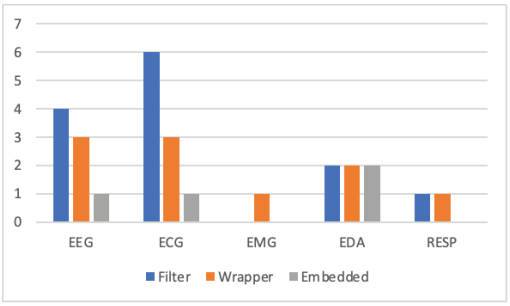


Fig. 3. The distribution of different major feature selection methods used across various physiological signals

A. Filter Methods

Filter methods assign a score to each feature and select the highest-ranked ones, operating independently of the classifier. Several studies have demonstrated the effectiveness of filter methods for anxiety classification using physiological signals. Li et al. [7] classified multiple anxiety levels using EEG, extracting 90 features and selecting 30 through correlation coefficients, resulting in 62.56% accuracy with an SVM classifier. Klados et al. [4] focused on EEG signals, extracting

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:15:41 UTC from IEEE Xplore. Restrictions apply.

TABLE I
PHYSIOLOGICAL SIGNALS STRENGTH AND WEAKNESS

Physiological Signal	Definition	Strengths	Weaknesses
EEG (Electroencephalography)	Measures electrical activity in the brain	High temporal resolution; captures changes in neural activity related to anxiety	Susceptible to noise and artefacts; requires careful setup and preprocessing
EDA (Electrodermal Activity)	Measures skin conductance, reflecting sweat gland activity linked to arousal	Non-invasive; highly responsive to sympathetic nervous system activity	Sensitive to external factors like temperature and movement
ECG (Electrocardiography - HR + HRV)	Records Heart Rate (HR) and Heart Rate Variability (HRV), indicating ANS activity	Well-studied markers of anxiety; Provides information on both sympathetic and parasympathetic activity	Can be affected by factors unrelated to anxiety (e.g., physical activity)
EMG (Electromyography)	Measures muscle activity, is often used to assess tension	Useful in identifying muscle tension	Prone to movement artefacts and require calibration for accurate results
RESP (Respiration)	Measures breathing patterns, including rate and depth	Provides insights into respiratory changes associated with anxiety, such as hyperventilation	Influenced by various external and physiological factors, reducing specificity to anxiety

466 features and employing CfsSubsetEval to achieve 93.75% accuracy with a Naive Bayes classifier. Nath et al. [21] developed an anxiety detection system for older adults using wristband sensors, selecting 24 features with P-value FSA and achieving 92% accuracy with Random Forest. Tripathy et al. [15] used single-channel ECG to detect anxiety stages, selecting 187 out of 200 features with ANOVA, resulting in 79.17% accuracy with a Decision Tree.

Lastly, Daneshmand et al. [16] developed an anxiety recognition approach using EEG signals, applying mRMR to select 15 out of 28 features, which resulted in 100

B. Wrapper Methods

Wrapper methods use optimization or search algorithms in combination with machine learning to identify the most informative features, incorporating feature selection directly within the learning process. Ihmig et al. [9] used Sequential Feature Selection (SFS) to retain six out of 25 features extracted from ECG, RESP, and EDA, achieving 89.8% accuracy with Bagged Trees. Similarly, Arsalan et al. [10] employed forward selection to choose five out of ten features for anxiety detection using EEG, achieving 76.92% accuracy with an MLP classifier.

Shikha et al. [1] extracted 504 features from EEG and applied RFECV to select 318 features, resulting in 83.5% accuracy using SSAE. Muhammad et al. [13] extracted 60 features from EEG and applied RFE to reduce them to 18, achieving 94.9% accuracy with Random Forest. Wen et al. [5] employed backward feature selection to reduce 11 heart rate features to four, achieving 90.5% accuracy with an SVM classifier.

Mo et al. [18] introduced a metaheuristic approach using the Improved Fireworks Algorithm for feature selection across multiple modalities, including heart rate, respiration rate, behavioural, audio, text, and questionnaire features. This method effectively selected 19 features, achieving 97.61% accuracy with the Adaptive Boosting (ADB) classifier.

In another study, Shikha et al. [20] utilized a Genetic Algorithm for feature selection in anxiety classification. The authors investigated different sliding window sizes and achieved the highest accuracy of 98.61% using Gradient Boosting (GB) with a 1-second sliding window.

C. Embedded methods

Embedded methods incorporate feature selection within the model training process, enabling the learning algorithm to determine which features contribute the most to predictive accuracy. Salkevicius et al. [6] used an embedded feature selection approach for anxiety recognition in a virtual reality therapy system. They extracted 33 features from multimodal physiological signals and reduced the set to 10 key features using Random Forest (RF), achieving 86.3

Lee et al. [8] developed a driver anxiety detection system using EEG, PPG, EDA, and pupil size signals. They initially extracted 445 features and employed LASSO to reduce the feature set to 31, achieving 77.01% accuracy using a logistic regression classifier with frontal EEG channels.

D. Hybrid Methods

Hybrid methods integrate multiple feature selection strategies, combining filter, wrapper, or embedded approaches to leverage the strengths of each. This combination enhances feature selection performance by mitigating the limitations inherent in individual techniques. Zheng et al. [3] proposed a wearable system utilizing PPG and EEG signals. They initially extracted 32 features and applied Pearson Correlation and PCA, ultimately achieving 62.5% accuracy with KNN. Similarly, George et al. [12] used a combination of P-value and PCA to reduce an initial set of 23 features to 11, achieving 82.2% accuracy with an SVM classifier. Baygin et al. [19] combined Neighborhood Component Analysis (NCA) with the Chi-square test for feature selection, achieving 99.32% accuracy with a KNN classifier using ECG signals. In another study, Vaz et al. [14] employed an unsupervised Feature Selection Algorithm (FSA) and Recursive Feature Elimination with Cross-Validation (RFECV) to reduce 109 features from ECG, EDA, and EMG to 34, achieving 71.1% accuracy with a Logistic Regression classifier.

Rodr´ıguez-Arce et al. [2] worked with multimodal physiological signals in an academic setting, applying CfsSubsetEval,

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:15:41 UTC from IEEE Xplore. Restrictions apply.

TABLE II
STUDIES UTILIZED FEATURE SELECTION FOR PHYSIOLOGY-BASED ANXIETY CLASSIFICATION USING MACHINE LEARNING.

Author	year	Subject	Signal	Origional Features	Feature Selection	Selected Features	Model	Accuracy (%)
Zheng et al. [3]	2016	20	EEG, PPG	32	Pearson Correlation, PCA	6	KNN, SVM	KNN: 62.5
Klados et al. [4]	2017	16	EEG	466	CfsSubsetEval function of WEKA	–	NB	NB: 93.75
Wen et al. [5]	2018	59	ECG	11	Backward Feature selection	4	SVM, NB, KNN, LDA, QDA	SVM - F1 score: 90.5
Salkevicius et al. [6]	2019	30	BVP, GSR, ST	33	RF	10	SVM	SVM: 86.3
Li et al. [7]	2019	12	EEG	90	correlation coefficients	30	SVM, KNN	SVM: 62.56
Lee et al. [8]	2020	31	EEG, PPG, EDA, Pupil size	445	LASSO	42	LR	LR: 77.01
Ihmig et al. [9]	2020	57	ECG, EDA, BR	25	Sequential FS	6	Bagged Trees	Bagged Trees: 89.8
Rodr´ıguez - Arce et al. [2]	2021	21	ST, GSR, HR, BR, SPo2	21	CfsSubsetEval and CorrelationAttributeEval) , (InfoGainAttributeEval	13	KNN, SVM, LR, RF	KNN: 95.56
Arsalan et al. [10]	2021	65	EEG	10	Wrapper method - forward selection approach	5	MLP, LR, RF	RF: 87.69
Nath et al. [11]	2021	41	EDA, BVP	47	P-value	24	RF	RF: 92
George et al. [12]	2021	52	ECG, RESP	23	P-value, PCA	11	SVM, KNN, DT, Rf, LR, Baging, Adabbost, Gradient boosting, Voting	LR: 82.83
Shikha et al. [1]	2021	23	EEG	504	RFECV	318	SSAE, DT, RF	83.93
Muhammad et al. [13]	2022	23	EEG	60	RFE	18	RF, MLP, SVM, KNN, DT	RF: 94.9
Vaz et al. [14]	2023	15	ECG, EDA, EMG	109	Unsupervised FSA, RFECV	34	LR, LDA, DT, SVM, ADB, RF, XGB	XGB: 91.6
Tripathy et al. [15]	2023	19	ECG	200	ANOVA	187	DT, LGBM, XGBoost, Rf, ERT	ERT: 91.92
Daneshmand et al. [16]	2023	71	ECG, EDA, RESP, and temperature	28	mRMR	15	KNN	KNN: 100
Mohan et al. [17]	2023	23	EEG	270	mRMR	–	Deep CNN, ANN, KNN, LSTM, RNN	Deep CNN: 97.6
Mo et al. [18]	2024	227	PPG, Behaviour audio	–	Improved fireworks algorithm	19	ADB, DT, Rf, LR, KNN, SVM, MLP	ADB: 97.61
Baygin et al. [19]	2024	19	ECG	–	NCA, Chi2	–	KNN+NCA, SVM+Chi2	SVM+Chi2: 99.94
Shikha et al. [20]	2024	299	EEG	–	Genetic Algorithm	–	DT, KNN, XGBoost, GB	GB: 98.61

CorrelationAttributeEval, and InfoGainAttributeEval to select 13 out of 21 features, which led to 95.56% accuracy with KNN.

III. CHALLENGES IN FEATURE SELECTION FOR ANXIETY CLASSIFICATION

Feature selection for anxiety classification using physiological signals involves numerous challenges, particularly due to the complexity of physiological data and the variability in individual responses. Some of the key challenges are as follows:

1) *Scalability Issues*: The high dimensionality of physiological data, such as multi-channel EEG or high-resolution HRV features, poses scalability challenges for feature selection, particularly for wrapper methods. These methods struggle to efficiently handle large datasets, which affects their utility for large-scale anxiety studies.
2) *Data Noise Sensitivity*: Physiological signals are inherently noisy due to factors such as movement artefacts, environmental influences, and individual variability. This noise complicates the feature selection process, making identifying consistent and reliable features for anxiety classification difficult. It is particularly problematic for anxiety detection, where subtle physiological changes are critical indicators.
3) *Generalizability*: Selected features may not always generalize well across diverse populations, especially when physiological signals vary significantly between individuals. Anxiety manifests differently among people due to various factors such as age, gender, and comorbidities, which complicates the selection of universally relevant

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:15:41 UTC from IEEE Xplore. Restrictions apply.

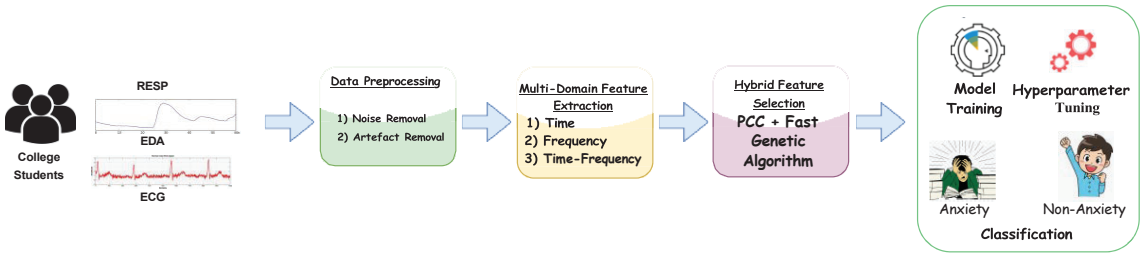


Fig. 4. Proposed anxiety classification framework using hybrid feature selection algorithm.

features.

4) *Capturing Feature Interactions*: Filter methods, while computationally efficient, cannot adequately capture intricate interactions between features, which may limit their effectiveness in anxiety classification. Physiological signals like HRV and EDA often exhibit complex relationships, and missing these interactions could lead to reduced model accuracy and reliability in detecting anxiety states.

IV. Future Directions and Solutions

Future research in physiology-based anxiety classification should address several key challenges and explore innovative directions to improve the robustness and reliability of anxiety detection models:

1) *Metaheuristic Feature Selection Algorithms:* Future research could explore the use of metaheuristic optimization techniques, such as Genetic Algorithms (GA), Particle Swarm Optimization (PSO), and Differential Evolution, to enhance feature selection. These algorithms can help navigate the complex feature space of physiological data to find optimal feature subsets while balancing accuracy and computational costs.
2) *Reinforcement Learning for Feature Selection:* Reinforcement Learning (RL) could be employed as a feature selection agent to select the most informative features in real-time adaptively. RL can improve the selection process by learning from the data and optimizing the feature subset based on classification performance, which would be especially useful in dynamic anxiety detection environments.
3) *Improving Generalizability Across Populations:* More research is needed to ensure that selected features generalize well across different populations. Various techniques, such as transfer learning and domain adaptation, could be explored to improve the model's ability to handle variability in physiological signals among diverse demographics, making the models more applicable in real-world settings.
4) *Multimodal Fusion for Robust Anxiety Detection:* Future research should improve multimodal data fusion techniques to leverage the complementary information provided by multiple physiological signals, such as EEG, EDA, and HRV. Advanced fusion methods, such as deep learning-based fusion or multi-view learning, could enhance model robustness and accuracy.
5) *Real-time Implementation and Computational Efficiency:* Developing adaptive and lightweight feature selection methods that operate in real-time is crucial for practical applications, particularly for wearable devices used in anxiety monitoring.
6) *Addressing Data Noise and Improving Signal Quality:* Future work should also focus on developing more effective preprocessing techniques to mitigate noise and artefacts in physiological signals. This could include using adaptive filtering methods or leveraging advanced signal processing techniques to improve the quality of features extracted from noisy data.

V. Proposed Architecture: Anxiety Classification based on ECG, EDA and RESP physiological signals and Machine Learning

Previous studies [2] have shown that hybrid Feature Selection Algorithms and metaheuristic algorithms [18] [20] often achieve the highest accuracies. Therefore, this review presents a framework for anxiety classification using a novel corr-FG hybrid FSA. The study utilizes a publicly available spider-phobic dataset [9], which includes physiological signals such as ECG, EDA, and RESP. Fig. 4 shows the proposed framework approach, which involves distinct phases:

1) *Data Preprocessing:* The study applies a Savitzky-Golay filter with a 40-point window and a 400 ms sigma to smooth the EDA signal and eliminate noise. For the respiration signal, the NeuroKit2 tool is utilized for cleaning and analysis.
2) *Feature Extraction:* The study extracts 31 ECG-derive Heart Rate Variability (HRV) features by identifying correct RR intervals, which represent the time between consecutive heartbeats. It also extracts 19 phasic and tonic EDA features and 11 RESP features.
3) *Feature Selection:* Feature selection is a crucial step in the anxiety classification methodology. The multidomain nature of physiological signals often results in high-dimensional feature sets. Therefore, this study proposes a hybrid feature selection algorithm. Initially it applies Correlation analysis to remove highly correlated features

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:15:41 UTC from IEEE Xplore. Restrictions apply.

that do not contribute additional value. Subsequently, a Fast Genetic Algorithm (Fast GA) is utilized to further refine the feature selection by identifying the subset of features that effectively enhance classification performance. The Fast GA is optimized for quick convergence by using a reduced population size of 20 and 30 generations, which are selected based on preliminary experimentation to balance computational efficiency and accuracy.

4) *Anxiety Classification:* The selected features are input into machine learning models, including Random Forest (RF), Support Vector Machine (SVM), K Nearest Neighbour (KNN) and Decision Tree (DT). Additionally, *Grid_search* is utilized to improve the model performance further.

TABLE III
PERFORMANCE COMPARISON OF DISTINCT CLASSIFICATION ALGORITHM AFTER CORR-FG HYBRID FSA

Classifier	Accuracy (Before FSA)	F1-Score (Before FSA)	Accuracy (Corr-FG FSA)	F1-Score (Corr-FG FSA)
RF	83.53	81.48	88.93	87.01
SVM	70.18	71.85	79.55	78.68
KNN	75.78	78.35	75.37	75.36

Table III presents the performance of various classifiers using the selected features. The proposed corr-FG hybrid FSA identifies 8 out of 63 features, representing the most influential attributes for effective anxiety detection. In the classification algorithm, the RF classifier shows 5.4% improvement in accuracy, while SVM shows a 9.37% increase in accuracy after feature selection. Furthermore, KNN obtained a comparable accuracy of 75.37% using 8 features.

VI. CONCLUSION

This review highlights the significance of feature selection in physiological-based anxiety classification. Given physiological data's complexity and high dimensionality, selecting the most informative features is crucial to improving model accuracy and reducing computational load. Analyzing filter, wrapper, embedded, and hybrid approaches reveals their advantages and limitations in enhancing anxiety classification performance. Further, it proposes a novel Corr-FG hybrid FSA for anxiety classification. Future research should focus on developing adaptive and computationally efficient feature selection techniques, which can provide more robust solutions for anxiety detection. This work aims to guide researchers in advancing feature selection methodologies for more effective anxiety detection systems.

REFERENCES

[1] S. Shikha, M. Agrawal *et al.*, "Stacked Sparse Autoencoder and Machine Learning Based Anxiety Classification Using EEG Signals," in *Proceedings of the First International Conference on AI-ML-Systems*, 2021, pp. 1–7.

[2] J. Rodr´ıguez-Arce, L. Lara-Flores *et al.*, "Towards an anxiety and stress recognition system for academic environments based on physiological features," *Computer methods and programs in biomedicine*, vol. 190, p. 105408, 2020.

[3] Y. Zheng, T. C. Wong *et al.*, "Unobtrusive and multimodal wearable sensing to quantify anxiety," *IEEE Sensors Journal*, vol. 16, no. 10, pp. 3689–3696, 2016.

[4] M. A. Klados, N. Pandria, A. Athanasiou, and P. D. Bamidis, "An automatic EEG based system for the recognition of math anxiety," in *IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE, 2017, pp. 409–412.

[5] W. Wen, G. Liu, Z.-H. Mao, W. Huang, X. Zhang, H. Hu, J. Yang, and W. Jia, "Toward constructing a real-time social anxiety evaluation system: Exploring effective heart rate features," *IEEE transactions on affective computing*, vol. 11, no. 1, pp. 100–110, 2018.

[6] J. S˘alkevicius, R. Damas˘evic˘ius, R. Maskeliunas, and I. Laukiene˙, "Anxiety level recognition for virtual reality therapy system using physiological signals," *Electronics*, vol. 8, no. 9, p. 1039, 2019.

[7] Z. Li, X. Wu, X. Xu, H. Wang, Z. Guo, Z. Zhan, and L. Yao, "The recognition of multiple anxiety levels based on electroencephalograph," *IEEE Transactions on Affective Computing*, 2019.

[8] S. Lee, T. Lee *et al.*, "Detection of drivers' anxiety invoked by driving situations using multimodal biosignals," *Processes*, vol. 8, no. 2, p. 155, 2020.

[9] F. R. Ihmig, F. Neurohr-Parakenings *et al.*, "On-line anxiety level detection from biosignals: Machine learning based on a randomized controlled trial with spider-fearful individuals," *Plos one*, vol. 15, no. 6, p. e0231517, 2020.

[10] A. Arsalan *et al.*, "A study on multi-class anxiety detection using wearable EEG headband," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–11, 2021.

[11] R. K. Nath *et al.*, "Machine learning based solutions for real-time stress monitoring," *IEEE Consumer Electronics Magazine*, vol. 9, no. 5, pp. 34–41, 2020.

[12] A. S. George, A. V. Kurup *et al.*, "Predicting autonomic dysfunction in anxiety disorder from ECG and respiratory signals using machine learning models," *International Journal of Online & Biomedical Engineering*, vol. 17, no. 7, 2021.

[13] F. Muhammad *et al.*, "Human state anxiety classification framework using EEG signals in response to exposure therapy," *Plos one*, vol. 17, no. 3, p. e0265679, 2022.

[14] M. Vaz, T. Summavielle, R. Sebastia˜o, and R. P. Ribeiro, "Multimodal classification of anxiety based on physiological signals," *Applied Sciences*, vol. 13, no. 11, p. 6368, 2023.

[15] R. K. Tripathy *et al.*, "Detection of Different Stages of Anxiety from Single channel Wearable ECG Senor Signal using Fourier-Bessel Domain Adaptive Wavelet Transform," *IEEE Sensors Letters*, 2023.

[16] F. Daneshmand-Bahman and A. Goshvarpour, "Anxiety recognition using a new eeg signal analysis approach based on sample density in a chebyshev chaotic map," in *Affective Computing in Healthcare: Applications based on biosignals and artificial intelligence*. IOP Publishing Bristol, UK, 2023, pp. 1–1.

[17] R. Mohan and S. Perumal, "Classification and detection of cognitive disorders like depression and anxiety utilizing deep convolutional neural network (CNN) centered on EEG signal." *Traitement Du Signal*, vol. 40, no. 3, 2023.

[18] H. Mo, S. C. Hui, X. Liao, Y. Li, W. Zhang, and S. Ding, "A multimodal data-driven framework for anxiety screening," *IEEE Transactions on Instrumentation and Measurement*, 2024.

[19] M. Baygin, P. D. Barua, S. Dogan, T. Tuncer, T. J. Hong, S. March, R.-S. Tan, F. Molinari, and U. R. Acharya, "Automated anxiety detection using probabilistic binary pattern with ECG signals," *Computer Methods and Programs in Biomedicine*, vol. 247, p. 108076, 2024.

[20] S. Shikha, D. Sethia, and S. Indu, "Investigating the effect of sliding window length for EEG-based anxiety classification," in *2024 Asia Pacific Conference on Innovation in Technology (APCIT)*. IEEE, 2024, pp. 1–6.

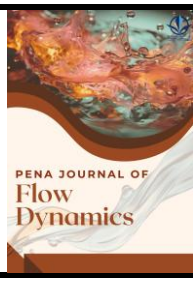
[21] R. K. Nath *et al.*, "Machine learning-based anxiety detection in older adults using wristband sensors and context feature," *SN Computer Science*, vol. 2, no. 5, pp. 1–12, 2021.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:15:41 UTC from IEEE Xplore. Restrictions apply.

PENA CENDEKIA PUBLISHING
NURTURING IDEAS, ENRICHING KNOWLEDGE

Pena Journal of Flow Dynamics

Journal homepage:
https://penacendekia.com.my/index.php/pjfd/
ISSN: XXXX-XXXX



A Study on the Viscosity of Viscoelastic Magnetic Abrasive Gel under Different Volumetric Proportions

Krovvidi Srinivas[1,*]

[1] Department of Mechanical Engineering, Delhi Technological University, Shabad Daulatpur, Main Bawana Road, Delhi, India

ARTICLE INFO

Article history:
Received 16 December 2024
Received in revised form 16 January 2025
Accepted 20 February 2025
Available online 31 March 2025

Keywords:
Viscoelastic medium; magnetic abrasive finishing; Nano finishing

ABSTRACT

In the present study, the Viscosity of the Viscoelastic Gel used in Magnetic Abrasive finishing is analysed based on the test conducted on the Compact Rheometer. Silicon-based Transformer oil is used as the base oil and Aluminium Stearate is used for increasing the abrasive holding capacity under very high magnetic field strength. The media prepared is commonly used for performing Nano-finishing of various metals' internal and external surfaces. The Viscosity of AP3 Grease, transformer oil and silicon oil is computed and compared using Rheometer, and the sample of viscoelastic Magnetic Abrasive media is prepared using these base oil. Six samples have been tested and the results have been analysed. The Media selected for the Viscoelastic magnetic abrasive finishing is tested after 30 days of preparation and compared its viscosity with the freshly prepared medium. The result highlighted the successful application of transformer oil and Silicon oil to prepare gel for the viscoelastic Magnetic Abrasive Finishing process.

1. Introduction

The importance of computational fluid flow has shaped the dynamics of the modern flow system and improved products and manufacturing quality. A recent example is the MAF and AFM process in which the machining operations occur due to the flow of Media inside the constrictive passage. The Media used in the Magnetic Abrasive flow Finishing is non-newtonian. The Media is a mixture of Abrasive, Gel and polymer, due to which viscosity plays a significant role in the machining operations. Many scholars have worked on the non-Newtonian fluid and contributed to the development and analysis of the non-Newtonian fluid in various aspects. Various research has been carried out for the flow governing peristaltic movement in porous walls, which acts as a reference for the non-newtonian fluid flow [1-6].

The viscoelastic Magnetic Abrasive finishing process is one of the leading solutions in the field of the nano finishing process. It consists of carbonyl particles along with abrasive and viscoelastic putty, allowing the viscoelastic putty to act along the line of the magnet. The carbonyl particles present in

* *Corresponding author.*
E-mail address: srinivaskrovvidi@dtu.ac.in

https://doi.org/10.37934/pjfd.1.1.4152

the viscoelastic putty get attracted toward the magnet. As a result, a magnetic brush is formed. The machining process occurs when this brush is rotated over the surface of the workpiece for effective machining.

A similar process may be applied when machining the internal surface of the hollow workpiece. Here the viscoelastic magnetic abrasive Media is allowed to reciprocate within the boundary of the workpiece, where the workpiece is placed co-axially along the cylindrical workpiece and the magnet is placed around it. The carbonyl particle in the media would be attracted towards the wall of the workpiece due to the magnetic field. Thus, effective machining occurs due to magnetic forces and the finishing process.

Many researchers modelled the viscoelastic Magnetic Abrasive Finishing process and studied the media's flow parameters. The researchers study the effect of Extrusion pressure, no of cycles, Abrasive mesh size, Media viscosity, abrasive, Type of Abrasive particles, Size of carbonyl particles, and Base oil for viscoelastic media and gel [7-11]. The concept of preparation of viscoelastic fluid is confined to the viscoelastic magnetic abrasive finishing and abrasive flow machining. The primary difference between the media of Abrasive Flow Machining and viscoelastic Magnetic abrasive finishing is that the media of AFM does not contain carbonyl particles, but the synthesis process remains the same [12-15].

2. Methodology

In this article, a brief discussion will be made about the linear Viscoelastic Models. Subsequently, a Viscoelastic Magnetic abrasive Medium preparation will be discussed. In total, 7 Viscoelastic Magnetic Abrasive Media have been prepared and their Viscosity has been tested using Modular Compact Rheometer (MCR 102). The graphical Results have shown the shear stress vs shear strain and shear stress vs viscosity and are also validated [16]. Based on this, a suitable Viscoelastic Magnetic Abrasive Medium has been proposed for the internal finishing of the Complex surface and the external surfaces. In the present research, the abrasive medium prepared is known as Viscoelastic Magnetic Abrasive Medium, specially prepared for performing both internal and external finishing of metallic specimens. The main problem found during the literature review is that the medium used for Magnetic abrasive finishing experiences sedimentation of the heavier particles [17,18].

In Magnetic Assisted finishing process, the Carbonyl Iron particle's density is higher than the Abrasives used. As the CIP particle settles down before and after exposure to the magnetic field, it becomes difficult to form the CIP chain under the Magnetic field. So, a Viscoelastic medium that addresses the sedimentation is prepared and experiments have been performed to find the viscosity before choosing a suitable medium. Materials with Viscoelastic properties exhibit both Viscous behaviour and Elastic behaviour. Most of the materials exhibit Viscoelasticity, both Natural and artificial. However, the dominance of Viscosity or Elasticity depends upon the individual material prepared and the operating conditions. All the Rheological materials will exhibit Viscoelastic properties. When a material is under Static loading, the Elastic behaviour dominates, whereas the Viscous behaviour dominates with Dynamic loading.

2.1 Viscosity Measurement

The Viscosity and Rheology measurement could be done using three different equipment pieces. Each piece of equipment has a specific purpose and would be used depending upon the type of loading. The types of equipment used for measuring Viscosity/Rheology are:

a) Rotational Viscometer: Used for Steady loading conditions. The probes used for this purpose are Cylindrical, Flat plates having Circular cross-sections and Conic sections. For measuring fluids with Very high viscosity, the flat plate Viscometer is usually preferred.
b) Oscillatory Rheometer (Viscometer): Used for finding the viscosity when the force Associated with it is Oscillatory or bending force.
c) Transient Condition Rheology tester: This equipment is the advanced version of the SteadyState Viscometer, in which the variation in the deformation of the specimen with time under impulsive loading would be measured.

In the present research work, Modular Compact Rheometer (MCR 102), a RotationalViscometer, is used to find the Viscosity of the Viscoelastic Magnetic Abrasive Finishing Medium, especially prepared.

The critical part associated with the Rheometer is as follows:

a) Measuring Head: Holds the rotating probe known as PP 50 measuring system.
b) MICR coupling: A part used for holding the rotating probe.
c) Flange ring: Holds the Temperature device holder.
d) Colour Display: Displays the temperature, position in mm, Normal force and status.
e) Soft keys: Used for initiating different operations of the instrument.
f) Adjustable feet: For making the platform plane.

Figure 1 shows the schematic diagram of the modular compact Rheometer and Figure 2 shows the experimental setup. Figure 2(i) shows the actual experimental setup, while Figures 2 (ii), Figure (iii) and Figure (iv) show the different types of measuring probes used for the measurement of viscosity. Plate-Plate filling and plate cone filling are used for the present work.

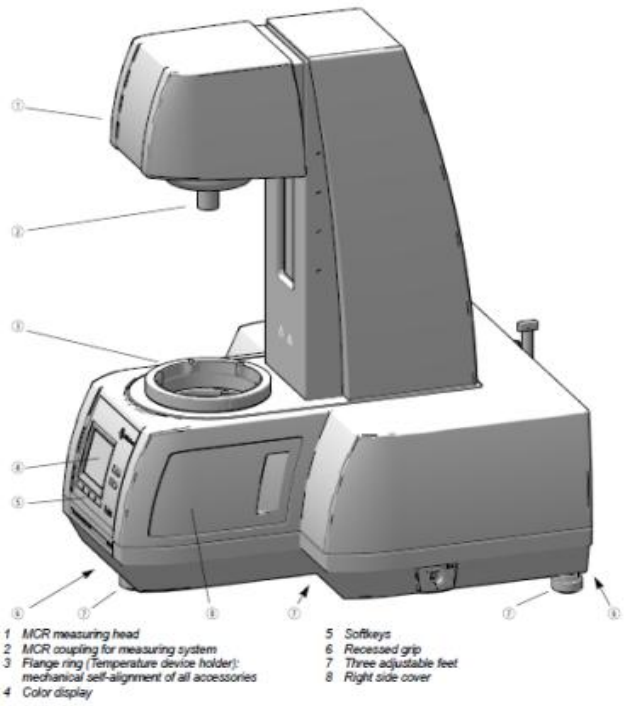


Fig. 1. Modular compact rheometer with important parts



Figure (i)
Modular Compact Rheometer

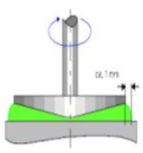


Figure (ii)
Cone-Plate
Filling

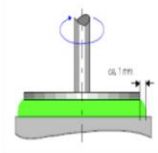


Figure (iii)
Plate –Plate
Filling

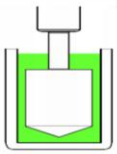


Figure (iv)
Concentric
Cylinder Filling

Fig. 2. Experimental setup (i) modular compact rheometer (ii) Cone Plate Filling (iii) Plate-Plate filling (iv) Concentric Cylinder Filling

2.2 Viscoelastic Magnetic Abrasive Medium Preparation

In the present work, Sevensamples have been prepared, and their viscosity has been measured using MCR 102. Initially, the viscosity of the Rheological fluids used in preparing the samples was measured and based on the experimental results, the base fluid for preparing the ViscoelasticMagnetic abrasive finishing (VEMAF) Medium has been decided. Subsequently, adding the required ingredients has taken place, producing the required Viscoelasticmedium [19-21]. By adding the mixture of Carbonyl Iron Particle (CIP) of CS grade and the Silicon Carbide abrasive in predefined proportion with the Viscoelastic Medium, the requiredViscoelastic Magnetic abrasive Medium has been prepared.

Three types of Viscousfluids and AP3 grease have been considered for preparing the samples, and they are as follows:

a) Transformer Oil (white)
b) Transformer Oil (Red)
c) Silicone Oil
d) AP3 Grease

While performing the viscosity measurement experiment for the base fluids mentioned at 1 to 3 above, a Plate-Plate filling type probe was used and for the AP3 Grease Cone-Plate A filling-type probe has been used. The sample has been placed below the probe, which is initially compressed and then the probe starts rotating, which in turn induces shear stress and shear strain on the sample. The sample quantity depends upon the gap between the probe and the platform. The experimental results are shown below.

From the Graphs for Viscosity Vs Shear strain and Shear Stress vs ShearStrain for White Transformer oil, it may be noted that after initial disturbance, viscosity is almost constant and obtained an average value of 12.183 mPa-s and the Shear stress vs Shearstrain is maintaining a constant slope from point 11 in the table.

From the tabulated data and the Graphs for Viscosity Vs Shear strain and Shear Stress vs ShearStrain for Red Transformer oil, it may be noted that viscosity drops to a lower value for a strain rate of 1.2 after that, viscosity is almost constant and obtained an average value of 20.466mPa-s and the Shear stress vs Shear strain is maintaining a constant slope from the point 11 in the graph.

From the tabulated data for Silicone oil, it may be noted that the viscosity of the Silicone oil is almost constant within the experimental range and has a magnitude of 236.2 mPa-s. Also, the Shear Stress-Shear Strain curve is almost linear. Based on the results obtained above, Silicone oil and Red transformer oil have been chosen as the base Rheological fluids for making the Viscoelastic Magnetic Abrasive Medium.

Figure 3 shows the viscosity measurement results for the transformer oil red, Transformer oil White and silicone oil. The graphs were made based on tabulated data obtained from the modular compact Rheometer.

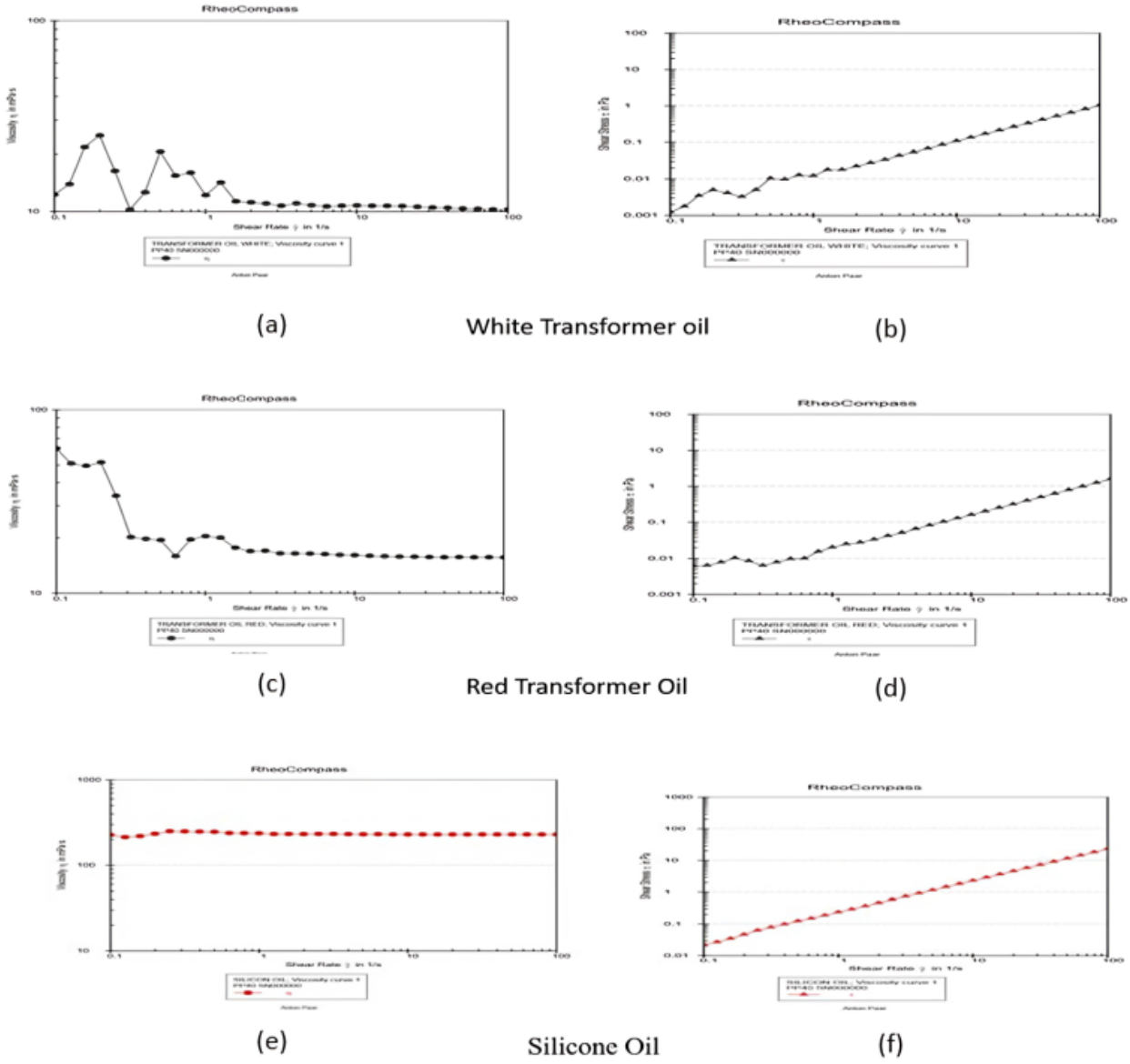


Fig. 3. Viscosity measurement for Transformer oil and the silicon oil (a) viscosity vs shear strain for white transformer oil (b) shear stress vs shear strain for white transformer oil (c) Viscosity VS shear strain for red transformer oil (d) Shear stress vs shear strain for red transformer oil (e) viscosity vs shear strain for silicon oil (f) Shear stress vs Shear strain for the silicone oil

2.3 Preparation of Silicone Oil-Based Vemaf Medium

The preparation has got the following steps.

2.3.1 Preparation of polymer

Mixing 1000 ml Silicone oil with 50 grams of Boric acid thoroughly and heating from room temperature up to 60° C makes the required polymer which would be white. Here, the cross-linking of Silicone oil's chained structure with Boric acid occurs. Figure 4(i) below shows a schematic representation of Cross-linking. Subsequently, 10 Lewis acid grams were added at 40° C of temperature, followed by adding ammonium Carbonate and heating up to boiling temperature along with Continuous stirring, bringing the final stage to the polymer, which is having pale yellow colour shown in Figure 4(ii) below. Cooling takes place in the open atmosphere. The prepared polymer will behave as a Viscous medium in the Viscoelastic Magnetic Abrasive Finishing Medium and has a marginal reduction in viscosity compared to the Silicon oil Figure 4(iv).

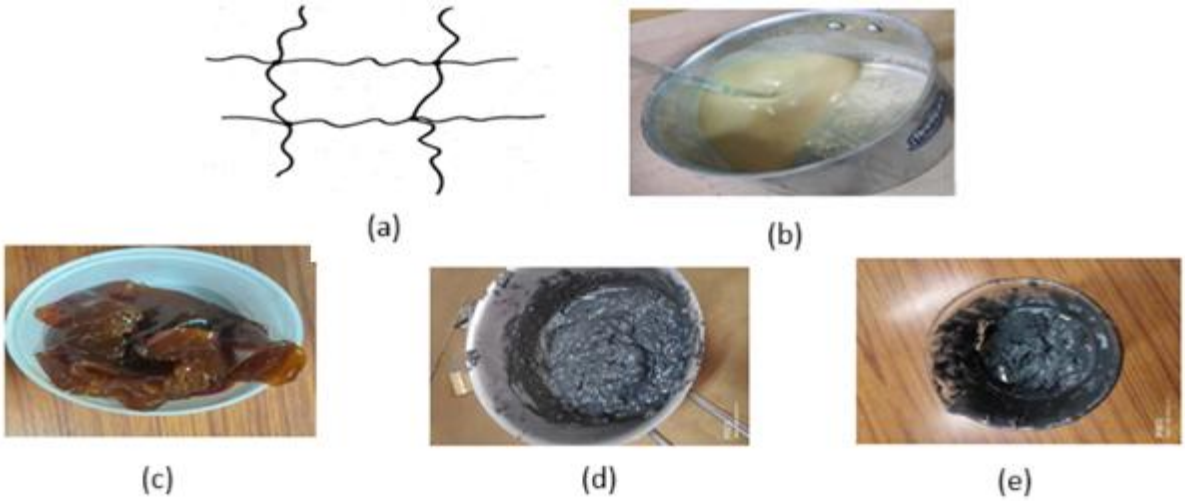


Fig. 4. (a) Cross-linking polymers (b) Prepared Polymer (c) gel (d) Silicone oil-based Magnetic Abrasive Media (e) Transformer oil-based viscoelastic magnetic abrasive based Media

2.3.2 Preparation of gel

The gel is prepared with the chemical reaction of the hydrocarbon oil and the aluminium stearate. 500 ml of the Hydrocarbon oil and 25 grams of Aluminium Stearate are mixed thoroughly and heated up to 120 degrees Celsius. The Product obtained is called the gel, having red colour and very high Viscosity [21-23].

2.3.3 Mixing of carbonyl iron particle and silicon carbide

Carbonyls Iron Particles (CIP) of CS grade and Silicon Carbide(SiC) of Abrasive Mesh Numbers 320, 400, 600 and 800 have prepared the mixture. The density of CIP is approximately 7.2 grams per cm3 and the Sic powder from the manufacturers' catalogue has a density of 2.6 grams per cm3. Carbonyl particles and Silicon Carbide have been taken in a ratio of 3:1 by weight. Separately, the mixture has been prepared for each abrasive Size. A ball mill has been used to mix CIP and the abrasive powder thoroughly.

2.3.4 Final mixing

Here, the polymer and the gel have been mixed in a ratio of 8:1 by weight Thoroughly by Gel mixing machine. Now the Polymer-Gel medium hence prepared, is mixed with CIP- Abrasive mixture in a ratio of 2:1 by weight and finally, the viscosity of the Viscoelastic Magnetic Abrasive medium has been found. This medium is used for the Internal finishing of the Spline shafts.

2.4 Preparation of Transformer Oil (Red)-Based Viscoelastic Magnetic Abrasive Medium

The following steps are followed in sequence for making the Viscoelastic Magnetic abrasive Medium.

a) 1000 ml Red transformer oil is heated till the moisture is removed. 500 ml Transformer oil is heated up to 50° C, and 50 grams of Sodium Stearate is added and thoroughly stirred until a thorough mixing occurs. This will make the medium thick compared to its initial condition.
b) Further heating up to around 150° C will saponify the mixture and it becomes further thick. The water content left is removed.
c) The remaining 500 ml transformer oil is heated up to 40O C, and Simultaneously 50 ml boric acid solution and 50 ml Sodium hydroxide solution are added and thoroughly stirred. This converts the transformer oil to highly viscous.
d) Viscous oil, both the saponified semisolid and the viscous transformer oil, were mixed thoroughly at 50° C and stirred for 60 minutes. An antioxidant solution of 1% concentration is added. This gives Viscoelastic medium.
e) Now the Carbonyl Iron Particle and SiC mixture in a 3:1 ratio by weight have been mixed Thoroughly in the ball mill.
f) Now the Viscoelastic medium and the CIP and SiC powder in 1:3 ratios by weight have been mixed and thorough stirring takes place by the Chemical stirrer. This makes the Viscoelastic Magnetic Abrasive Medium.

The Viscoelastic Magnetic Abrasive Medium prepared based on Transformer oil (Red colour) has a Viscosity of 2926.2 mPa-s (Milli Pascal-Second). This lies in the viscosity of semisolids. Also, Shear-thinning is observed as a decrease in Viscosity with Shear strain.

2.5 Preparation of Viscoelastic Magnetic Abrasive Medium Based On AP3 Grease

The Carbonyl Iron Particle and Silicone Carbide abrasive mixture have been added to the AP3 grease, and a semisolid is prepared. The viscosity of each sample is found.

2.5.1 AP3 grease-based viscoelastic magnetic abrasive medium lab sample 5

In this sample, 30 grams of AP3 Grease and 10 grams of CIP and SiC mixture is thoroughly Stirred, and the sample has been tested for viscosity using MCR 102. The viscosity obtained is 1640mPa-s. The results are given below. Graph.3.17 and Graph.3.18 show the relation between Viscosity Versus Shear strain and Shear stress versus Shear Strain.

2.5.2 AP3 grease-based viscoelastic magnetic abrasive medium lab sample 6

In this sample, 30 grams of AP3 grease and 15 grams of CIP- SiC mixture is thoroughly mixed and the sample has been tested for viscosity. The graphs show a relation between Viscosity Versus Shear strain and Shear stress versus Shear Strain.

2.5.3 AP3 grease-based viscoelastic magnetic abrasive medium lab sample 7

In this sample, 30 grams of AP3 grease and 20 grams of CIP- SiC mixture is thoroughly mixed, and the sample has been tested for viscosity.

2.5.4 AP3 grease-based viscoelastic magnetic abrasive medium lab Sample 8

In this sample, 30 grams of AP3 grease and 30 grams of CIP- SiC mixture is thoroughly mixed and the sample has been tested for viscosity.

3. Results and Discussion

All the prepared samples were tested under the Compact Rheometer and their Viscosity vs strain rate curves were obtained. Also, Graphs were plotted for the shear stress vs strain rate. Figure 5 shows the obtained graph for various samples. The Viscoelastic Magnetic Abrasive medium is 30 days old viscoelastic magnetic abrasive media, and the graph of viscosity vs strain rate for polymer and gel is also plotted in the following Figure (Figure 6). The results are shown in the graph.

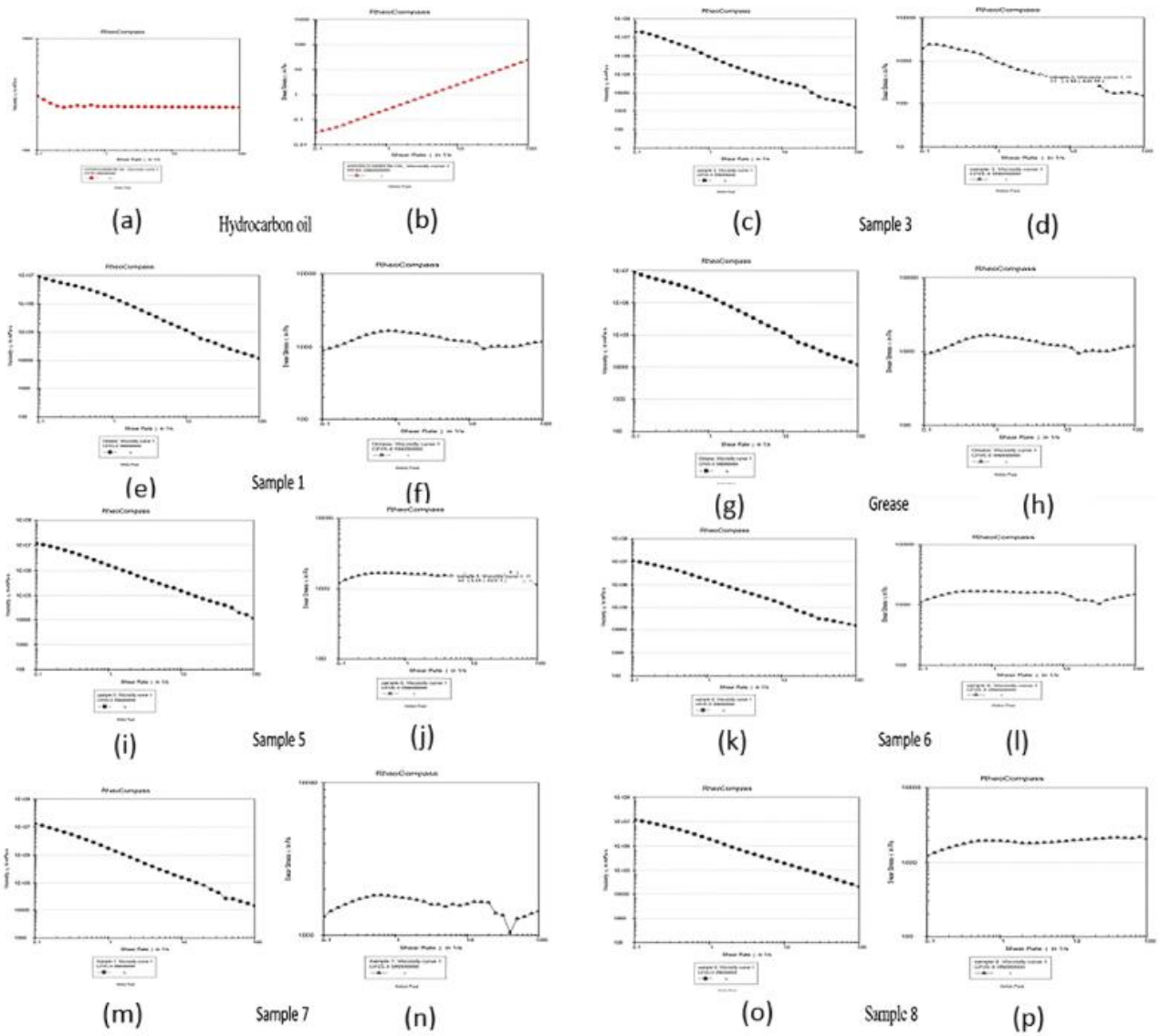


Fig. 5 Viscosity vs shear strain and shear stress vs shear strain graph for hydrocarbon oil, Sample 3, Sample 1, Grease, sample 5, Sample 6, sample 7 and sample 8. Fig 5 (a) Viscocity vs Strain rate hydrocarbon oil (b) shear stress vs strain rate (c) viscosity vs strain rate sample 3 (d) shear stress vs strain rate for sample 3 (e) viscosity vs strain rate sample 1(f) shear stress vs strain rate sample 1 (g) Viscocity vs Strain rate grease (h) shear stress vs strain rate for grease (i) viscosity vs strain rate sample 5 (j) shear stress vs strain rate sample 5 (k) viscosity vs strain rates sample 6 (l) shear stress vs strain rate sample 6 (m) viscosity vs strain rate sample 7 (n) Shear stress vs strain rate Sample 7 (o) viscosity vs strain rate Sample 8 (p)shear stress vs strain rate sample 8

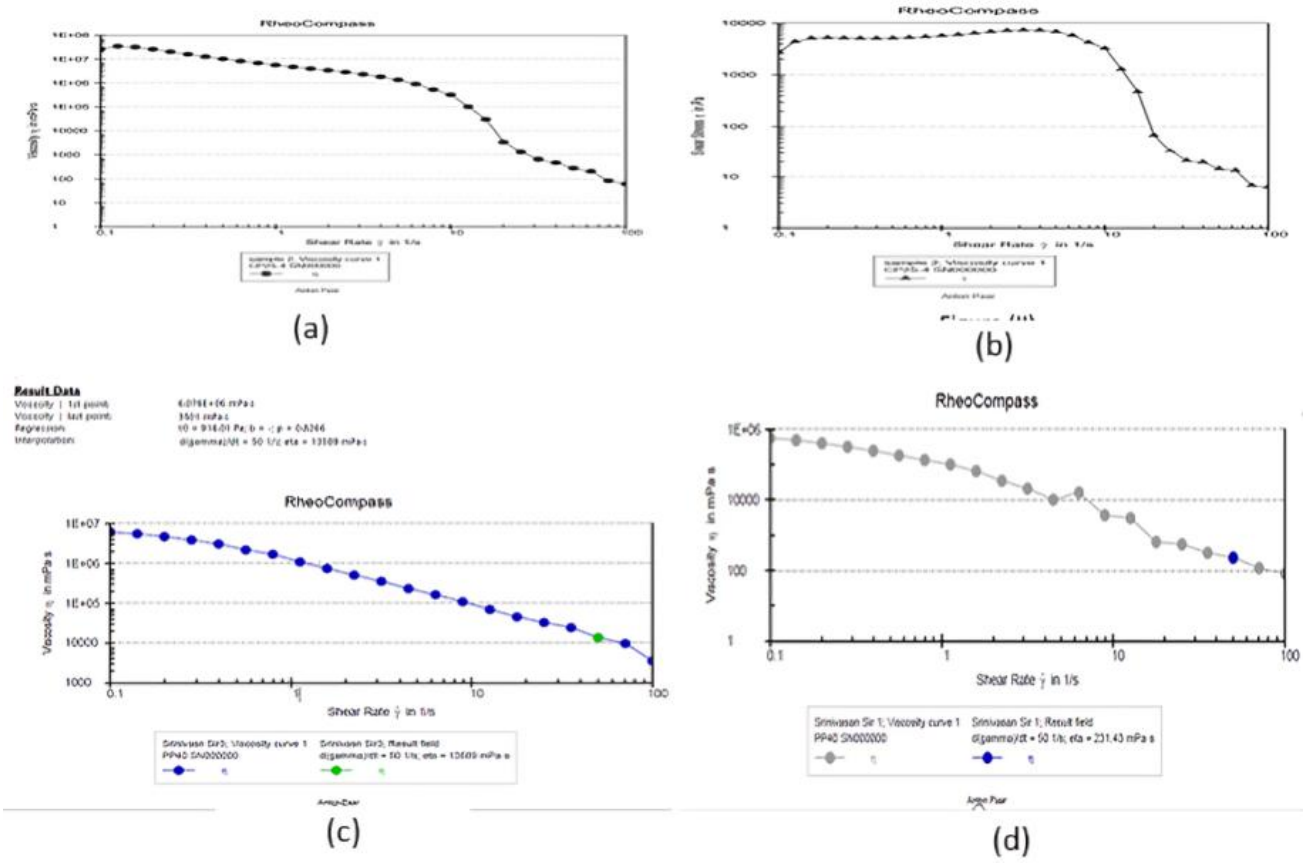


Fig. 6 Viscocity vs strain rate for 30 days old sample for VMAF Media, Gel and polymer 6 (i) Viscocity vs strain rate for 30days old VMAF sample (ii) Shear stress vs strain rate for a 30-day old sample of VMAF (iii) Viscocity vs Strain rate for gel (iv) Viscocity vs strain rate for polymers

4. Conclusions

In the present research, an attempt was made to prepare a gel for the VMAF process using different types of hydrocarbon oil as the base material. The researcher also computed the viscosity of transformer oil (Red and White), Grease, Hydrocarbon oil and silicon oil wittth the help of a Rheometer and predicted their application for the Viscoelastic Magnetic Abrasive Finishing Process. The researchers concluded the following points:

a) A total of six Viscoelastic Magnetic abrasive media have been prepared and their viscosity tested.
b) The Viscosity of the AP3 Grease-based Viscoelastic Magnetic Abrasive Media has shown an increase in the viscosity with the increase in the proportion of CIP-Abrasive Powder. However, with the increased temperatures during the finishing operations, there is a higher chance of reduction in the viscosity.
c) White transformer oil could not be used due to its instability during the initial loading period.
d) Silicone oil-based Viscoelastic Magnetic Abrasive Medium is preferred for internal Finishing. The sample also showed flow ability.
e) The viscosity of the Silicone oil has drastically increased from 954.06 mPa-s to 5688 mPas in 30 days from the date of preparation, and at this higher viscosity, this medium can't be used for finishing operation.

The viscosity of the Media is the major parameter for machining the workpiece, as the viscosity holds the abrasive particles and the polymer provides the flow to the media. The gel provides viscosity. Thus, selecting the proper base oil for the preparation of gel results in effective machining for the AFM and MAF process.

Acknowledgement
Any grant did not fund this research, and The authors appreciate the constructive comments of the reviewers, which led to definite improvement in the paper.

References

[1] Choudhari, Rajashekhar, Manjunatha Gudekote, Hanumesh Vaidya, and Kerehalli Vinayaka Prasad. "Peristaltic flow of Herschel-Bulkley fluid in an elastic tube with slip at porous walls." *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences* 52, no. 1 (2018): 63-75.

[2] Gudekote, Manjunatha, Hanumesh Vaidya, Divya Baliga, Rajashekhar Choudhari, Kerehalli Vinayaka Prasad, and Viharika Viharika. "The effects of convective and porous conditions on peristaltic transport of non-Newtonian fluid through a non-uniform channel with wall properties." (2021).

[3] Baliga, Divya, Manjunatha Gudekote, Rajashekhar Choudhari, Hanumesh Vaidya, and Kerehalli Vinayaka Prasad. "Influence of velocity and thermal slip on the peristaltic transport of a herschel-bulkley fluid through an inclined porous tube." *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences* 56, no. 2 (2019): 195-210.

[4] Vaidya, Hanumesh, Rajashekhar Choudhari, Manjunatha Gudekote, and Kerehalli Vinayaka Prasad. "Effect of variable liquid properties on peristaltic transport of Rabinowitsch liquid in convectively heated complaint porous channel." *Journal of Central South University* 26, no. 5 (2019): 1116-1132. https://doi.org/10.1007/s11771-019-4075-x

[5] Vaidya, Hanumesh, C. Rajashekhar, G. Manjunatha, and K. V. Prasad. "Effect of variable liquid properties on peristaltic flow of a Rabinowitsch fluid in an inclined convective porous channel." *The European Physical Journal Plus* 134 (2019): 1-14. https://doi.org/10.1140/epjp/i2019-12536-y

[6] Vaidya, Hanumesh, C. Rajashekhar, G. Manjunatha, and K. V. Prasad. "Peristaltic mechanism of a Rabinowitsch fluid in an inclined channel with complaint wall and variable liquid properties." *Journal of the Brazilian Society of Mechanical Sciences and Engineering* 41 (2019): 1-14. https://doi.org/10.1007/s40430-018-1543-4

[7] Srinivas, K., Q. Murtaza, and A. K. Aggarwal. "Effect of shape of magnet on the machining of workpiece." *International Journal of Recent Technology and Engineering (IJRTE), ISSN* (2019): 2277-3878. https://doi.org/10.35940/ijrte.B1367.0982S1119

[8] Srinivas, K., Q. Murtaza, and A. K. Aggarwal. "Effect of permanent magnetic pole orientation on-field strength in viscoelastic magnetic abrasive finishing process." *International Journal of Mechanical and Production Engineering Research and Development* 9, no. 5 (2019): 43-52.

[9] Srinivas, Krovvidi, Q. Murtaza, and A. K. Aggarwal. "Modelling of viscoelastic fluid in finishing processes." *International journal of scientific & technology research* 8, no. 9 (2019).

[10] Srinivas, K., Q. Murtaza, and A. K. Aggarwal. "Variation of flow parameters and Magnetic flux in Viscoelastic Magnetic Abrasive Finishing Process." *International Journal of Applied Engineering Research* 14, no. 8 (2019): 1940-1946.

[11] Srinivas, Krovvidi, and Anant Bharadwaj. "Pressure variation in abrasive flow machining: modelling and simulation." *IJAPIE* 3, no. 4 (2018): 15-19.

[12] Bhardwaj, Anant, Krovvidi Srinivas, and Rajiv Chaudhary. "Modelling Of Wind Turbine Magnet for Magnetic Abrasive Finishing and Magnetic Field Assisted Abrasive Flow Machining Process." *Journal of Engg. Research ICARI Special Issue pp* 121 (2021): 130. https://doi.org/10.36909/jer.ICARI.15291

[13] Bhardwaj, Anant, Krovvidi Srinivas, and Rajiv Chaudhary. "Modelling and simulation of abrasive flow machining." *Int J Adv Prod Ind Eng* 6, no. 2 (2021): 01-09.

[14] Anant. Bhardwaj, "computational and experimental analysis of parameters in centrifugal force assisted abrasive flow machining process, Diss. 2019, Delhi Technological University.

[15] Bhardwaj, Anant, Parvesh Ali, R. S. Walia, Qasim Murtaza, and S. M. Pandey. "Development of hybrid forms of abrasive flow machining process: a review." *Advances in Industrial and Production Engineering: Select Proceedings of FLAME 2018* (2019): 41-67. https://doi.org/10.1007/978-981-13-6412-9_5

[16] Srinivas. Krovvidi, "A study of the viscoelastic magnetic abrasive finishing process, Diss. 2020, Delhi Technological University.

[17] A. B. Metzner, "The Significant Rheological Characteristics of Lubricants", Journal of lubrication Technology, Transactions of ASME, July 1968.

[18] Schiessel, H., R. Metzler, A. Blumen, and TF0921 Nonnenmacher. "Generalized viscoelastic models: their fractional equations with solutions." *Journal of physics A: Mathematical and General* 28, no. 23 (1995): 6567. https://doi.org/10.1088/0305-4470/28/23/012

[19] An, Haining, Stephen J. Picken, and Eduardo Mendes. "Enhanced hardening of soft self-assembled copolymer gels under homogeneous magnetic fields." *Soft Matter* 6, no. 18 (2010): 4497-4503. https://doi.org/10.1039/c0sm00216j

[20] Chirikov, D. N., S. P. Fedotov, L. Yu Iskakova, and A. Yu Zubarev. "Viscoelastic properties of ferrofluids." *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics* 82, no. 5 (2010): 051405.

[21] Park, Bong Jun, Fei Fei Fang, and Hyoung Jin Choi. "Magnetorheology: materials and application." *Soft Matter* 6, no. 21 (2010): 5246-5253.

[22] Yu, Miao, Benxiang Ju, Jie Fu, Shuzhi Liu, and Seung-Bok Choi. "Magnetoresistance characteristics of magnetorheological gel under a magnetic field." *Industrial & Engineering Chemistry Research* 53, no. 12 (2014): 4704-4710.

[23] Abro, Kashif Ali, Mirza Mahmood Baig, and Mukruram Hussain. "Influences of magnetic field in viscoelastic fluid." *Int. J. Nonlinear Anal. Appl* 9, no. 1 (2018): 99-109.

Achieving the positivity of the secret key in a BB84 like quantum key distribution protocol

Rashi Jain* and Satyabrata Adhikari†
Department of Applied Mathematics,
Delhi Technological University, Delhi-110042, India

Woodhead [Phys. Rev. A **88**, 012331 (2013)] derived the lower bound of the secret key rate for a Bennett-Brassard (BB84) like quantum key distribution protocol under collective attacks. However, this lower bound does not always assure the generation of the secret key and thus the protocol may have to be aborted sometimes. Thus, we modify the Woodhead's lower bound of the secret key rate in such a way that the secret key is always generated in a BB84 like quantum key distribution protocol. Exploiting the obtained modified lower bound of the secret key rate, we analyze two state dependent quantum cloning machines such as (i) Wootters-Zurek QCM and (ii) Modified Buzek-Hillery QCM constructed by fixing the cloning machine parameters of Buzek Hillery quantum cloning machine (QCM), which may be used by the eavesdropper to extract information from the intercepted state. We, thereafter, show that it is possible for the communicating parties to distill a secret key, even in the presence of an eavesdropper. Moreover, we also discuss the effect of the efficiency of the QCM on the generation of the secret key for a successful key distribution protocol.

arXiv:2504.16434v1 [quant-ph] 23 Apr 2025

I. INTRODUCTION

One of the most challenging problems in the modern world is to communicate securely, over an insecure channel. Quantum cryptography plays a vital role in providing security in the areas related to quantum technologies [1]. Quantum key distribution (QKD) is one of the most advanced applications of quantum cryptography. In a QKD protocol, a secure key is established between two or more trusted communicating parties following the laws of quantum mechanics. This key is then used for secure communication over a quantum channel. Various QKD protocols have been proposed since early 1980s [2, 3]. The security of the key being distributed is the utmost priority for the communicating parties. The distribution of the secret key is a challenging task as there might be some external noise or an eavesdropper present in the quantum channel, which may tamper with the key.

There are broadly two classes of QKD implementation namely, prepare-and-measure QKD and, entanglement-based approach [4]. In this paper, we will focus on the prepare-and-measure QKD scheme. In this scheme, the sender, say Alice, prepares and transmits quantum states randomly, on one of the previously decided bases to the receiver, say Bob. Upon receiving the states, Bob performs quantum measurements upon them on any one of the basis in a random manner. The state preparation basis and state measurement basis are then discussed in the public channel to assure the correctness of the key. This forms a *raw key*. Alice and Bob then chooses randomly any subset of the raw key, discuss their prepared states and calculate the error occurred in the selected subset. If the error is too high, they abort the protocol, else they continue. The remaining qubits form a *sifted key*, which can be improved by applying error correction schemes [5, 6] and privacy amplification schemes [7, 8]. In an ideal case, Bob's string is identical to Alice's string, and the key rate is considered to have no dark counts and perfect detector efficiency [9]. Although, in realistic situations, this might not be the case. An eavesdropper, say Eve, might be present in between, who is in general assumed to posses every possible power that is compatible within the laws of quantum mechanics. The most general form of attacks considered in QKD are joint attacks. An important class of joint attacks is the collective attack [10]. Eve may intercept Alice's state before reaching to Bob in the midway and she may apply an unitary transformation on it in order to know about the secret key. Since the above procedure is repeating several times so after the classical communication has been completed between Alice and Bob, Eve separately applies the same operation to each state that she intercepted from Alice. This is known as a *collective attack* [11].

The interference of Eve generate some errors in the key distribution scheme. This raises the problem of determining the number of key bits that can be securely extracted from a QKD protocol, or in other words, how much error can be tolerated by the QKD scheme such that Alice and Bob are still able to distill a key. In this regard, in 2004, Gottesman et al, proposed the achievable key rate for the BB84 protocol [2], which is given by [12]

$$r = 1 - H_2(\delta) - H_2(\delta_p) \tag{1}$$

where $H_2(\delta)$ denotes the Shannon entropy of a random variable δ, representing the fraction of the sifted key bits being sacrificed to perform error correction, and $H_2(\delta_p)$ represents the fraction of the sifted key bits being sacrificed to perform privacy amplification. Later, in 2005, Devetak and Winter proposed a lower bound on the asymptotic key rate that is secure against an adversary under collective attacks [13]. Kraus et al [14] proposed

* rashijain_23phdam07@dtu.ac.in
† satyabrata@dtu.ac.in

lower and upper bounds on the secret key using one way classical communication QKD protocols. Till date there have been many theories in the literature for the upper bound of key rate in various kinds of QKD protocols [15–18], but only a few works comprises of lower bound of secret key rates [14, 19]. Therefore, in this paper, we derive the possible lower bound, for a successful BB84 like protocol. We call BB84 like protocol is successful, if the key rate is positive. We further find the upper bound on the error rate such that the two communicating parties are able to distill a secret key, even in the presence of Eve, where she uses a quantum cloning transformation to extract information from the intercepted state from the sender, Alice.

The aim of this paper is three-fold. Firstly, we observe that the lower bound for a key rate introduced by [20] is based on Eve's fidelity and the errors introduced in the system for a BB84-like protocol. However, this lower bound does not assure the success of the protocol. Thus, our aim is to find the estimate of the lower bound possible on the secret key rate such that the protocol is completed successfully, i.e., the key is formed positively. Secondly, our aim is to find whether Eve's cloning transformation has some effect on the key rate, or the introduced errors. To study this, we find the maximum possible upper bound of the introduced errors such that the communicating parties are able to distill a secret key. Thirdly, we aim to find if there exists an upper bound on the efficiency of the QCM that may effect the generation of the secret key.

The rest of the paper is organized as follows: In Section II, we will discuss the BB84 like QKD protocol and provide the theoretical proposal for the generation of Alice's state in a laboratory, which she uses for QKD. We also clearly define the role of the eavesdropper in the QKD protocol. In Section III, we provide the estimate of the lower bound for a positive key rate. Moreover, for a given fidelity between the states resulting from the output of the QCM, we find the upper bound on the error rate upto which the protocol proves to be successful. In Section IV, we study the state dependent symmetric QCMs, such as the Wootter-Zurek QCM, and the modified Buzek-Hillery QCM. We find the upper bound on error rate, for a given fidelity and analyze the value of secret key rate, such that the protocol is successful. In Section V, we study the effect of efficiency of the QCM that may be used by Eve in a successful QKD protocol. Finally, we conclude with a brief summary in Section VI.

II. BB84 LIKE PROTOCOL

Consider the following setting of BB84 like protocol, which is similar to the BB84 protocol: Two communicating parties, namely, Alice and Bob are connected via an untrusted quantum channel and they wish to establish a secure secret key for further communication. To start with, let us suppose that Alice possesses an unbiased source that produces the two-qubit state of the following form

$$|\Phi\rangle_{AA} = |0\rangle_A \otimes |S\rangle_A + |1\rangle_A \otimes |S'\rangle_A \tag{2}$$

where $|S\rangle \equiv |\phi\rangle$ or $|\psi\rangle$ and $|S'\rangle \equiv |\phi'\rangle$ or $|\psi'\rangle$. These four quantum states, either from the "Z basis" or "X basis" (these bases may not be identical to the actual Z and X basis known in quantum mechanics) at random, with equal probability [15]. Let us further consider that the two orthogonal states in "Z basis" are $\{|\phi\rangle, |\phi'\rangle\}$, which is of the form

$$\begin{aligned} |\phi\rangle &= \alpha|0\rangle + \beta|1\rangle \\ |\phi'\rangle &= \alpha|1\rangle - \beta|0\rangle \end{aligned} \tag{3}$$

where the real numbers α and β satisfy the normalization condition $\alpha^2 + \beta^2 = 1$. The orthogonal states of the "X basis" can be given by $\{|\psi\rangle, |\psi'\rangle\}$. The density operators corresponding to these states can be given by $\rho = |\phi\rangle\langle\phi|$, $\rho' = |\phi'\rangle\langle\phi'|$, $\sigma = |\psi\rangle\langle\psi|$ and $\sigma' = |\psi'\rangle\langle\psi'|$. Alice performs measurement in the computational basis $\{|0\rangle, |1\rangle\}$ and according to the measurement outcome, she may get one of the four states $\{|\phi\rangle, |\phi'\rangle, |\psi\rangle, |\psi'\rangle\}$. Alice then sends the resulting single qubit state obtained from the measurement, to Bob. Bob chooses either "Z basis" or "X basis" at random and performs measurement on the received state, and keeps the sequence obtained in terms of bits 0 and 1 respectively. Thereafter, Alice and Bob discuss their state preparation basis and state measurement basis, respectively, through the public channel. This forms a raw key between Alice and Bob. We note here that in the variant of BB84 protocol, which is called as the BB84 like protocol, the sender Alice and the receiver Bob use only the "Z basis" results for the secret key generation. Then they select a subset of this raw key to calculate the "Z basis" bit error rate denoted by δ_z. They use the remaining data to extract the keys.

A. Generation of Alice's two-qubit state in the laboratory

In this section, we will show the procedure of the experimental generation of the two-qubit state possessed by Alice. Let us start with a pseudopure state (PPS), given by $|00\rangle$, which is experimentally generated by the process adopted by [21]. The two-qubit state possessed by Alice can be constructed by applying a sequence of various quantum gates upon the PPS state. The two-qubit state possessed by Alice is given by

$$\begin{aligned} |\Phi\rangle_{AA} &= |0\rangle_A \otimes |\phi\rangle_A + |1\rangle_A \otimes |\phi'\rangle_A \\ &= \frac{\alpha}{\sqrt{2}}(|00\rangle_{AA} + |11\rangle_{AA}) + \frac{\beta}{\sqrt{2}}(|01\rangle_{AA} - |10\rangle_{AA}) \end{aligned} \tag{4}$$

where $|0\rangle_A$ and $|1\rangle_A$ are the states of the classical registers possessed by Alice and $|\phi\rangle_A = \alpha|0\rangle_A + \beta|1\rangle_A$ and

$|\phi'\rangle_A = \alpha|1\rangle_A - \beta|0\rangle_A$ denotes the Z basis states emitted by Alice, intended for Bob.
The quantum circuits for the state construction consists of single-qubit rotation gate, controlled-NOT gates, controlled rotation (CROT) gate, followed by a single-qubit Hadamard gate. The unitary transformations applied are such that the normalization of the output state is preserved.
The single-qubit Hadamard gate for the basis $\{|0\rangle, |1\rangle\}$ is defined as

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{5}$$

For the computational basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, the ideal CNOT gate flips the target qubit (second qubit), when the control qubit (first qubit) is $|1\rangle$, which is given by

$$\mathrm{CNOT}_{12} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{6}$$

Similarly, a controlled rotation gate (CROT-gate) is given by

$$\mathrm{CROT}_{12}^{\theta} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ 0 & 0 & \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix} \tag{7}$$

where θ represents the arbitrary angle of rotation.
The experimental construction of the two-qubit state possessed by Alice is given by

$$\begin{aligned}
|00\rangle &\xrightarrow{\mathrm{U}_1^{2\theta_1}} \cos\theta_1|00\rangle + \sin\theta_1|10\rangle \\
&\xrightarrow{\mathrm{CNOT}_{12}} \cos\theta_1|00\rangle + \sin\theta_1|11\rangle \\
&\xrightarrow{\mathrm{CROT}_{12}^{2\theta_2}} \cos\theta_1|00\rangle + \sin\theta_1\cos\theta_2|11\rangle + \sin\theta_1\sin\theta_2|10\rangle \\
&\xrightarrow{\mathrm{H}_1} \frac{1}{\sqrt{2}}\Big[(\cos\theta_1 + \sin\theta_1\sin\theta_2)|00\rangle + (\sin\theta_1\cos\theta_2)|01\rangle \\
&\qquad + (\cos\theta_1 - \sin\theta_1\sin\theta_2)|10\rangle + (-\sin\theta_1\cos\theta_2)|11\rangle\Big] \\
&\xrightarrow{\mathrm{CNOT}_{12}} \frac{1}{\sqrt{2}}\Big[(\cos\theta_1 + \sin\theta_1\sin\theta_2)|00\rangle + (\sin\theta_1\cos\theta_2)|01\rangle \\
&\qquad + (-\sin\theta_1\cos\theta_2)|10\rangle + (\cos\theta_1 - \sin\theta_1\sin\theta_2)|11\rangle\Big]
\end{aligned} \tag{8}$$

The single-qubit unitary transformation $U_1^{2\theta_1}$ is a transformation that rotates the first qubit by arbitrary angle $\theta_1 \in (0, \pi/2)$. The controlled-NOT operation is applied by the implementation of CNOT_{12}, considering the first qubit as the control qubit and second qubit as the target qubit.

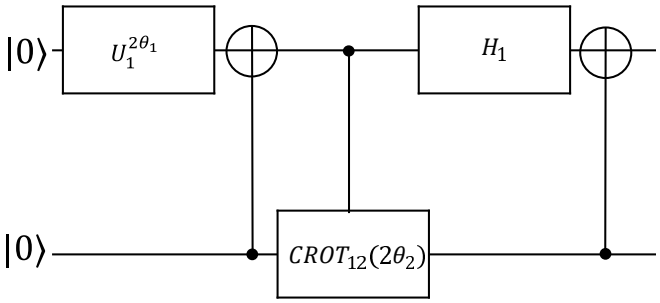


FIG. 1: Quantum circuit [21] showing the sequence of implementation of single-qubit unitary transformation, two-qubit controlled NOT gate, two-qubit controlled rotation gate and single-qubit Hadamard gate required for the construction of two-qubit state possessed by Alice, given by (11).

$\mathrm{CROT}_{12}^{2\theta_2}$ implements a controlled rotation by an arbitrary angle $\theta_2 \in [0, \pi/2)$, with the first qubit as control qubit and second qubit as the target qubit. The final state obtained is dependent upon θ_1 and θ_2.
Now, we choose the values of θ_1 and θ_2 in such a way that the coefficients of the like terms of the state given in (4) matches with the coefficients of the output state obtained in (8). Therefore, we have

$$\cos\theta_1 + \sin\theta_1\sin\theta_2 = \cos\theta_1 - \sin\theta_1\sin\theta_2 = \alpha \tag{9}$$

$$\sin\theta_1\cos\theta_2 = \beta \tag{10}$$

Here, we choose $\theta_2 = 0$, which reduces equation (8) to

$$\frac{1}{\sqrt{2}}\Big[\cos\theta_1(|00\rangle + |11\rangle) + \sin\theta_1(|01\rangle - |10\rangle)\Big] \tag{11}$$

From (11), we obtain the required values of state parameters as

$$\begin{aligned} \alpha &= \cos\theta_1 \\ \beta &= \sin\theta_1 \end{aligned} \tag{12}$$

which satisfies the normalization condition $\alpha^2 + \beta^2 = 1$, for $\theta_1 \in (0, \pi/2)$.

B. Role of eavesdropper

If we assume that a potential eavesdropper, say Eve is present in between the communicating parties, say, Alice and Bob then her role is to secretly intercept Alice's state that is being transferred to Bob. Eve's intention is to extract as much information as possible on the key exchanged between Alice and Bob. To execute her plan of stealing the information, she make use of a quantum cloning machine (QCM) to copy the intercepted state. After successfully applying $1 \longrightarrow 2$ quantum cloning transformation U, she keeps one copy with herself and

sends the other copy to Bob, in order to remain hidden in the key distribution process.
Let us consider that Eve's cloning transformation U acts on the intercepted state in the following way:

$$U(|S\rangle_{in}\langle S| \otimes |0\rangle_b\langle 0| \otimes |Q\rangle_x\langle Q|)U^\dagger, \quad S \equiv \phi \text{ or } \phi' \tag{13}$$

where $|Q\rangle_x$ denotes the QCM state vector, and the subscript "in" and "b" denote the input state and the copy mode, respectively. Eve obtains the copied state by applying a partial trace operation over the machine state vector $|Q\rangle_x$ and the input state $|S\rangle_{in}$ and keeps one of the copied state described by the density operator ρ_E or ρ'_E with herself.
Since the sender Alice and the receiver Bob have no idea about the existence of an eavesdropper in between so, Bob, after receiving the state may proceed with the protocol if the key rate $r > 0$, or they abort it and start over again if the key rate $r < 0$ [22, 23]. The key rate r depends upon the "Z basis" error rate δ_z and it may increase due to the noise present in the communication channel or the interference of the eavesdropper who is present in between the communicating parties. Thus, it would be interesting to study the effect of Eve's cloning transformation on δ_z which may lead to the failure of the protocol and thus may reveal Eve's presence in the key distribution process. Hence, our task is to explore the possible bound of the key rate where the key distribution protocol is successful, even when an eavesdropper is present in the protocol.

III. LOWER BOUND FOR A POSITIVE KEY RATE

In this section, we would like to explore the possible lower bound of the secret key rate when the key distribution protocol is successful, even in the presence of the eavesdropper. Shor and Perskill gave the lower bound on the key rate r, which is given by [24]

$$r \geq 1 - h(\delta_x) - h(\delta_z) \tag{14}$$

where δ_x and δ_z are the X and Z basis error rates, respectively. Woodhead modified the lower bound (14) on r using the cloning bound, for a variant of BB84 scheme, which is given by [20]

$$r \geq 1 - h(u) - h(\delta_z) \tag{15}$$

where $u = \frac{1}{2}(1 + F(\rho_E, \rho'_E))$ and $h(x) = -x\log_2(x) - (1-x)\log_2(1-x)$, is the binary (Shannon) entropy function [25]. The fidelity $F(\rho_E, \rho'_E)$ measures the overlapping between the output states ρ_E and ρ'_E obtained after applying the cloning transformation on the two input states $|\phi\rangle$ and $|\phi'\rangle$, respectively at Eve's site. The fidelity $F(\rho_E, \rho'_E)$ can be calculated as $F(\rho_E, \rho'_E) = ||\sqrt{\rho_E}\sqrt{\rho'_E}||_1$, where $||\cdot||_1$ denotes the trace norm.
It is interesting to note that the right hand side of (15) might be positive or negative, based on the values of $h(u)$ and $h(\delta_z)$. A key distribution protocol is considered to be successful if the key rate is positive, i.e., $1-h(u)-h(\delta_z) > 0$. Therefore, for a successful QKD protocol, we aim to find the lower bound for $1-h(u)-h(\delta_z)$ such that the key rate r is always positive.
For the sake of simplicity, let us assume that $1 - h(u) - h(\delta_z) \equiv R$. We find the lower bound of R denoted by R_{lb} such that r is always positive, i.e., $0 < R_{lb} < R < r$. Now, expressing R in terms of fidelity $F(\rho_E, \rho'_E)$ and δ_z, we get

$$R = \frac{1}{\log 2}\left[\log 2 + \left(\frac{1+F(\rho_E,\rho'_E)}{2}\right)\log\left(\frac{1+F(\rho_E,\rho'_E)}{2}\right) + \left(\frac{1-F(\rho_E,\rho'_E)}{2}\right)\log\left(\frac{1-F(\rho_E,\rho'_E)}{2}\right) + \delta_z\log\delta_z + (1-\delta_z)\log(1-\delta_z)\right] \tag{16}$$

In order to find the lower bound on R such that $0 < R_{lb} < R$, we rewrite (16) as

$$R = \frac{1}{\log 2}[a + \delta_z\log\delta_z + (1-\delta_z)\log(1-\delta_z)] \tag{17}$$

where $a = \log 2 + \left(\frac{1+F(\rho_E,\rho'_E)}{2}\right)\log\left(\frac{1+F(\rho_E,\rho'_E)}{2}\right) + \left(\frac{1-F(\rho_E,\rho'_E)}{2}\right)\log\left(\frac{1-F(\rho_E,\rho'_E)}{2}\right)$.
Let us consider the inequality

$$\frac{x}{1+x} \leq \log(1+x), \ x > -1 \tag{18}$$

We will use this inequality when $-1 < x < 0$ and by replacing x with $-\delta_z$, the inequality (18) reduces to

$$-\delta_z \leq (1-\delta_z)\log(1-\delta_z), \ 0 < \delta_z < 1 \tag{19}$$

Using the inequality (19) in the expression (17), we get

$$R \geq \frac{1}{\log 2}[a + \delta_z\log\delta_z - \delta_z] \equiv R_{lb} \tag{20}$$

Further, applying the inequality $\log\delta_z > \frac{\delta_z^2 - 1}{2.5\delta_z}$ for $\delta_z \in (0, 0.305)$, the expression of R_{lb} given in (20) further

reduces to

$$R_{lb} > \frac{1}{\log 2}\left[a + \frac{\delta_z^2 - 1}{2.5} - \delta_z\right] \tag{21}$$

For $R_{lb} > 0$, we have to find the values of fidelity $F(\rho_E, \rho'_E)$ and δ_z for which the right hand side of (21) is positive, i.e., $a + \frac{\delta_z^2 - 1}{2.5} - \delta_z > 0$ holds. Simplifying this condition, we get a quadratic equation in δ_z, which is given by

$$\delta_z^2 - 2.5\delta_z + 2.5a - 1 > 0 \tag{22}$$

Solving the inequality (22), we get

$$0 < \delta_z < \frac{2.5 - \sqrt{10.25 - 10a}}{2} \tag{23}$$

and

$$\frac{2.5 + \sqrt{10.25 - 10a}}{2} < \delta_z \tag{24}$$

Since, in general, $\delta_z \in (0, 1)$, so we eliminate (24). Therefore, the inequality (23) holds for $F(\rho_E, \rho'_E) \in (0.8281, 0.9922)$.

Substitute a in terms of $F(\rho_E, \rho'_E)$ in the inequality (23) and solving, we find that for a given $F(\rho_E, \rho'_E)$, $\delta_z \in (0, \delta_{z_1})$, where $\delta_{z_1} \in (0, 0.305)$. Therefore, the inequality $0 < R_{lb} < R$ holds when $\delta_z \in (0, \delta_{z_1})$, where $\delta_{z_1} \in (0, 0.305)$. Thus, in the given range of $F(\rho_E, \rho'_E)$ and δ_z, we find that the key rate r is always positive and hence it results into the success of the key distribution protocol.

IV. ANALYSIS OF VARIOUS SYMMETRIC QCMS

In this section, we would like to study the effect of Eve's cloning transformation on the secret key rate r. The success and failure of the protocol depend on its estimate R given in (16), which can be positive or negative, respectively. The estimated value of R depends on the type of cloning transformation used by the eavesdropper. Eve clone the input states $|\phi\rangle$ and $|\phi'\rangle$ and obtains the output states ρ_E and ρ'_E, respectively using various symmetric QCMs. The overlapping between these output states $F(\rho_E, \rho'_E)$ depends on the input state parameters α and the cloning machine parameters. We, therefore, study various types of symmetric QCMs used by the eavesdropper and derive the relationship between $F(\rho_E, \rho'_E)$, the input state parameters, and the cloning machine parameters. For the given values of the input state parameters and the cloning machine parameters, we find the values of $F(\rho_E, \rho'_E)$ and analyze the value of R for the corresponding range of δ_z given in (23). If $R_{lb} > 0$, then the secret key can be established successfully between two distant partners Alice and Bob.

A. Wootter-Zurek QCM

In 1982, Wootters and Zurek [26] analyzed the copying process of an arbitrary input state $|\phi\rangle_a$ which is a linear combination of basis states vectors $|0\rangle_a$ and $|1\rangle_a$, the subscript a denotes the original system. The cloning transformation is defined on the basis vectors $|0\rangle$ and $|1\rangle$, and is given by

$$\begin{aligned} |0\rangle_a|0\rangle_b|Q\rangle_x &\longrightarrow |0\rangle_a|0\rangle_b|Q_0\rangle_x \\ |1\rangle_a|0\rangle_b|Q\rangle_x &\longrightarrow |1\rangle_a|1\rangle_b|Q_1\rangle_x \end{aligned} \tag{25}$$

where $|Q\rangle_x$ denotes the QCM state vector and, the subscript b denotes the copy mode of the input state. The copying state vectors satisfies the conditions

$$\langle Q_0|Q_1\rangle = 0, \langle Q_i|Q_i\rangle = 1, i = 0, 1 \tag{26}$$

The WZ copying machine is input state-dependent, i.e., the quality of the copies at the output of the quantum cloner depends on the input state.

Let us assume that Eve uses WZ QCM as described by (25) to copy the intercepted state that Alice wishes to send to Bob. Therefore, the cloning transformation can be read as

$$\begin{aligned} |0\rangle_A|0\rangle_E|Q\rangle_M &\longrightarrow |0\rangle_B|0\rangle_E|Q_0\rangle_M \\ |1\rangle_A|0\rangle_E|Q\rangle_M &\longrightarrow |1\rangle_B|1\rangle_E|Q_1\rangle_M \end{aligned} \tag{27}$$

where the subscript A refers to Alice's qubit that Eve intercepts, the subscript B refers to the qubit that Eve sends to Bob after copying the intercepted state in order to remain hidden in the key distribution process, Eve's copy state is expressed in the basis state $|i\rangle_E, i = 0, 1$, and $|Q_i\rangle_M$ refers to Eve's cloning machine state vectors at the output.

Let us assume that Alice sends the state $|\phi\rangle = \alpha|0\rangle + \beta|1\rangle$ to Bob. Eve intercepts the state in the midway and applies WZ QCM given by (27) and obtains the state at the output as ρ_{BEM}. The two-qubit copy state described by the reduced density operator ρ_{BE} after applying cloning operation is given by

$$\rho_{BE} = \mathrm{Tr}_M[\rho_{BEM}] = \alpha^2|00\rangle_{BE}\langle 00| + \beta^2|11\rangle_{BE}\langle 11| \tag{28}$$

The single-qubit state ρ_E, which is in possession of Eve is given by

$$\rho_E = \mathrm{Tr}_B[\rho_{BE}] = \alpha^2|0\rangle\langle 0| + \beta^2|1\rangle\langle 1| \tag{29}$$

Further, if Alice sends the qubit $|\phi'\rangle = \alpha|1\rangle - \beta|0\rangle$, then Eve performs the same cloning operation on the intercepted qubit, which results in the two-qubit state after eliminating the machine state vector which is given by

$$\rho_{BE'} = \mathrm{Tr}_M[\rho_{BE'M}] = \beta^2|00\rangle_{BE'}\langle 00| + \alpha^2|11\rangle_{BE'}\langle 11| \tag{30}$$

Eve keeps the single-qubit state with herself and it is described by the density operator ρ'_E, which is given by

$$\rho'_E = \beta^2|0\rangle\langle 0| + \alpha^2|1\rangle\langle 1| \tag{31}$$

The overlapping between two density matrices ρ_E and ρ'_E can be described as [27]

$$F(\rho_E, \rho'_E) = \left(\sum_i \sqrt{\lambda_i(\rho_E, \rho'_E)}\right)^2 \tag{32}$$

where λ_i is the ith eigenvalue of the product of density matrices ρ_E and ρ'_E.
Therefore, the fidelity between Eve's output states ρ_E and ρ'_E when she uses WZ QCM is given by (See Appendix A.1 for calculations (A3))

$$F_{WZ}(\rho_E, \rho'_E) = 4\alpha^2(1-\alpha^2) \tag{33}$$

It can be observed that $F_{WZ}(\rho_E, \rho'_E)$ only depends upon Alice's secret state parameter α^2. It is interesting to note here that $F_{WZ}(\rho_E, \rho'_E)$ displays a symmetric behavior when α^2 is lying in the interval $(0, 0.5)$ and $(0.5, 1)$. Therefore, we will only consider the values of $\alpha^2 \in (0, 0.5)$. Further, we can find the values of α^2, for which the lower bound $R_{lb} > 0$ is satisfied i.e., the key distribution protocol is successfully completed with $R > 0$. To achieve this, we have analyzed the values of α^2 and found that when $\alpha^2 \in (0.293, 0.456)$, the fidelity $F_{WZ}(\rho_E, \rho'_E) \in (0.8281, 0.9922)$. With the help of the range of $F_{WZ}(\rho_E, \rho'_E)$, we find the corresponding range of δ_z for which $R_{lb} > 0$. This is depicted in fig 2.

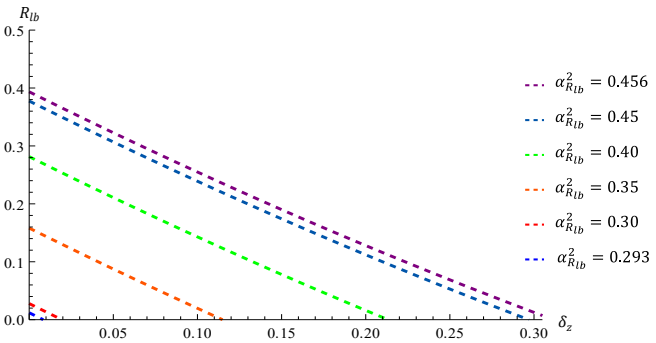


FIG. 2: The lower bound of R for a successful QKD protocol ($r > 0$) when the eavesdropper uses WZ QCM to extract information from the intercepted state is shown. The $x-$ axis represents the range of δ_z, for a given value $F_{WZ}(\rho_E, \rho'_E)$, which depends only on Alice's secret state parameter $\alpha^2 \in (0.293, 0.456)$. The $y-$ axis represents the value of lower bound R_{lb} such that $0 < R_{lb} < R$.

It can be observed from fig 2 that as Alice increases the value of her secret state parameter $\alpha^2 \in (0.293, 0.456)$, the overlapping between Eve's output states $F(\rho_E, \rho'_E)$ also increases when Eve uses WZ QCM. For an increasing $F(\rho_E, \rho'_E)$, the range of δ_z also increases significantly. This is valid because increasing fidelity implies better cloning, and therefore increase in error rate. The estimate of the lower bound of the key rate r, given by R_{lb} varies accordingly.
Since, the protocol is successful in generating the secret key for $R_{lb} > 0$, so Alice is restricted to choosing the state parameter α^2 in the interval $(0.293, 0.456)$. Otherwise, there might arise a case when Alice chooses $\alpha^2 \in (0, 0.293) \cup (0.456, 0.5)$ then R_{lb} might be negative, and thus resulting in the failure of the key distribution protocol. Therefore, the question arises that whether there exist some other cloning machine that may be used by the eavesdropper for which Alice can choose the secret parameter α^2 in the interval $(0, 1)$ with $R_{lb} > 0$ and even then Eve manages to remain hidden in the key distribution protocol?

B. Modified Buzek-Hillery symmetric QCM

In this subsection, we modified the QCM prescribed by Buzek and Hillery [28] and checked whether it fulfilled the asked question in the previous subsection. We modify the BH QCM such that it maybe used by Eve to clone Alice's intercepted qubit such that $R_{lb} > 0$ and she remains hidden in the key distribution protocol.
In 1996, Buzek and Hillery introduced the input state independent QCM, which has the property that the quality of the output state was independent of the input state. The BH quantum cloning transformation is given by

$$\begin{aligned} |0\rangle_A|0\rangle_E|Q\rangle_x \longrightarrow & |0\rangle_B|0\rangle_E|Q_0\rangle_x \\ & + [|0\rangle_B|1\rangle_E + |1\rangle_B|0\rangle_E]|Y_0\rangle_x \\ |1\rangle_A|0\rangle_E|Q\rangle_x \longrightarrow & |1\rangle_B|1\rangle_E|Q_1\rangle_x \\ & + [|0\rangle_B|1\rangle_E + |1\rangle_B|0\rangle_E]|Y_1\rangle_x \end{aligned} \tag{34}$$

The unitarity of the transformation gives

$$\langle Q_i|Q_i\rangle + 2\langle Y_i|Y_i\rangle = 1, i = 0, 1 \tag{35}$$
$$\langle Y_i|Y_j\rangle = \langle Q_i|Q_j\rangle = 0, i \neq j, i, j = 0, 1 \tag{36}$$
$$\langle Q_i|Y_i\rangle = \langle Y_i|Q_i\rangle = 0, i = 0, 1 \tag{37}$$

In addition to the above conditions, the following also holds

$$\langle Y_0|Y_0\rangle = \langle Y_1|Y_1\rangle = \xi, 0 \leq \xi \leq 1/2 \tag{38}$$
$$\langle Y_i|Q_j\rangle = \langle Q_i|Y_j\rangle = \eta/2, i \neq j, i, j = 0, 1 \tag{39}$$

where $0 \leq \eta \leq 1\sqrt{2}$. Here ξ and η denotes the cloning transformation parameters. It can be noted here that if we choose $\xi = 0$, then it implies $\eta = 0$ also, and BH cloning transformation reduces to WZ QCM. So, here we modify the BH QCM by choosing $\eta = 0$ and $\xi \neq 0$. In this case, BH QCM will become a state dependent QCM. Let us assume that Eve uses the modified QCM for cloning the intercepted state. Suppose that Alice sends

the state $|\phi\rangle = \alpha|0\rangle + \beta|1\rangle$ to Bob, which is intercepted by Eve in the mid way. If she applies the modified BH quantum cloning transformation described in (34) on the intercepted state then the combined state becomes ρ_{BEM}. The resultant single-qubit state ρ_E with Eve, after tracing out the copy qubit and the cloning machine state vector, is given by

$$\rho_E = \alpha^2 + \xi(\beta^2 - \alpha^2)|0\rangle\langle 0| + \beta^2 + \xi(\alpha^2 - \beta^2)|1\rangle\langle 1| \tag{40}$$

Similarly, if Alice sends $|\phi'\rangle = \alpha|1\rangle - \beta|0\rangle$ to Bob, then Eve may intercept it and applies the modified BH quantum cloning transformation (34) on it. In this case, the single-qubit state with Eve is given by

$$\rho'_E = \beta^2 + \xi(\alpha^2 - \beta^2)|0\rangle\langle 0| + \alpha^2 + \xi(\beta^2 - \alpha^2)|1\rangle\langle 1| \tag{41}$$

The overlap between the two mixed states described by the density operator ρ_E and ρ'_E, when Eve uses the modified BH QCM can be given by (Refer Appendix A.2 for calculations (A5))

$$F_{BH}(\rho_E, \rho'_E) = 4[\alpha^2(1-\alpha^2)(1-2\xi)^2 + \xi(1-\xi)] \tag{42}$$

It can be clearly seen here that $F_{BH}(\rho_E, \rho'_E)$ not only depends upon Alice's secret state parameter α^2, but also on the cloning machine parameter ξ.

The expression of $F_{BH}(\rho_E, \rho'_E)$ obtained in (42) lies in the interval $(0, 1]$, for $\alpha^2 \in (0, 1)$ and $\xi \in (0, 0.5]$. In this case also, we find that $F_{BH}(\rho_E, \rho'_E)$ is a symmetric function with respect to the parameter α^2. Alice choose the value of α^2 to prepare her secret state without having any idea about the cloning machine state parameter ξ, used by Eve in the construction of modified BH QCM to extract information from the intercepted state. We have shown that for a given value of $\xi \in (0, 0.455)$, there exists certain values of $\alpha^2 \in (0, 0.455)$ for which Eve manages to remain hidden in the key distribution process, and the key is distributed successfully between the sender Alice and receiver Bob.

The results in the given table I can be validated by considering the fact that as the fidelity for Eve increases, the interval of the error rate also increases due to Eve's intervention. We have shown that, it is possible to form the secret key between the sender and receiver, in spite of the increasing interval of the error rate.

ξ	α^2	$F_{BH}(\rho_E, \rho'_E)$	δ_z
0.1	0.241	0.8283	(0,0.0003)
	0.30	0.8976	(0,0.0947)
	0.35	0.9424	(0,0.1748)
	0.40	0.9744	(0,0.2495)
	0.445	0.9936	(0,0.3049)
0.2	0.155	0.8286	(0,0.0006)
	0.20	0.8704	(0,0.0543)
	0.25	0.9100	(0,0.1149)
	0.30	0.9424	(0,0.1748)
	0.35	0.9676	(0,0.2318)
	0.40	0.9856	(0,0.2823)
	0.426	0.9921	(0,0.3044)
0.3	0.001	0.8406	(0,0.0153)
	0.05	0.8704	(0,0.0544)
	0.10	0.8976	(0,0.0947)
	0.15	0.9216	(0,0.1350)
	0.20	0.9424	(0,0.1748)
	0.25	0.9600	(0,0.2133)
	0.30	0.9744	(0,0.2495)
	0.35	0.9856	(0,0.2823)
	0.39	0.9923	(0,0.3049)
0.4	0.001	0.9602	(0,0.2137)
	0.05	0.9676	(0,0.2318)
	0.10	0.9744	(0,0.2496)
	0.15	0.9804	(0,0.2665)
	0.20	0.9856	(0,0.2823)
	0.25	0.9900	(0,0.2969)
	0.28	0.9923	(0,0.3049)
0.455	0.001	0.9919	(0,0.3038)
	0.011	0.9923	(0,0.3049)

TABLE I: For a given value of $\xi \in (0, 0.455)$, there exists values of Alice's secret state parameter α^2 such that the lower bound required for a successful QKD protocol, i.e., $R_{lb} > 0$ is achieved, even when Eve uses the BH QCM.

V. EFFICIENCY OF THE QCM

Alice prepares her secret states $|\phi\rangle$ or $|\phi'\rangle$ and then sends it to Bob. Eve intercepts Alice's state in the midway, before reaching to Bob. She then applies a quantum cloning transformation to extract the information from the intercepted state. After applying the quantum cloning transformation on the intercepted state $|\phi\rangle$ or $|\phi'\rangle$, the possible output is represented by the density operator ρ_E or ρ'_E, respectively.

In this section, we will study the efficiency of the QCM by calculating the Hilbert-Schmidt (HS) distance between the input (ideal) state, described by the density operator ρ_{id} or ρ'_{id} and the output state of the cloning machine, described by the density operator ρ_E or ρ'_E. We have shown that the HS distance between ρ_{id} (ρ'_{id}) and ρ_E (ρ'_E) is upper bounded by the trace distance between ρ_E and ρ'_E, by using the relationship between trace distance and the fidelity. Finally, we have analyzed the inequality between the HS distance and the trace distance to find the effect of the efficiency of the QCM on the successful key generation, i.e., $R_{lb} > 0$.

The distance between the density matrices ρ^{in} and ρ^{out} can be calculated by trace distance $D(\rho^{in}, \rho^{out})$, defined as

$$D(\rho^{in}, \rho^{out}) = \frac{1}{2}\mathrm{Tr}|\rho^{in} - \rho^{out}| = \frac{1}{2}||\rho^{in} - \rho^{out}||_1 \quad (43)$$

where $||M||_1 = \mathrm{Tr}\sqrt{M^\dagger M}$ denotes the trace norm for an arbitrary operator M. In higher dimensional systems, trace distance might require diagonalization of density matrices, and computing the diagonalized matrix can be exponentially difficult, making trace distance a less preferred choice [29]. To overcome this difficulty, an alternative measure is the Hilbert-Schmidt (HS) distance [30–32], which can be defined as

$$D_{HS}(\rho, \sigma) = \mathrm{Tr}[(\rho - \sigma)^2] = ||\rho - \sigma||_2^2 \quad (44)$$

where ρ and σ are two arbitrary quantum states. $||M||_2 = \sqrt{\mathrm{Tr}(M^\dagger M)}$ denotes the Frobenius norm which is easier to calculate both numerically and analytically. Furthermore, employing HS distance can be cost efficient as it can be efficiently computed on a quantum computer, thereby reducing the cost of the experiment [33]. It provides a direct interpretation as an information distance between quantum states [34]. The efficiency of a QCM can be analyzed using the HS distance.

A. Relation between the efficiency of the QCM and the secret key rate for a successful QKD protocol

To establish the relation between the efficiency of the QCM and the generation of the secret key, we start with the relationship between trace distance $D(\rho_E, \rho'_E)$ and fidelity $F(\rho_E, \rho'_E)$, which is given by [35]

$$D(\rho_E, \rho'_E)^2 \leq 1 - F(\rho_E, \rho'_E)^2 \quad (45)$$

In 2019, Coles et. al. proposed a strong bound between the trace distance $D(\rho, \sigma)$ and the HS distance $D_{HS}(\rho, \sigma)$ between the two density matrices ρ and σ as [36]

$$\frac{1}{2}D_{HS}(\rho, \sigma) \leq D(\rho, \sigma)^2 \quad (46)$$

Taking $\rho \equiv \rho_E$ and $\sigma \equiv \rho'_E$, the inequality (46) reduces to

$$D_{HS}(\rho_E, \rho'_E) \leq 2D(\rho_E, \rho'_E)^2 \quad (47)$$

Using the triangular inequality of the distance measure, we get

$$\begin{aligned} &D_{HS}(\rho_E, \rho'_E) \leq D_{HS}(\rho_E, \rho_{id}) + D_{HS}(\rho'_E, \rho_{id}) \\ \implies &D_{HS}(\rho_E, \rho'_E) - D_{HS}(\rho_E, \rho_{id}) \leq D_{HS}(\rho'_E, \rho_{id}) \end{aligned} \quad (48)$$

Since, $D_{HS}(\rho'_E, \rho_{id}) \geq 0$, so we impose the condition

$$D_{HS}(\rho_E, \rho_{id}) \leq D_{HS}(\rho_E, \rho'_E) \quad (49)$$

Using (47) and (49), we have

$$D_{HS}(\rho_E, \rho_{id}) \leq 2D(\rho_E, \rho'_E)^2 \quad (50)$$

Similarly, we get

$$D_{HS}(\rho'_E, \rho'_{id}) \leq 2D(\rho_E, \rho'_E)^2 \quad (51)$$

Equations (50) and (51) are the required relationships between the HS distance between ρ_E (ρ'_E) and ρ_{id} (ρ'_{id}) and the trace distance between ρ_E and ρ'_E.
We find from equation (23), that if $F(\rho_E, \rho'_E) \in (0.8281, 0.9922)$, then $0 < R_{lb} < R$, i.e., the secret key can be formed between Alice and Bob. Then for a given value of $F(\rho_E, \rho'_E) \in (0.8281, 0.9922)$, we have $\delta_z \in (0, \delta_{z_1})$, where $\delta_{z_1} \in (0, 0.305)$. Thus, for a given value of $F(\rho_E, \rho'_E) \in (0.8281, 0.9922)$, a range of $D(\rho_E, \rho'_E) \in (0, 0.56058)$ can be obtained by using the inequality given in (45). Hence, for a given value of $D(\rho_E, \rho'_E)$, an upper bound of $D_{HS}(.,.)$ can be obtained. This implies that, for a given fidelity $F(\rho_E, \rho'_E)$, we can estimate an upper bound on the efficiency of the QCM. This further suggest us that there is an upper bound of the efficiency of the QCM, beyond which QKD protocol may be aborted.

B. Analysis of the efficiency of WZ QCM and modified BH QCM and its effect on successful key generation

In the previous section, we have established the relationship between the HS distance $D_{HS}(.,.)$ and the trace distance $D(.,.)$, which is given by the inequality (50-51). In this section, we are interested in analyzing the efficiency of two particular QCMs, namely, WZ QCM and modified BH QCM, when the QKD protocol is successful.

1. Efficiency of WZ QCM in successful QKD protocol

The efficiency of WZ QCM for $R_{lb} > 0$ can be analyzed by the table (II) given below.

α^2	$F_{WZ}(\rho_E, \rho'_E)$	$D(\rho_E, \rho'_E)^2$	$D_{HS}^{WZ}(\rho_E, \rho_{id})$
0.293	0.8286	[0,0.3134)	[0,0.6268)
0.30	0.84	[0,0.2944)	[0,0.5888)
0.35	0.91	[0,0.1719)	[0,0.3438)
0.40	0.96	[0,0.0784)	[0,0.1568)
0.45	0.99	[0,0.0199)	[0,0.0398)
0.456	0.9922	[0,0.0154)	[0,0.0311)

TABLE II: Analysis of WZ QCM in Alice's secret state parameter α^2 and the HS distance $D_{HS}(\rho_{id}, \rho_E)$ between the input and the output mode.

It can be clearly observed from table II that as the fidelity $F(\rho_E, \rho'_E)$ increases from 0.8286 to 0.9922, the upper bound of HS distance between ρ_{id} (ρ'_{id}) and ρ_E (ρ'_E)

reduces rapidly. This implies that if the eavesdropper uses the WZ QCM to clone the input state $|\phi\rangle$ or $|\phi'\rangle$, then we find that the WZ QCM acts efficiently for a larger value of Alice's state parameter α^2. In this case, Eve manages to remain hidden in the QKD process in spite of generating the secret key,.

2. Efficiency of modified BH QCM in successful QKD protocol

ξ	α^2	$F_{BH}(\rho_E, \rho'_E)$	$D(\rho_E, \rho'_E)^2$	$D_{HS}^{BH}(\rho_{id}, \rho_E)$
0.1	0.241	0.8283	[0,0.3139)	[0,0.6279)
	0.30	0.8976	[0,0.1943)	[0,0.3886)
	0.35	0.9424	[0,0.1118)	[0,0.2237)
	0.40	0.9744	[0,0.0504)	[0,0.1009)
	0.445	0.9936	[0,0.0154)	[0,0.0308)
0.2	0.155	0.8286	[0,0.3134)	[0,0.6268)
	0.20	0.8704	[0,0.2424)	[0,0.4848)
	0.25	0.9100	[0,0.1719)	[0,0.3438)
	0.30	0.9424	[0,0.1118)	[0,0.2237)
	0.35	0.9676	[0,0.0637)	[0,0.1275)
	0.40	0.9856	[0,0.0285)	[0,0.0571)
	0.426	0.9921	[0,0.0157)	[0,0.0314)
0.3	0.001	0.8406	[0,0.2933)	[0,0.5866)
	0.05	0.8704	[0,0.2424)	[0,0.4848)
	0.10	0.8976	[0,0.1943)	[0,0.3886)
	0.15	0.9216	[0,0.1506)	[0,0.3013)
	0.20	0.9424	[0,0.1118)	[0,0.2237)
	0.25	0.9600	[0,0.0784)	[0,0.1568)
	0.30	0.9744	[0,0.0505)	[0,0.1011)
	0.35	0.9856	[0,0.0286)	[0,0.0572)
	0.39	0.9923	[0,0.0154)	[0,0.0308)
0.4	0.001	0.9602	[0,0.0781)	[0,0.1562)
	0.05	0.9676	[0,0.0637)	[0,0.1275)
	0.10	0.9744	[0,0.0505)	[0,0.1011)
	0.15	0.9804	[0,0.0388)	[0,0.0776)
	0.20	0.9856	[0,0.0286)	[0,0.0572)
	0.25	0.9900	[0,0.0199)	[0,0.0398)
	0.28	0.9923	[0,0.0154)	[0,0.0308)
0.455	0.001	0.9919	[0,0.0161)	[0,0.0321)
	0.011	0.9923	[0,0.0154)	[0,0.0308)

TABLE III: Analysis of modified BH QCM based on Eve's cloning machine's parameter ξ, Alice's secret state parameter α^2 and, the HS distance between the input state ρ_{id} and the output state ρ_E.

When Eve uses the modified BH QCM to extract information from the intercepted state, it can be noted that the fidelity $F(\rho_E, \rho'_E)$ of the modified BH QCM not only depends on Alice's state parameter α^2 but also on the cloning machine parameter ξ. Therefore, from table III we can see that for a given value of α^2, the upper bound of HS distance decreases, i.e., the efficiency of the cloning machine increases as ξ increases. This means that for a given value of α^2 and a larger value of ξ, Eve can copy the intercepted state in a better way. Therefore, though Eve can extract information using the QCM but its effect cannot be seen in the generation of secret key. Thus, Alice and Bob are able to generate the secret key as $R_{lb} > 0$, in this case, in spite of Eve's intervention in the QKD protocol.

VI. CONCLUSION

To summarize, we have analyzed the lower bound of key rate proposed by Woodhead [20] for a variant of BB84 protocol. This lower bound does not assure the success of the protocol, i.e., successful key distribution between the communicating parties. This means that if the key rate is found to be negative, the protocol is aborted. To overcome this problem, we modified the lower bound given by Woodhead in such a way that the success of the protocol is assured. For a successful key distribution scheme, we find that for a given fidelity, $F(\rho_E, \rho'_E) \in (0.8281, 0.9922)$, there exists error rate $\delta_z \in (0, \delta_{z_1})$, such that $\delta_{z_1} \in (0, 0.305)$. The modified lower bound depends only upon the overlapping between the density operators ρ_E and ρ'_E, that is obtained by Eve after applying the cloning transformation on the intercepted state $|\phi\rangle$ and $|\phi'\rangle$ respectively, and it guaranteed a successful QKD scheme. In particular, we have studied two types of state dependent QCMs, namely, WZ QCM and modified BH QCM that may be used by the eavesdropper to extract the information from the intercepted state. The modified BH QCM can be constructed by setting the cloning machine parameters η and ξ as $\eta = 0$ and $\xi \neq 0$. In both of these cases, we have shown that it is possible for Alice and Bob to distill a secret key in spite of the presence of the eavesdropper. Moreover, we use the inequality between $F(\rho_E, \rho'_E)$ and $D(\rho_E, \rho'_E)$ to determine the range of $D(\rho_E, \rho'_E)$, for a given value of $F(\rho_E, \rho'_E)$. Since the Hilbert Schmidt distances $D_{HS}(\rho_{id}, \rho_E)$ and $D_{HS}(\rho'_{id}, \rho'_E)$ represent the quality of the output of the cloning machine so we have estimated $D_{HS}(\rho_{id}, \rho_E)$ and $D_{HS}(\rho'_{id}, \rho'_E)$ in terms of $D(\rho_E, \rho'_E)$ and thereby discussed the effect of efficiency of the QCM on the generation of the secret key that is required for a successful key distribution protocol. The use of asymmetric quantum cloning transformations in the proposed methodology is still unexplored, and we leave it as an open question how our results extend in this case.

VII. DECLARATIONS

A. Author contribution statement

RJ, SA: Conceptualization, Methodology, Software, Writing- Original draft preparation, Visualization, Writing- Reviewing and Editing.

B. Competing interests

The authors have no competing interests to declare that are relevant to the content of this article.

[1] N. Gisin, G. Ribordy, Q. Tittel, and H. Zbinden, Rev. of Mod. Phys. **74**, 145 (2002).
[2] C. H. Bennett, and G. Brassard, In: Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India (IEEE, New York, 1984), pp. 175-179.
[3] R. Jain, and S. Adhikari, Eu. Phys. J. D **78**, 145 (2024).
[4] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
[5] A. R. Dixon, and H. Sato, Sci. Rep. **4**, 7275 (2014).
[6] V. A. Pastushenko, and D. A. Kronberg, Entropy **25**, 956 (2023).
[7] T. Horvath, L. B. Kish, and J. Scheuer, Europhy. Lett. **94** 28002, (2011).
[8] B. Y. Tang, B. Liu, Y. P. Zhai, C. Q. Wu, and W. R. Yu, Sci. Rep. **9**, 15733 (2019).
[9] S. Pirandola, et al., Adv. Opt. Photonics **12**, 1012 (2020).
[10] E. Biham, M. Boyer,G. Brassard, J. V. Graaf, and T. Mor, Algorithmica, **34**, 372 (2002).
[11] R. Renner, Arxiv preprint arXiv:0512258 (2005).
[12] D. Gottesman, H.-K. Lo, N. L¨utkenhaus, and J. Preskill, Quant. Inf. Comp. **4**, 325 (2004).
[13] I. Devetak and A. Winter, Proc. R. Soc. A **461**, 207 (2005).
[14] B. Kraus, N. Gisin, and R. Renner, Phys. Rev. Lett. **95**, 080501 (2005).
[15] Ø. Marøy, L. Lydersen, and J. Skaar, Phys.Rev.A **82**, 032337 (2010).
[16] M. Koashi, New J. Phys. **11**, 045018 (2009).
[17] M. Curty, T. Moroder, X. Ma, H. K. Lo, and N. Lutkenhaus, Phys. Rev. A **79**, 032335 (2009).
[18] E. Kaur, K. Horodecki, and S. Das, Phys. Rev. App. **18**, 054033 (2022).
[19] Z. Zhang, Q. Zhao, M. Razavi, and X. Ma, Phys. Rev. A **95**, 012333 (2017).
[20] E. Woodhead, Phys. Rev. A **88**, 012331 (2013).
[21] S. Dogra, K. Dorai, and Arvind, Phys. Rev. A. **91**, 022312 (2015).
[22] P.Wang, Y. Zhang, Z. Lu, X. Wang and Y. Li, New J. Phys. **25**, 023019 (2023).
[23] P.Sohr, S. Ecker, L. Bulla, M. Bohmann and R. Ursin, Phy. Rev. App. 22, 024059 (2024).
[24] P.W. Shor and J. perskill, Phys. Rev. Lett. **85**, 441 (2000).
[25] T. M. Cover, Elements of information theory. John Wiley & Sons (1999).
[26] W. K. Wooters, and W. H. Zurek, Nature London, **299**, 802, (1982).
[27] A. Muller, arXiv:2309.10565v4, (2023).
[28] V. Bužek, and M. Hillery, Phys. Rev. A **54**, 1844 (1996).
[29] V.V. Dodonov, O.V. Man'ko, V.I. Man'ko & A. Wünsche, Jour. Mod. Opt. **47**, 633 (2000).
[30] C. Witte, and M. Trucks, Phys. Lett. A **257**, 14 (1999).
[31] M. Ozawa, Entanglement measures and the Hilbert-Schmidt distance, Phys. Lett. A **268**, 158 (2000).
[32] R. Jain, and S. Adhikari, Phys. Scr. **100**, 035113 (2025).
[33] L. Cincio, Y. Subasi, A. T. Sornborger, and P. J. Coles, New J. Phys. **20**, 113022 (2018).
[34] J. Lee, M. S. Kim, and C. Brukner, Phys. Rev. Lett. **91**, 087902 (2003).
[35] M. A. Nielsen, and I. L. Chuang, Quantum Computation and Quantum Information. Cambridge University Press, New Delhi (2008).
[36] P. Coles, M. Cerezo, and L. Cincio, Phys. Rev. A **100**, 022103 (2019).

Appendix A: Fidelity between ρ and σ

In this section, we will calculate the fidelity between two density matrices ρ and σ.
The fidelity $F(\rho,\sigma)$ is given by [27]

$$F(\rho,\sigma)=\left(\sum_i\sqrt{\lambda_i(\rho\sigma)}\right)^2 \tag{A1}$$

where λ_i is the ith eigenvalue of the product $\rho\sigma$.

1. Fidelity of WZ QCM for $|\phi\rangle$ and $|\phi'\rangle$

The fidelity between the density matrices ρ_E and ρ'_E given in (29-31) maybe calculated as

$$\rho_E\rho'_E=\begin{bmatrix}\alpha^2\beta^2 & 0\\ 0 & \alpha^2\beta^2\end{bmatrix} \tag{A2}$$

The eigenvalues of $\rho_E\rho'_E$ are given by $\alpha^2\beta^2,\alpha^2\beta^2$. Therefore, substituting these values in (A1), we get

$$\begin{aligned}F_{WZ}(\rho_E,\rho'_E)&=\left(\sqrt{\alpha^2\beta^2}+\sqrt{\alpha^2\beta^2}\right)^2\\&=\left(2\sqrt{\alpha^2\beta^2}\right)^2\\&=4\alpha^2\beta^2\end{aligned} \tag{A3}$$

2. Fidelity of the modified BH QCM

The eigenvalues of the product of the density matrices ρ_E and ρ'_E obtained by Eve after applying the cloning transformation (34) and tracing out Bob's bits, can be given by $\{(\alpha^2+\xi(\beta^2-\alpha^2))(\beta^2+\xi(\alpha^2-\beta^2)),(\alpha^2+\xi(\beta^2-\alpha^2))(\beta^2+\xi(\alpha^2-\beta^2))\}$. Therefore, the fidelity

$F_{BH}(\rho_E, \rho'_E)$ is given by

$$\begin{aligned} F_{BH}(\rho_E, \rho'_E) &= \Big(\sqrt{(\alpha^2 + \xi(\beta^2 - \alpha^2))(\beta^2 + \xi(\alpha^2 - \beta^2))} \\ &\quad + \sqrt{(\alpha^2 + \xi(\beta^2 - \alpha^2))(\beta^2 + \xi(\alpha^2 - \beta^2))} \Big)^2 \\ &= 4(\alpha^2 + \xi(\beta^2 - \alpha^2))(\beta^2 + \xi(\alpha^2 - \beta^2)) \end{aligned} \tag{A4}$$

Now using the normalization condition $\beta^2 = 1 - \alpha^2$, we obtain

$$F_{BH}(\rho_E, \rho'_E) = 4\left(\alpha^2(1-\alpha^2)(1-2\xi)^2 + \xi(1-\xi)\right) \tag{A5}$$

arXiv:2504.16774v1 [eess.IV] 23 Apr 2025

ORIGINAL ARTICLE

Journal Section

Advanced Chest X-Ray Analysis via Transformer-Based Image Descriptors and Cross-Model Attention Mechanism

Lakshita Agarwal[1] | **Bindu Verma**[2*]

[1]Department of Information Technology, Delhi Technological University, Delhi, 110042, India

[2]Assistant Professor, Department of Information Technology, Delhi Technological University, Delhi, 110042, India

Correspondence
Bindu Verma, Assistant Professor, Department of Information Technology, Delhi Technological University, Delhi, 110042, India
Email: bindu.cvision@gmail.com

Funding information
No Funding Required

The examination of chest X-ray images is a crucial component in detecting various thoracic illnesses. The study introduces a new image description generation model that integrates a Vision Transformer (ViT) encoder with cross-modal attention and a GPT-4-based transformer decoder. The ViT captures high-quality visual features from chest X-rays, which are fused with text data through cross-modal attention to improving the accuracy, context, and richness of image descriptions. The GPT-4 decoder transforms these fused features into accurate and relevant captions. The model was tested on the National Institutes of Health (NIH) and Indiana University (IU) Chest X-ray datasets. On the IU dataset, it got scores of 0.854 (B-1), 0.883 (CIDEr), 0.759 (METEOR), and 0.712 (ROUGE-L). On the NIH dataset, it had the best on all metrics: BLEU 1–4 (0.825, 0.788, 0.765, 0.752), CIDEr (0.857), METEOR (0.726), and ROUGE-L (0.705). This framework has the ability to enhance chest X-ray evaluation, assisting radiologists in more precise and efficient diagnosis.

KEYWORDS
Deep Learning, Image Description Generation, Transformer-based

[*]Equally contributing authors.

1 | INTRODUCTION

Automated image description generation is an emerging technology in medical imaging that falls under the domain of computer vision. It enables the creation of textual summaries that accurately describe visual content [1]. Chest X-ray imaging is crucial for diagnosing and treating thoracic diseases and providing important information about pulmonary and cardiovascular disorders. The integration of deep learning techniques with artificial intelligence (AI) has brought a revolution in medical image analysis in recent years [2]. These technologies enhance the precision of diagnoses and promote the efficiency of radiological workflows by expediting report generation and minimizing interpretation time [3 4].

The proposed CrossViT-GPT4 architecture combines the cross-model attention mechanism of the Vision Transformer (ViT) with the language model of GPT-4. ViT captures spatial information at the pixel level, converts it into manageable sequences, and extracts comprehensive features through cross-model attention, which helps capture long-range dependencies. This integration of textual and visual data improves image description generation by enhancing sentence accuracy, contextual relevance, and detail, while offering scalability for diverse applications. The approach was tested on established chest X-ray datasets like NIH and IU, demonstrating its effectiveness. However, challenges such as computational complexity, limited annotated medical data, domain-specific terminology, clinical integration, and accuracy in analyzing flawed images remain. The proposed methodology combined ViT and cross-model attention with GPT-4, generating detailed annotations that improve the medical terminology as well as the spatial relationships with pixel-level information. It promises to outperform previous approaches towards improving diagnostic workflows and patient care.

The paper's contributions are as follows:

- The research introduces a new method, CrossViT-GPT4, that combines Vision Transformer (ViT) with cross-model attention and GPT-4 to generate detailed annotations for chest X-ray images.
- The ViT model with cross-model attention leverages contextual comprehension to integrate pertinent medical terminology and spatial relationships, whereas GPT-4 utilizes pixel-level information to generate sequences.
- The suggested model outperformed existing approaches in extensive tests conducted on benchmark datasets of the Indiana University Chest X-ray dataset and the National Institutes of Health (NIH) Chest X-ray dataset. This demonstrates the potential to enhance analysis workflows, clinical outcomes, and patient care by enabling automated production of image descriptions for better diagnosis and treatment planning.
- Ultimately, the study offers a numerical evaluation of all methods and details of the advantages and disadvantages of the proposed model.

The paper is organized as follows: Previous research in this field that is relevant is discussed in Section 2. Comprehensive information about the proposed method that is employed is provided in Section 3. The utilized datasets and the final projected descriptions and outcomes are examined in Section 4. The work's future directions and conclusion are summarised in Section 5.

2 | RELATED WORK

The primary objective of the image description model is to identify different objects and represent their relationships through accurate, semantically correct sentences. Various methods have been adapted to medical imaging tasks [5]. Sun et al.[6] proposed a feature-augmented (FA) module validated on datasets such as MS-COCO, which uses the multi-modal pre-trained CLIP model and channels attention within the encoder to enhance image description and captioning performance. Yao et al.[7] proposed a CNN-RNN framework for anatomical structure and abnormality identification in radiological imaging reports. Li et al.[8] used BERT-based models to generate radiological reports from chest X-ray images, showing the significance of contextual language understanding in medical image analysis. Shaikh et al.[9] proposed an encoder-decoder transformer model combined with a pre-trained CheXNet model, evaluated on the IU X-ray dataset, for chest X-ray report generation. The CheXReport model [10] obtains the best-reported performance on the MIMIC-CXR dataset, given the use of Swin Transformer blocks. Retrieval-based methods that incorporate deep neural networks are also presented [11]. In another work, Conditional Self Attention Memory-Driven Transformer was proposed, which outperformed all existing state-of-the-art approaches with a high BLEU score value for radiological report production by taking ResNet152 v2 for feature extraction and the self-attention memory-driven transformer for text generation [12]. Despite these developments, current models fail to integrate textual and visual information effectively and are not able to handle medical language well, which makes the descriptions less understandable and accurate. Moreover, because these models are highly sensitive to errors, misunderstandings in medical contexts could potentially affect diagnoses.

The suggested approach, CrossViT-GPT4, enhances previous research by integrating the benefits of GPT-4, which excels in contextual language modeling, with ViT and a cross-model attention technique for spatial feature extraction. This combination offers a comprehensive solution for automatically generating image descriptions in chest X-ray analysis.

3 | OVERVIEW OF THE PROPOSED WORK

The automated generation of image descriptions has received recent attention due to its intricate nature, which integrates natural language processing (NLP) with computer vision. The objective of the proposed research is to enhance the extraction of spatial and channel-wise information from images. The model is fine-tuned over multiple epochs (75) using the cross-entropy loss function and the Adam optimizer. Figure 1 represents the overall framework of the proposed approach.

Each step of the proposed model is discussed below:

3.1 | Data pre-processing

The input images (I) are initially partitioned into patches (P) of a predetermined size to prepare the input data for the encoder. The patches undergo flattening and linear projection, and each individual patch is thereafter embedded and compressed into a sequence of vectors. The data pre-processing step for the task is shown in Figure 1(a). Equation (1) denotes compressing each patch into a lower-dimensional space:

$$E_{\text{patch}} = \text{Linear}(\text{Flatten}(I_{\text{patch}})) + \text{PE} \tag{1}$$

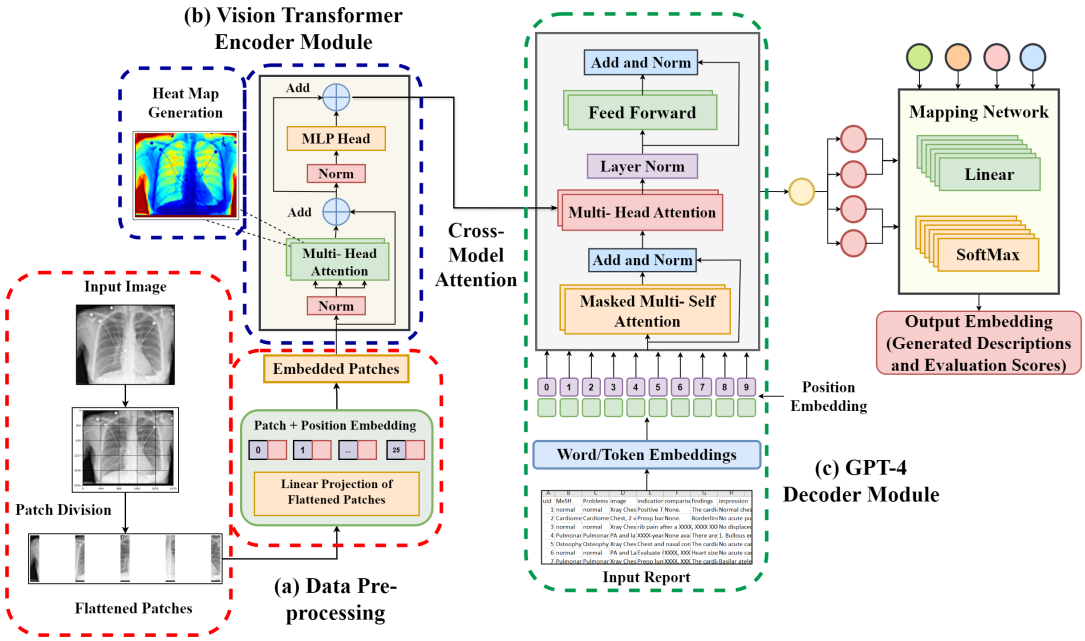


FIGURE 1 A representation of the proposed framework's structure: CrossViT-GPT4- During the initial stage, the ViT encoder utilizes cross-model attention to extract high-level visual features from the pre-processed images. GPT-4 decoder uses tokenization to extract individual words from the provided input caption file. It then utilizes the dense layers of the network to produce image descriptions.

where, I_{patch} represents the image patches, Linear denotes the linear transformation matrix used for embedding, and Flatten is the operation that flattens each patch into a vector. The positional encoding (PE) uses sine and cosine functions to include spatial information. This lets the model understand the order of events, as shown in Equations (2) and (3):

$$\text{PE}(\text{pos}, 2i) = \sin\left(\frac{\text{pos}}{10000^{(2i/\text{dim})}}\right) \tag{2}$$

$$\text{PE}(\text{pos}, 2i+1) = \cos\left(\frac{\text{pos}}{10000^{(2i/\text{dim})}}\right) \tag{3}$$

3.2 | Encoder Module of the Proposed Architecture

The encoder part efficiently collects the spatial and visual characteristics of the image, thereby preparing the model for subsequent tasks such as image labeling. The vision transformer (ViT) encoder is used to receive the sequences of tokenized patches that have been enhanced with positional data. The architectural structure for the Vision transformer encoder is illustrated in Figure 1(b). For multi-head self-attention, the attention is computed in parallel across

h heads and concatenated, as denoted in Equations (4), (5) and (6):

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{4}$$

$$\text{MultiHead}(Q, K, V) = \text{Concat}\left(\text{Head}_1, \ldots, \text{Head}_h\right)W^O \tag{5}$$

$$\text{Head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \tag{6}$$

where, Q, K, and V represent the queries, keys, and values matrices, respectively. d_k is the dimension of the keys, W_i^Q, W_i^K, and W_i^V are projection matrices for each head and W^O is the output projection matrix. The multi-head attention module further generates heat maps to offer insights into the attention mechanism of the model, as illustrated in Figure 1(b). These heat maps aid in identifying the specific areas of the input image that impact key description elements. Additionally, the cross-model attention mechanism is used to transmit these extracted features to the GPT-4 decoder module for final mapping.

3.3 | Cross-model attention mechanism

Through cross-modal attention, which dynamically focuses on pertinent features from each modality, the proposed model is able to match visual details with textual descriptions. In image description generation, the model creates sentences by focusing on various aspects of an image and lining them up with the context that the text provides. Cross-model attention mechanism can be represented by the following Figure 2. The alignment enhances the model's

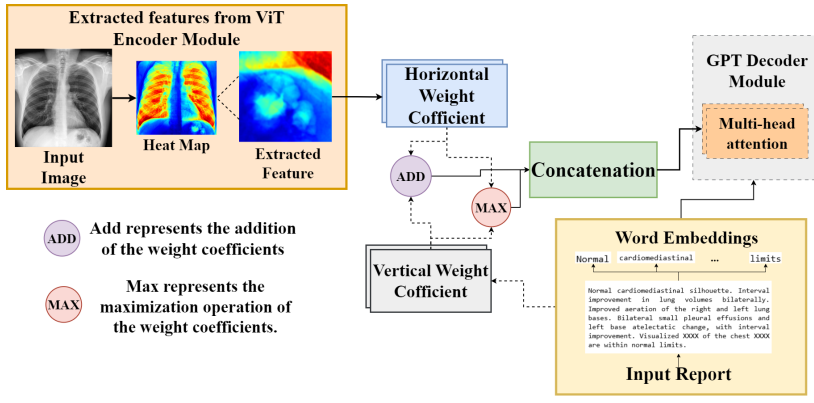


FIGURE 2 Cross-Model Attention Mechanism.

capacity to generate precise and contextually appropriate descriptions, eventually leading to a richer understanding

and more coherent outputs. The cross-modal attention mechanism is defined as Equation (7):

$$\text{CrossAttention}(Q_t, K_i, V_i) = \text{Softmax}\left(\frac{Q_t K_i^T}{\sqrt{d_t}}\right) V_i \tag{7}$$

where, Q_t represents text queries, K_i and V_i are the keys and values from the image features and d_t is the dimension of the text queries.

3.4 | Decoder Module of the Proposed Architecture

The primary architecture of the decoder is based on the GPT-4 transformer model, which is specifically designed for context-based tasks. The structure consists of layers of feed-forward and multiple self-attention neural networks. The diagram in Figure 1(c) illustrates the structural framework of GPT-4. Using self-attention mechanisms, every layer progressively enhances input embedding, enabling the model to discern distant relationships and contextual information. The self-attention within the GPT-4 decoder is given by Equation (8):

$$\text{SelfAttention}(Q_t, K_t, V_t) = \text{Softmax}\left(\frac{Q_t K_t^T}{\sqrt{d_t}}\right) V_t \tag{8}$$

where, Q_t, K_t, and V_t represent the text queries, keys, and values respectively and d_t is the dimension of the text queries. To create the next word in the sequence, the output layer predicts the probability distribution across the vocabulary using word/token embedding, representing words as vectors. Due to this methodology, GPT-4 can generate cohesive and contextually suitable information by extensively working on more substantial sentences. The final textual output is obtained by passing the combined attention outputs through a feed-forward (FF) network and then generating Logits for each token in the vocabulary as represented in Equation (9):

$$\text{Logits} = \text{FF}\left(\text{CrossAttention} + \text{SelfAttention}\right) \tag{9}$$

Linear and Softmax layers in network mapping are essential for producing the end result and predicting evaluation scores when generating image descriptions. The model can represent intricate relationships with them appropriately. Equation (10) represents the expression denoted as $\text{Linear}(W, X)$.

$$\text{Linear}(W, X) = WX + b \tag{10}$$

The weight matrix is symbolized by W, the input vector is indicated by X, and the bias vector is designated by b. Next, the updated characteristics are normalized into a probability distribution throughout the lexicon of potential words or tokens using the Softmax layer, as represented by Equation 11.

$$\text{Softmax}(\mathbf{z})_i = \frac{e^{z_i}}{\sum_j e^{z_j}}, \tag{11}$$

where, the input vector is denoted as z. The accuracy of these predictions is evaluated by comparing them to the descriptions in the reference file. The evaluation metrics used are ROUGE-L, CIDEr, METEOR, and BLEU 1-4, which assess the degree of similarities between the predicted and actual descriptions. These metrics enable the evaluation of the model's ability to effectively integrate the context and semantics of the images into the generated descriptions.

3.5 | Algorithm of the Proposed Work:

The following describes Algorithm 1 for the suggested model along with its score evaluation:

Algorithm 1: Algorithm for the proposed model, CrossViT-GPT4

Input: Images I

Output: Generated Descriptions $\text{Output}_{\text{text}}$, Evaluation Scores

Data Pre-processing:
Partition the image I into patches P, flatten and linearly project each patch;
Apply positional encoding to patches;
ViT Encoding:
Input the sequence of embedded patches E_{patch} into the ViT encoder;
Compute multi-head attention: $\text{MultiHead}(Q, K, V)$;
Cross-Modal Attention:
Compute cross-modal attention: $\text{CrossAttention}(Q_{\text{t}}, K_{\text{i}}, V_{\text{i}})$;
GPT-4 Decoding:
for *each time step t in decoding* **do**
    Compute self-attention within the text: $\text{SelfAttention}(Q_{\text{t}}, K_{\text{t}}, V_{\text{t}})$;
    Generate logits for the next token: Logits;
    Compute probability distribution and select the next token: $\text{Output}_{\text{text}} = \text{Softmax}(\text{Logits})$;
end
Evaluation:
Evaluate generated descriptions using BLEU, CIDEr, METEOR, and ROUGE-L scores;

4 | EXPERIMENTAL ANALYSIS

The experiments were performed using Google Colab Pro+, which provides high-performance resources, including 52 GB of RAM, an NVIDIA A100 GPU with 40 GB of VRAM, and a virtual CPU equivalent to an Intel Xeon processor. The framework was implemented using Keras and TensorFlow 2.12. The ViT model has 86.7 million parameters and GPT-4 has 1.76 trillion. Input images are (batch_size, 224, 224, 3), and output shapes vary by layer. The developed framework was evaluated on two benchmark datasets: the Indiana University Chest X-Ray dataset (IU X-Ray) and the NIH Chest X-ray dataset. The IU X-ray dataset consists of 3,955 XML radiologist reports and 7,471 PNG images while the NIH dataset consists of 112,120 frontal X-rays with annotations for 13 thoracic illnesses across 30,805 individuals.

4.1 | Ablation Study:

The suggested model's image description components were clarified using ablation analysis. This study investigates the capabilities of the proposed architecture, CrossViT-GPT4. In Table 1, we compare the performance of various models on two major chest X-ray datasets: the IU Chest X-Ray dataset and the NIH Chest X-Ray dataset. The results presented in the following table summarize the performance of each model across these metrics.

TABLE 1 Evaluation of Models on the IU and NIH Chest X-Ray Datasets

Dataset	Model	B-1	B-2	B-3	B-4	C	M	R-L
IU Chest X-Ray	CNN-LSTM Model	0.289	0.272	0.268	0.236	0.677	0.273	0.219
	Encoder-Decoder Transformer	0.398	0.373	0.354	0.343	0.658	0.286	0.316
	ViT+ Attention	0.283	0.258	0.236	0.205	0.574	0.194	0.188
	ViT + attention + LSTM	0.601	0.572	0.554	0.498	0.587	0.436	0.342
	ViT + GPT 2.0	0.704	0.663	0.634	0.626	0.792	0.564	0.474
	CNN + GPT 3.0	0.761	0.753	0.736	0.715	0.867	0.725	0.657
	CrossViT-GPT4 (Proposed)	**0.854**	**0.817**	**0.804**	**0.785**	**0.883**	**0.759**	**0.712**
NIH Chest X-Ray	CNN-RNN Model	0.211	0.178	0.161	0.148	0.445	0.198	0.149
	CNN-LSTM Model	0.277	0.252	0.246	0.227	0.465	0.242	0.197
	Encoder-Decoder Transformer	0.383	0.363	0.334	0.312	0.578	0.265	0.287
	ViT with Attention	0.284	0.268	0.253	0.248	0.534	0.187	0.176
	ViT with LSTM	0.438	0.425	0.409	0.387	0.589	0.436	0.386
	ViT + GPT 2.0	0.687	0.658	0.644	0.612	0.763	0.582	0.604
	CNN + GPT 3.0	0.782	0.774	0.752	0.706	0.835	0.704	0.664
	CrossViT-GPT4 (Proposed)	**0.825**	**0.806**	**0.795**	**0.772**	**0.857**	**0.726**	**0.705**

The table presents the performance comparison of different image description generation models on IU and NIH Chest X-ray datasets with respect to BLEU scores from B-1 to B-4, CIDEr (C), METEOR (M), and ROUGE-L (R-L). The proposed CrossViT-GPT4 model is seen to outperform all other models for both the datasets with respect to all metrics. It reports a BLEU score of 0.854 on B-1, 0.817 on B-2, 0.804 on B-3, and 0.785 on B-4 along with CIDEr of 0.883, METEOR of 0.759, and ROUGE-L of 0.712 on the IU Chest X-ray dataset. Correspondingly, it has a score of 0.825 on B-1, 0.806 on B-2, 0.795 on B-3, 0.772 on B-4, CIDEr of 0.857, METEOR of 0.726, and ROUGE-L of 0.705 on the NIH Chest X-ray dataset.

4.2 | Results and Analysis:

The suggested model, CrossViT-GPT4, provides an excellent approach for illustrating images in the IU and NIH chest X-ray datasets. It combines the ViT encoder and the GPT-4 decoder modules. The results for the chest X-ray datasets from IU and NIH are shown in able 2 and 3, respectively.

The tables provide a quantitative comparison between the actual data and the reports generated by the model for different test images. The results evaluate the model's capacity to forecast diseases and produce descriptions that showcase the model's precision by utilizing X-ray images for both datasets.

TABLE 2 Quantitative Results Obtained on IU Chest X-Ray Dataset

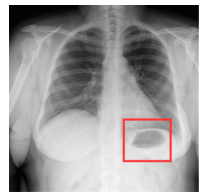
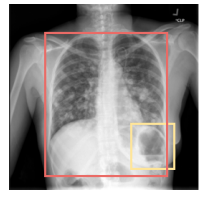
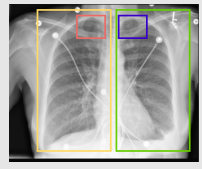
Test Image	Ground Truth Report	Disease Prediction	Predicted Report
	The heart is normal in size. The mediastinal contours are within normal limits. There is mild prominence of the superior mediastinum which is somewhat lucent and reflects mediastinal and vascular structures. No focal consolidation is seen. There is no pleural effusion.		Mediastinal contours within normal limits. Mild mediastinum somewhat lucent.
	Lateral view of the chest shows an unchanged cardiomediastinal silhouette. The cardiac silhouette remains moderately enlarged, exaggerated by epicardial fat pads. Interstitium is prominent. No focal airspace consolidation or pleural effusion. There is spine spondylosis.		Unchanged cardiomediastinal silhouette. Interstitium prominent. Spine spondylosis.

TABLE 3 Quantitative Results Obtained on NIH Chest X-Ray Dataset

Test Image	Ground Truth Report	Disease Prediction	Predicted Report
	1. The heart is normal size. 2. The mediastinum is unremarkable. 3. There is no pleural effusion, pneumothorax, or focal airspace disease. 4. There is stable irregularity of the posterior left 6th rib which represents an old fracture.		No pleural effusion, pneumothorax, or focal airspace disease
	1. The heart size and cardiomediastinal silhouette are normal. 2. There is hyperexpansion of the lungs with flattening of the hemidiaphragms. 3. There is no focal airspace opacity, pleural effusion, or pneumothorax. 4. There multilevel degenerative changes of thoracic spine.		Cardiovascular silhouette normal. No focal airspace opacity, pleural effusion, or pneumothorax.

4.3 | Comparison With Other state-of-the-art Methods:

Compared to the most advanced methodologies currently available, the proposed method demonstrates a substantial enhancement in disease prediction accuracy and report production for X-ray images. The method, CrossViT-GPT4,

effectively performs on both the IU and NIH chest X-ray datasets by utilizing a fusion model that incorporates the ViT encoder and GPT-4.0 decoder.

The study presented in Table 4 thoroughly analyzes various models used to generate image descriptions on the IU X-ray and NIH datasets. The table compares various models for generating image descriptions on the IU Chest X-Ray

TABLE 4 Comparative Analysis on IU and NIH Chest X-Ray Datasets

Model	BLEU	METEOR	CIDEr	Rouge-L
IU Chest X-Ray Dataset				
R2Gen [13]	0.470	0.187	-	0.371
R2Gen + ChexNet [14]	0.508	0.222	-	0.365
Cross-modal PROtotype driven NETwork (XPRONET) [15]	0.525	0.220	-	0.411
Contrastive attention [8]	0.492	0.193	-	0.381
Knowledge-injected U-Transformer [16]	0.525	0.242	-	0.409
AERMNet [17]	0.486	0.219	0.560	0.398
CrossViT-GPT4 (Proposed)	**0.854**	**0.759**	**0.883**	**0.712**
NIH Chest X-Ray Dataset				
Semantic Attention [18]	0.467	0.192	0.560	0.204
Co-Attention [19]	0.756	0.597	0.755	0.675
Clinical-BERT [20]	0.383	0.144	-	0.275
ChestBioX-Gen [21]	0.668	0.189	0.416	0.674
CrossViT-GPT4 (Proposed)	**0.825**	**0.726**	**0.857**	**0.705**

and NIH Chest X-Ray datasets, evaluating their performance using BLEU, METEOR, CIDEr, and Rouge-L metrics. On the IU Chest X-Ray dataset, R2Gen [13] demonstrates intermediate performance, while R2Gen + ChexNet [14] improves BLEU and METEOR scores but slightly reduces Rouge-L. Both XPRONET [15] and Knowledge-injected U-Transformer [8] achieve a BLEU score of 0.525, with the latter excelling in METEOR and Rouge-L. Contrastive Attention exhibits moderate performance, whereas AERMNet [17] provides balanced results across metrics. The proposed CrossViT-GPT4 model outperforms all these approaches, achieving significantly higher scores across all metrics: BLEU (0.854), METEOR (0.759), CIDEr (0.883), and Rouge-L (0.712). On the NIH Chest X-Ray dataset, the CrossViT-GPT4 model similarly leads with the highest scores across all metrics: BLEU (0.825), METEOR (0.726), CIDEr (0.857), and Rouge-L (0.705). In comparison, Semantic Attention [18], Co-Attention [19], ChestBioX-Gen [21], and Clinical-BERT [20] exhibit less balanced or lower performance. These results underscore the superior capability of the proposed model in generating accurate and semantically rich image descriptions across both datasets.

5 | CONCLUSION AND FUTURE DIRECTION

Combining the Vision Transformer encoder module (ViT) with cross-model attention and the Generative Pre-trained Transformers 4 (GPT 4.0) decoder module can result in an efficient framework, CrossViT-GPT4, for producing visual

descriptions. The model uses vision-based feature extraction and language modeling to thoroughly analyze and characterize complex medical images. When textual and visual components are combined, chest X-ray pathology reports can be described and explained more precisely. Transformer-based designs are scalable and adaptable, promoting medical image analysis research and innovation. Their adaptability enables them to efficiently manage diverse datasets and tasks. By enhancing medical image processing, the suggested approach may increase diagnostic precision and assist clinical decision-making platforms. Poor image quality can seriously affect model accuracy by restricting the ability to extract features and generate accurate descriptions. Future research could address this by creating reliable pre-processing methods or incorporating sophisticated denoising and super-resolution techniques to improve performance. Additional research could also increase the applicability and adaptability of the proposed method in a number of scenarios by utilizing a range of large and diverse medical imaging datasets.

Acknowledgment

The authors sincerely thank the Cyber Forensics and Malware Analysis Laboratory, Department of IT, DTU, Delhi, India, for providing the necessary resources to complete the research.

Data Availability Statement:

The paper contains links to all datasets.

Competing Interests:

The authors declare they have no competing financial interests.

Conflict of Interest:

The authors declare no conflict of interest.

references

[1] Beddiar DR, Oussalah M, Seppänen T. Automatic captioning for medical imaging (MIC): a rapid review of literature. Artificial intelligence review 2023;56(5):4019–4076.

[2] Agarwal L, Verma B. Comparison of Deep Learning Models for Automatic Image Descriptors. In: 2023 IEEE 20th India Council International Conference (INDICON) IEEE; 2023. p. 914–919.

[3] Revathi B, Kowshalya AM. Automatic image captioning system based on augmentation and ranking mechanism. Signal, Image and Video Processing 2024;18(1):265–274.

[4] Agarwal L, Verma B. From methods to datasets: A survey on Image-Caption Generators. Multimedia Tools and Applications 2024;83(9):28077–28123.

[5] Pan X, Ye T, Han D, Song S, Huang G. Contrastive language-image pre-training with knowledge graphs. Advances in Neural Information Processing Systems 2022;35:22895–22910.

[6] Sun JT, Min X. Research on image caption generation method based on multi-modal pre-training model and text mixup optimization. Signal, Image and Video Processing 2024;p. 1–19.

[7] Yao T, Li Y, Pan Y, Wang Y, Zhang XP, Mei T. Dual vision transformer. IEEE transactions on pattern analysis and machine intelligence 2023;.

[8] Liu F, Yin C, Wu X, Ge S, Zou Y, Zhang P, et al. Contrastive attention for automatic chest x-ray report generation. arXiv e-print arXiv:210606965 2021;.

[9] Shaikh Z, Bharti J. Transformer-Based Chest X-ray Report Generation Model. In: International Conference on Soft Computing and Signal Processing Springer; 2023. p. 227–236.

[10] Zeiser FA, da Costa CA, de Oliveira Ramos G, Maier A, da Rosa Righi R. CheXReport: A transformer-based architecture to generate chest X-ray reports suggestions. Expert Systems with Applications 2024;p. 124644.

[11] Reale-Nosei G, Amador-Domínguez E, Serrano E. From vision to text: A comprehensive review of natural image captioning in medical diagnosis and radiology report generation. Medical Image Analysis 2024;p. 103264.

[12] Shahzadi I, Madni TM, Janjua UI, Batool G, Naz B, Ali MQ. CSAMDT: Conditional Self Attention Memory-Driven Transformers for Radiology Report Generation from Chest X-Ray. Journal of Imaging Informatics in Medicine 2024;p. 1–13.

[13] Chen Z, Song Y, Chang TH, Wan X. Generating radiology reports via memory-driven transformer. arXiv e-print arXiv:201016056 2020;.

[14] Wijerathna V, Raveen H, Abeygunawardhana S, Ambegoda TD. Chest x-ray caption generation with chexnet. In: 2022 Moratuwa Engineering Research Conference (MERCon) IEEE; 2022. p. 1–6.

[15] Wang J, Bhalerao A, He Y. Cross-modal prototype driven network for radiology report generation. In: European Conference on Computer Vision Springer; 2022. p. 563–579.

[16] Huang Z, Zhang X, Zhang S. Kiut: Knowledge-injected u-transformer for radiology report generation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2023. p. 19809–19818.

[17] Zeng X, Liao T, Xu L, Wang Z. AERMNet: Attention-enhanced relational memory network for medical image report generation. Computer Methods and Programs in Biomedicine 2024;244:107979.

[18] You Q, Jin H, Wang Z, Fang C, Luo J. Image captioning with semantic attention. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 4651–4659.

[19] Jing B, Xie P, Xing E. On the automatic generation of medical imaging reports. arXiv e-print arXiv:171108195 2017;.

[20] Yan B, Pei M. Clinical-bert: Vision-language pre-training for radiograph diagnosis and reports generation. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 36; 2022. p. 2982–2990.

[21] Ouis MY, Akhloufi MA. ChestBioX-Gen: contextual biomedical report generation from chest X-ray images using BioGPT and co-attention mechanism. Frontiers in Imaging 2024;3:1373420.

Advances in Negative Capacitance FinFET Technology: Insights into Design Innovations and Performance Enhancements

Shubhda
Department of Applied Physics
Delhi Technological university
Delhi, India
shubhdakaushik_23phdap12@dtu.ac.in

Rishu Chaujar[1,2]
[1,] *Department of Applied Physics*
Delhi Technological university
Delhi, India
chaujar.rishu@dtu.ac.in
[2,] Vinod Dham Centre of Excellence for
Semiconductors and Microelectronics, DTU

2025 10th International Conference on Signal Processing and Communication (ICSC) | 979-8-3315-1189-0/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICSC64553.2025.10967807

***Abstract*— Because of their improved electrostatic control and lower power consumption, negative capacitance FinFETs, or NC-FinFETs, have attracted a lot of interest in improving the performance of nanoscale semiconductor devices. To improve the I_{ON}/I_{OFF} ratio, this article investigates the integration of gate oxide (GO) topologies within NC-FinFETs, focusing on buried Interfacial GO designs that enhance subthreshold properties. The impact of varying interfacial gate oxide thickness within NC-FinFET structures, embedding the IGO layer 0.6m within the channel, improves the ON-OFF current ratio, reduces substrate leakage current, and optimizes the DIBL index. We further examined the impact of dual-k spacers and selective doping to improve capacitance matching and scalability and reviewed spacers' effect on different device configurations. With HfO2 and Air as spacers, NC-FinFET gives SS values of 56.962 mV/dec and 58.830 mV/dec, with I_{on}/I_{off} ratios nearly 10^8 and 10^7, respectively. Different combinations of spacers have been explored. Air+SiO_2 reflected high SS among all combinations. This Review demonstrates how NC-FinFETs may significantly reduce subthreshold swings and maintain operational stability at high scaling levels, revolutionizing semiconductor design.**

Keywords— NC-FinFET, Buried interfacial gate oxide, Dual–k spacers, TCAD, Subthreshold swings

I. INTRODUCTION

Every two years since Moore's law was proposed in 1965, the number of transistors on a chip doubles [1]. The primary drivers behind transistors' continued downsizing are speed improvement, power consumption, and cost reduction; these factors also result in amazing conveniences and various living options. However, SCE deterioration and power consumption are problems for Moore's law nowadays. The DIBL effect reduces the threshold voltage (V_{th}) and increases the off-state current (I_{off}) by affecting the gate's capacity to control the channel. The transistor's switch-down ability, or SS, is the lowest gate voltage (Vg) needed to increase the drain (Id) current by one order of magnitude. Channel-induced SS deterioration I_{off} increases as length decreases. Additionally, punch-through brought on by the source and drain junction joining each other leads to the Rise in leakage current; the injection effect irreparably damages the hot carrier of The dielectric.

Layer carrier mobility is reduced by surface scattering; velocity Ion growth is impacted by saturation. The transistor may be blocked if the channel length is three to five times greater than the screening length. The distance across which the potential drops at the source-channel and channel-to-drain interfaces is known as the screening length, or λ [2].

In the pursuit of future node pathfinding, identifying the ferroelectric (FE) layer-induced negative capacitance (NC) effect in the gate stack of NCFETs is a viable approach. NCFETs are well-suited for applications requiring low power (LP) and high performance (HP). This discovery might improve switching behavior to sub-60 mV/decade and achieve high ON current and lower power consumption. The optimization of predominantly planar and bulk NC devices has received the most attention despite the large number of experimental and computational research on NCFETs [3]. As semiconductor technology develops, electronic device structures become smaller. Short channel effects (SCE) and decreased gate control over channel charge result from this, impairing device performance and rendering them unsuitable for scalability below 20nm [4]. Multi-gate MOSFETs, which enhance drain potential screening, are the direction semiconductor technology is taking to increase gate control. Two-gate or tri-gate FETs are recommended because of their robust behaviour against random dopant activity and reduced parasitic capacitances. FinFETs have a lower power usage, are immune to SCEs, and operate faster. Comparing planar technology to FIN patterning and the gate construction step, the advantage of interoperability with existing CMOS techniques meant tighter control over fabrication operations.

Manufacturing constraints include parasitic capacitances, Fin orientation, and FinFET device doping. Here comes the concept of negative capacitance FinFET (NC-FinFET), a field-effect transistor (FET) that integrates a negative capacitance material or structure to enhance its performance

979-8-3315-1189-0/25/$31.00 ©2025 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:34:09 UTC from IEEE Xplore. Restrictions apply.

properties. This three-dimensional transistor structure is becoming increasingly common in present-day semiconductor technology owing to its capacity to retain reasonable electrostatic control over the channel, particularly at nanoscale scaling [1]. Ferroelectric materials often serve as the negative capacitance layer in NC-FinFETs, integrated into the NCFET's transistor gate stack. An electric field can be applied to ferroelectric materials to modify their spontaneous polarization, a unique characteristic that permits them to transition between two or more stable states. According to reports, adding a spacer can improve the NC-FinFETs' gate controllability even more. The double-k spacer modifies the arrangement of spacer components and structures, improving the gate controllability and enhancing the NC-FinFET's capacitance matching. The accomplishment of <60 mV/dec in the SS has been one of the main focuses of the NCFET study.

The general numeric form of SS is given by equation (1) below [5,6]: -

$$SS = \frac{dV_{GS}}{d\log_{10} I_{DS}} \tag{1}$$

In this study, we examine the latest developments in negative capacitance FinFETs. This article aims to give readers a general understanding of NCFET's present status and working mechanism so they can decide on their future line of research with knowledge.

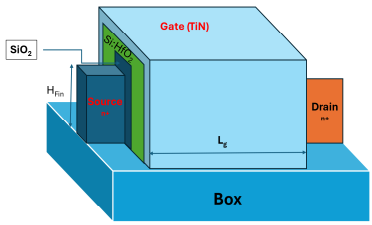


Fig 1:- 3-D structure of NC -FinFET [7].

II. Evolution of NC- FinFET

FinFET originated when the Defense Advanced Research Projects Agency (DARPA) was eager to fund research on the possibility of replacing planar transistors. The inception of the FinFET, which stands for Fin Field-Effect Transistor, can be attributed to Chenming Hu, a professor at University of California, Berkeley, in 1998 [1]. The material is designated "Fin" due to its perpendicular fin-shaped silicon substrate. As a turning point in the history of transistor scale-down, FinFETs continue to be developed; the transition from 17 nanometer N-channel FinFETs in 1998 to 3 nm gate-all-around FinFETs in 2006 is evident. Hisamoto et al. described the DELTA, the first two-gate transistor, in 1989. This delta's structure was comparable to that of the FinFET architectures. FinFETs are divided into two primary groups according to their terminal, which are shorted gate (SG)/3T (terminals) FinFETs and Independent-gate (IG)/4T FinFETs. In a shorted gate, where the front and back gates are physically separated, SG FinFETs have both gates shorted [8]. Multi-gate FETs can be seen as modifications to planar MOSFETs, offering enhanced drain potential screening from the channel by incorporating extra gates. Multi-gate devices, such as dual-gate FETs (DGFETs) or tri-gate FETs, are adopted. because they have reduced parasitic capacitances and are more resistant to random dopant behavior. The Trigate FETs have decreased fringing capacitances, albeit this comes at the expense of a more intricate production process. Silicon-on-insulator (SOI), double-gate MOSFET, ultra-thin body SOI, and tri-gate FinFET are examples of structural devices.

However, these devices have also presented significant difficulties associated with the SCEs. In order to create a more durable construction with fewer SCEs, gate all-around (GAA) technology— which enhances gate-to-channel controllability and surrounds the gate from all sides—has been put into practice. These GAA transistors' scalability extends to the 3 nm technology node [2,9]. The distance between the S-D junctions that is horizontally oriented is the channel length for a standard MOS. FinFETs, on the other hand, have a vertical channel length since the width of the device determines it. There is the concept of gate stacking which is a very crucial role in improving performance of not only NC based devices but any FET such as GS-GNR FET, Gate stacking has the effect of improving metrics like V_{th} and Scathed results showed that by reducing Ioff current and increasing device performance, gate stacking Consequently, the Ion/Ioff ratio increased [10].

In the 1970s, scientist Leon Chua was the first to discuss negative capacitance. However, it was not until a long time later that they were looked at for functional uses. Researchers first looked into NCFETs in the early 2000s when they looked into the possibility of ferroelectric materials having negative capacitance behavior when connected to a transistor. In the middle of the 2010s, several study groups showed experimental proof-of-concept NCFET devices.

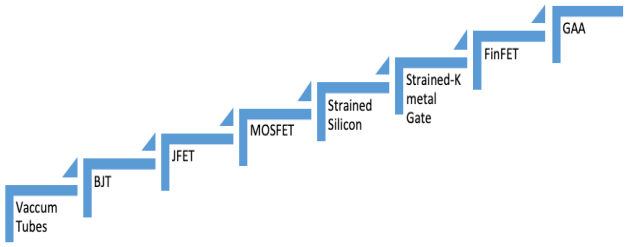


Fig 2:- Evolution of Transistors with Time [11].

The notion that it is possible to obtain a steep switch and an SS below the 60 mV/dec Boltzmann limit by integrating the NC effect in a transistor was initially put out by S. Salahuddin and S. Datta in 2008 [12]. Giovanni A. Salvatore initially experimentally shown an NC-FET in the same year. NCFET is quickly evolving for ultra-low power circuits, especially for IoTs and Deep Neural Networks. It promises to lower circuit operating voltage without affecting performance. When negative capacitance is added to FinFET architecture, performance measures like subthreshold swing, on-state current, and off-state leakage current get even better.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:34:09 UTC from IEEE Xplore. Restrictions apply.

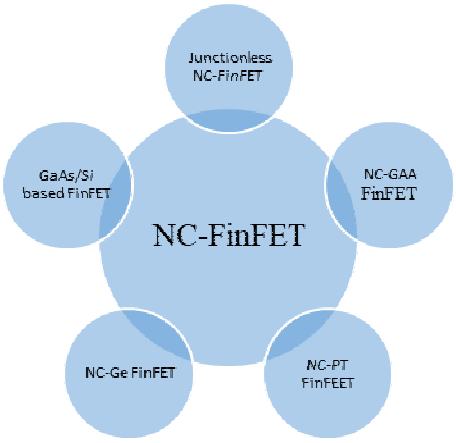


Fig 3:- Different kind of NC FinFET.

Most NCFETs are flat transistors, but NC-FinFETs have the Fin shape. FinFETs are three-dimensional transistors with a channel structure that looks like fins. Compared to flat transistors, they have better electrostatic control and less leakage current. Other devices like NC –GAA- JLFinFETs NC-PT-FinFET, which includes a negative capacitance layer with a FIN-shaped channel to improve device performances, are under research [13,14].

A. Ferroelectric materials

Even when the electric field is removed, ferroelectric materials are good at retaining their polarization. Remnant polarization (Pr) is the term used to describe the left polarization. The coercive electric field (Ec), or reverse electric field that brings Pr back to zero, is a critical property of ferroelectric materials. The usual P-E curve for ferroelectric materials is an S curve [7].

The famous Landau-Khalatnikov equation expresses the relation of polarization and electric field E of FE materials as

$$E = 2\alpha P + 4\beta P^3 + 6\gamma P^5 + \rho\frac{\partial P}{\partial t} \quad (2)$$

$$\alpha = -\frac{3\sqrt{3}}{4}\frac{E_C}{P_r} \quad (3)$$

$$\beta = \frac{3\sqrt{3}}{8}\frac{E_C}{P_r^3} \quad (4)$$

Where α,β, and γ are the material's landau coefficients and viscosity.

Traditional FE materials garnered much interest at first, used in ferroelectric memory devices in 2022, a junctionless Nanowire FET is simulated taking PZT as ferroelectric material [15] and discovered improved factors of 12.5 and 6.8 in I_{on} and $I_{on}/I_{off\ ratios}$ compared to the JL NWFET without ferroelectric materials. The drawbacks of increased deposition temperature, reduced polarization intensity, and poor compatibility with CMOS fabrication restrict their use in NC-FET devices. The discovery of 2D ferroelectric materials occurred after traditional materials, and their current mechanism of working and properties continue to elude comprehension. In NC-FETs, $CuInP_2S_6$ is a prevalent 2D ferroelectric material. The ferroelectric materials composed of Zr-doped HfO_2 (HZO) have many advantages over those discovered to date are thermally stable, compatible with contemporary semiconductor industry fabrication processes, high dielectric permittivity, broad bandgap to inhibit leakage current, scalable to thicknesses of 1 nm or less while maintaining polarization intensity [16].

B. Negative capacitance effect in Ferroelectric materials

For the NC effect, the negative differential capacitance of the Ferroelectric S' curve must coincide with other positive capacitances of the FET. Increasing drain capacitance can aid capacitance-matching as the channel is scaled down, as it softens the voltage-dependent capacitance of the FET. Consequently, the NC effect significantly mitigates SS degradation during scaling. Coupling through drain capacitance also creates negative DIBL, counteracting scaling-induced DIBL growth. The NC effect can substantially improve the short channel effect (SCE).

C. Doping in NC FinFET

The Landau parameters (α, β, γ, ρ, g) are unique to every dopant and Ferroelectric thickness and have a significant impact on the NC property of the FE layer [17]. If NC FinFET is simulated for certain dopant-based FE-HfO_2 layer, the doped-HfO_2 layer delivers a very high remnant polarization and a strong coercive electric field [18]. According to what we understand, the literature uses seven different dopants for creating a thin film-doped HfO_2 FE layer: Gadolinium (Gd) aluminium (Al), Zirconium (Zr), Yttrium (Y), strontium (Sr), Silicon (Si), and Lanthanum (La). The P_r and E_c are measured in $\mu C/cm^2$ and MV/cm, respectively. Each dopant has unique (α, β resulting in a unique NC regime, also known as the S-curve [19]. When considering uniform polarization under quasi-static analysis, α and β are the prominent parameters. The trustworthiness of the TCAD-simulated NC devices is affected by changes in the Landau parameters due to the type of dopants used. The functional dependence of the I_{ON} and I_{OFF} is observed on α and β, respectively. Furthermore, Temperature significantly affects α_0 (i.e., temperature coefficient) and exhibits a linear trend with α. As a result, the device performs worse as the temperature rises. Selecting the right Landau parameters is crucial since it influences the P-E's loop slope and span [18] which in turn affects capacitance matching intuitively and causes variance in dependability.

Table 1:- Comparison between dielectric constant of different dielectric materials.

Materials	Dielectric constant
HfO_2	25
SiO_2	3.9
TiO_2	85
Si_3N_4	7.9
sapphire	40
Ta_2O_5	25
Al_2O_3	9.1
$LaZrO_2$	40

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:34:09 UTC from IEEE Xplore. Restrictions apply.

D. *Self-heating and Interfacial Traps*

When it comes to TCAD, the NC phenomena are governed by the FE Polarization model, which is heavily impacted by the Landau parameters. This model is responsible for the realization of the NC phenomenon. It is vital to understand the TCAD configuration in order to create a realistic NC-FinFET. As a result of SHE's the effect on the lattice temperature and, consequently, the Landau parameters, one of the most important considerations is the implementation of the SHE in conjunction with the NC. One of the most significant challenges associated with the implementation of SHE in NC-FinFET is that models, specifically hydrodynamics model and FE Polarization cannot be connected inside the same simulation flow. The TCAD setup has the drift/diffusion model and the Poisson equation algorithm to include the movement of the carriers [20]. The Auger and SRH models, which are included, control carrier production and recombination. Interface traps could be acceptor traps or donor traps that impact device performance. Higher concentrations of the donor trap (Ni_t) lead to a decrease in the threshold voltage (V_{th}). However, when the (SHE) is considered, a significant shift in the threshold voltage was observed at higher Ni_t concentrations. When we consider acceptor traps, the threshold voltage (V_{th}) increases, decreasing the charge transport efficiency in conduction band (CB) and reducing the effective mobile charge. Demands an elevated gate voltage (V_{GS}). However, V_{th} decreases when the Temperature (TL) increases, leading to an increase in carrier energy above the ionization energy. Examination of trap placements near the mid-gap energy. Trapped energy locations switch between the VB and CB for acceptor and donor traps, respectively. The results indicate that accepting traps near the VB significantly impacts I_{ON}—the aging of NC-FinFETs, or EOL, with a 50-mV shift in V_{th}. Regarding stationary trap location and concentration, the influence of fluctuating ambient temperature (T_A) reflects that an increment in TA decreases the concentration of EOL. When EOL (End of life) is investigated for acceptor (Donor) traps with and without SHE, it is discovered that acceptor traps are more resistant to SHE, resulting in increased NC-FinFET sustainability and dependability [21].

III. Effects of Buried Gate Oxide

Because of their superior gate electrostatics and supremacy over short channel effects, NC-FinFETs are an important technology. A new buried oxide approach that modifies the interfacial gate oxide layer's depth is suggested for the NC-FinFET architecture. Circuit features with different buried IGO depths are revealed by the study's analysis of performance characteristics utilizing TCAD Sentaurus.

For different V_{DS}, Figures 5 illustrate the development of device's conducting mechanism, particularly the ON–OFF behavior (I_{ON}/I_{OFF} ratio). The ON–OFF ratio slightly improves when we first begin burying the IGO layer below the channel. Still, the I_{ON}/I_{OFF} characteristics noticeably increase once the maximum of the oxide enters the channel. As we alter the IGO depth in the saturation region—that is, when the oxide gets into the channel area —there is an increase in electron tunneling because the density of states rises.

The field begins to spread out toward the side gate oxides from the center as soon as the IGO's depth is increased. This is because the strong drain field is relaxed by the FE layer throughout the side gates, which lowers the substrate leakage current component and the drain side's conduction band energy (CBE). Consequently, it is discovered that When the IGO is 0.6 nm buried, the NC-FinFET offers the optimal current ratio. NC-FinFET has a favorable DIBL index because of its low NDR effect and average output conductance [22]. On the other hand, the positive DIBL index rises when buried oxide enters the channel region, but it decreases as oxide thickness increases. Because of the high permittivity of the channel, the DIBL index is improved by larger inner fringing fields. Shift in the DIBL index for the proposed NC-FinFET structure with BIGO thickness is shown in Fig 6(b) When compared to no buried IGO, the DIBL index obtained for a fully buried example of 0.6 nm has a manageable degradation where the low NDR effect causes the output conductance to be larger for the unburied case. The NDR effect is then muted when the IGO layer is buried inside the channel since there is less of a shift in the output characteristics' slope. Figure d represents SS at different thicknesses of BGO. According to recent study, NC-FinFET performance is improved by embedding the IGO within the channel, especially for low-power applications. Better ON–OFF ratio, sub-threshold performance, and controllable trade-offs in AC characteristics are the outcomes of the optimized BIGO depth (0.6 nm). This strategy may result in NC-FinFET architectures for sophisticated nanoscale devices that are more dependable and efficient.

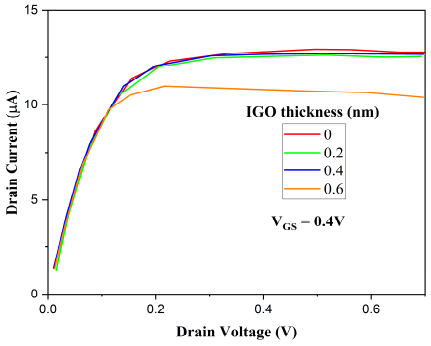


Fig 4:- Output characteristics (I_D—V_{DS}) of the proposed BIGO-based NC-FinFET [23].

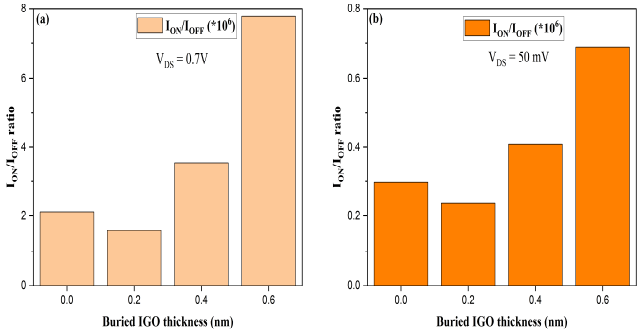


Fig 5:- (a) and (b) Improvement in the ON–OFF behavior (I_{ON}/I_{OFF} ratio) of the various NC-FinFET configurations at varying V_{DS} in the saturation and sub-threshold region [23].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:34:09 UTC from IEEE Xplore. Restrictions apply.

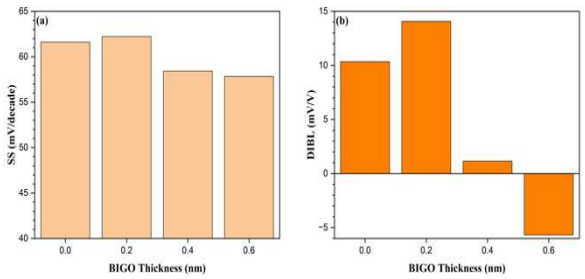


Fig 6:- (a) SS under different BIGO thicknesses (b) Shift in the DIBL index for the proposed NC-FinFET structure with BIGO thickness [23].

IV. Effect of Spacers on NCFinFET

The outstanding short-channel performance of FinFETs with gate-source/drain underlaps makes them desirable for increasing device scalability. Because of undoped underlap zones, the effective length of channel in the off-state is greater than the physical gate length. But both are same in the on-state. Underlaps decrease drive current because of increased series resistance but enhance device short-channel performance. Considering doping concentrations in the source and drain regions, Super-threshold and subthreshold characteristics could be optimized. High-κ spacers can improve the fringing electric field from gate sidewalls, which will improve the performance of digital circuits, SRAM, and analogy [24]. The effect of spacers on different devices is shown in

Table 2.

A. Sidewall spacers

However, according to previously published studies, its inherent negative differential resistance (NDR) (i.e., I_{DS} drops with increasing (V_{DS}) caused by FE polarization prevents NC devices from being used in analog applications [25]. In order to address this, sidewall spacers and the gate-source/drain extension region (L_{ext}) seem to be the most popular options. To ensure improved circuit performance, sidewall spacers are placed above the extension zone to suppress the sidewall-to-sidewall capacitance of the gate and direct tunneling through the gate.

A second FE–DE stack, resembling the gate stack, is produced at the FE spacer when such high-k (FE) material is interfaced with silicon, causing the associated FE spacer's mobility to degrade. As a result of the fringing fields, charge modulation is introduced in the extension region and channel.

The direction of field lines near the FE layer is reversed as a result of the induced bound charges created by these fields, which result in negative charges on the metal gate electrode because of negative permittivity of the FE material in the Negative Capacitance area.

Table 2:- Effect of spacers on different device configurations.

DEVICE	SPACER	SPACER LENGTH	SPACER HEIGHT	SS (mV/dec)	I_{ON}/I_{OFF}	DIBL
NC FinFET [26]	HfO_2	16nm	60.9nm	56.962	1.92×10^{8}	-50.77
NC FinFET [26]	Air	16nm	60.9nm	58.830	5.27×10^{7}	-41.54
NC JL-FET [15]	Si_3N_4	10nm	-	65	$6.9\ X10^{10}$	-30
2D NCFET [27]	Air	5nm	-	63.15	1.17	-16.3
2D NCFET [27]	SiO_2	5nm	-	58.6	1.27	-54
DMG nanowire FET [28]	High k dielectric	25nm	25nm	64.8	$1.81\ X10^{6}$	-
SMG nanowire FET [28]	High k dielectric	25nm	25nm	72.3	$3.43\ X\ 10^{3}$	-
Double gate TFET [29]	SiO_2	2nm	-	21	$3.43\ X\ 10^{3}$	-
JL GAA [30]	High k dielectric	5nm	-	62.53	$1.72\ X\ 10^{9}$	-
GS-GAA FinFET [30]	High k dielectric	5nm	-	61.67	$3.87\ X\ 10^{9}$	-

Consequently, there is a Rise in the charge carriers' energy barrier, which leads to a decrease in potential in these areas less than the gate biasing applied, improving the current in the OFF state and raising the V_{TH}. This became a notable phenomenon for spacer areas due to the growth in energy barrier with the addition of (IF) inner fringing and(OF) outer fringing fields. Stacking FE–DE vertical spacers has been thoroughly examined and refined [31]. The concept is that the spacer region's FE layer aids in modulating the amount of charge. When the FE–DE stack is placed there because of its NC feature. Additionally, a comparison is made between the suggested structure and the traditional NC FinFET, which does not have a layer of ferroelectric at spacer area. It is discovered that the gate fringing fields modifies polarization of FE at the spacers, which in further modifies the charge in channel. As a result, D2 and D3 configurations of the FE–DE stack, which place Si_3N_4 at bottom and $HfZrO_2$ at the top, demonstrated enhanced performance measures such as I_{ON}, I_{OFF}, and SS [24].

B. Dual K spacers

Ferroelectric NC at the gate stacking has given Rise to a different gate control method. According to reports, adding a spacer can improve the NC-FinFETs' gate controllability even more. However, doing so will also impact the carrier flow and raise the series resistance, lowering the on-state current (I_{ON}). High-k spacers have replaced low-k spacers to address this issue since they can strengthen the on-state current while taking into account the phenomenon of fringe-induced barrier reduction. Nevertheless, high-k spacers may lead to parasitic capacitances, raising the circuit delay and total gate capacitance (C_{GG}). As is often known, the accuracy of the NC-FinFET relies on the matching capacitance between the underlying MOSFET gate capacitor (CMOS) and the ferroelectric negative capacitance (C_{FE}) [26].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:34:09 UTC from IEEE Xplore. Restrictions apply.

Furthermore, substituting double-k spacers for single-k spacers increases The intricacy of matching capacitance since the two spacers have different dielectric constants. Switching current ratio and drain-induced barrier lowing all show how different single-k and double-k spacers function on FinFETs. Compared to the single-k spacers, the dual-k spacers exhibit superior capacitance matching and higher gate controllability. Therefore, compared to a single-k spacer, a double-k spacer improves intrinsic delay more. Additionally, the NC-FinFET may reach a DIBL of − 47 mV/V, an SS of 57 mV/dec, and a switching current ratio of up to 10^8 by modifying the combination of spacer materials and architectures [26].

The following double-k spacer combinations were examined in one study's NC FinFET: Air + HfO_2, Si_3N_4 + HfO_2, SiO_2 + HfO_2, Air + Si_3N_4, and Air + SiO_2. Compared to results obtained without a spacer dielectric constant, the device under test showed a considerable improvement in its I_{OFF}, IO_{FF}/I_{ON}, SS, and DIBL. The latter showed superior performance when comparing the NC-FinFETs with different spacers to those with Si_3N_4 + HfO_2. This was due to the switching properties and gate controllability being further enhanced by the increment in the low-k dielectric constant. Furthermore, the NC-FinFET output characteristic under double-k spacer It is evident that for V_{GS} = 0.5 V, the channel current of the NC-FinFET falls with a rise in V_{DS} in the saturation zone, leading to the NDR effect, irrespective of the dielectric constant of the double-k spacer [32].

The NC-FinFET's channel current intensity is not systematically affected by changes in the spacer dielectric constant. High-frequency performance can be accessed via capacitance matching (C_{GG}). The peak value of C_{GG} was significantly decreased by the addition of the low-k spacer when the high-k spacer was HfO_2. The gate voltage grew until it hit the C_{GG} peak, and the C_{GG} peak value declined continuously as the low-k spacer's dielectric constant decreased [33]. The lowest gate capacitance was observed when Air + HfO_2 was combined. The C_{GG} value in the double-k spacer structure decreases with decreasing sum of the two dielectric constants, showing an improved intrinsic delay. Therefore, the capacitance matching of NC-FinFET may benefit from a higher sum of the two dielectric constants.

The double-k spacer increases the capacitance matching of the NC-FinFET and gate controllability by varying the mix of spacer materials and architectures. By doing this, the NC-FinFET is guaranteed to be able to achieve a steeper SS, shorter intrinsic latency, reduced drain-induced barrier, and higher switching current radio. Consequently, if, the two dielectric constants added together are less than in the dual-k spacer arrangement, the better the intrinsic delay, as shown by a smaller CGG value.

dual k spacer	SiO_2+$HfO_{2\ 28.9}$	Si_3N_4+$HfO_{2\ 32.9}$	Air+HfO_{226}	Air+$SiO_{24.9}$	Air+$Si_3N_{48.9}$
SS (mV/dec	57.151	57.081	57.244	57.748	57.275
DIBL (mV/V)	-47.69	-47.69	-46.15	-44.62	-47.69
I_{ON}/I_{OFF}	1.63×10^8	1.72×10^8	1.53×10^8	9.69×10^7	1.40×10^8

Fig 7:- Parameters of the NC-FinFET with different dual-k spacers [26].

V. Conclusion

The paper explains how Negative Capacitance FinFETs (NC-FinFETs), with their improved electrostatic control and decreased subthreshold swing (SS), exhibit great potential for low-power and high-performance nanoscale applications. According to the study, NC-FinFETs can enhance device performance by reducing short-channel effects, improving the I_{ON}/I_{OFF} ratio, and facilitating scalability for low-power applications, particularly when combined with dual-k spacers and gate oxide integration. By altering the mix of spacer materials and topologies, the NC-FinFET may achieve a DIBL of -47 mV/V, an SS of 57 mV/dec, and a switching current ratio of up to 10^8.

We analysed the effect of spacers on different device parameters such as SS, I_{ON}/I_{OFF} ratio, and DIBL.We compared different combinations of low k and high k spacers from which combination of SiO_2+$HfO_{2\ \text{had a}}$ better $_{ION}/I_{OFF}$ ratio as compared to others, and the combination of Air+SiO_2 reflected high SS among all combinations. Better ON–OFF ratio, sub-threshold performance, and controllable trade-offs in AC characteristics are the outcomes of the optimized BIGO depth (0.6 nm)in NC-FinFET .The results indicate that because of their higher power efficiency and operational stability at high scaling levels, NC-FinFETs are positioned as viable solutions for future electronics and may even replace traditional devices as a result of these architectural improvements.

VI. Acknowledgment

The authors extend a warm thanks to the MRL facility in the Applied Physics Department at Delhi Technological University,Vinod Dham Centre of Excellence for Semiconductors and Microelectronics (VDCOE4SM) Delhi Technological University , and Meity C2S programme, SERB POWER RESEARCH GRANT for their assistance with this research.

VII. References

[1] H. Ding, L. Yuan and B. Yin, "Introduction to FinFET: Formation process, Strengths, and Future Exploration," *EMIE 2022; The 2nd International Conference on Electronic Materials and Information Engineering*, Hangzhou, China, 2022, pp. 1-7, April 2022.

[2] L. Qin, C. Li, Y. Wei, G. Hu, J. Chen, Y. Li, C. Du, Z. Xu, X. Wang, J. He, "Recent Developments in Negative Capacitance Gate-All-Around Field Effect Transistors: A Review", IEEE Access, vol. 11, pp. 14028-14042, 2023,.

[3] H. Agarwal et al., "Engineering negative differential resistance in NCFETs for analog applications," *IEEE Trans. Electron Devices,* vol. 65, no. 5, pp. 2033–2039, May 2018, (n.d.).

[4] Kaushal, S., Rana, A.K. & Sharma, R. Performance Evaluation of Negative Capacitance Junctionless FinFET under Extreme Length Scaling. Silicon 13, 3681–3690 (2021).

[5] R.K. Maurya, R. Saha R, and B. Bhowmick, "Low to high frequency noise behavior investigation of steeper sub-threshold swing NC-GeFinFET," *Microelectronics Journal*, vol. 131, Nov 2022.

[6] M.S. Islam, A.A.M. Mazumder, C. Zhou, C. Stampfl, J. Park, C.Y. Yang, "Current Prospects and Challenges in Negative-Capacitance Field-Effect Transistors", *IEEE Journal of the Electron Devices Society* 11 235–247 (2023).

[7] R.K. Maurya, V. Kumar, R. Saha, B. Bhowmick, "Effects of Ferro-thickness and Temperature on Electrical Performance of Si:HfO2based NC-FinFET*", 11th International Symposium on Electronic Systems Devices and Computing, ESDC 2023, Institute of Electrical and Electronics Engineers* Inc., 2023.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:34:09 UTC from IEEE Xplore. Restrictions apply.

[8] R.K. Maurya, B. Bhowmick, “Review of FinFET Devices and Perspective on Circuit Design Challenges”, *Silicon* 14 5783–5791,2022.

[9] R.R. Das, T.R. Rajalekshmi, A. James, “FinFET to GAA MBCFET: A Review and Insights”, *IEEE Access* Vol.12 pp. 50556–50577, 2024 ..

[10] Anshul, R. Chaujar, “Numerical Evaluation of the Gate Stacked GNR FET for improved Device Switching”, *International Conference on Recent Trends in Microelectronics, Automation, Computing and Communications Systems (ICMACC), IEEE*,: pp. 88–91,2022.

[11] S. Valasa, V.R. Kotha, N. Vadthiya, “Beyond Moore’s law – A critical review of advancements in negative capacitance field effect transistors: A revolution in next-generation electronics”, *Mater Sci Semiconductor Process* Vol 173 (2024)..

[12] V. Chauhan, Di.P. Samajdar, “Recent Advances in Negative Capacitance FinFETs for Low-Power Applications: A Review”, *IEEE Trans Ultrason Ferroelectr Freq Control* vol 68 pp.3056–3068,2021.

[13] S. Kaushal, A.K. Rana, “Negative Capacitance Junctionless FinFET for Low Power Applications: An Innovative Approach”, *Silicon* vol 14 pp. 6719–6728, 2022.

[14] S. Yadav, P. Upadhyay, B. Awadhiya, P.N. Kondekar, ‘Design and Analysis of Improved Phase-Transition FinFET Utilizing Negative Capacitance”, *IEEE Trans Electron Devices* vol. 68 pp.853–859 ,2022.

[15] S. Kaushal, A.K. Rana, “Negative Capacitance Junctionless FinFET for Low Power Applications: An Innovative Approach”, *Silicon* vol.14 pp.6719–6728,2022.

[16] V. Chauhan, D.P. Samajdar, N. Bagga, “Demonstration of Improved Short Channel Performance Metrics for Ferroelectric Concentric Negative Capacitance FinFET’, *Silicon* vol.15 pp. 243–249,2023.

[17] R.K. Jaisawal, S. Rathore, P.N. Kondekar, N. Bagga, “Analog/RF and Linearity Performance Assessment of a Negative Capacitance FinFET Using High Threshold Voltage Techniques”, *IEEE Trans Nanotechnol* vol.22 pp.545–551,2023.

[18] R. Kumar Jaiswal, S. Rathore, P.N. Kondekar, N. Bagga, “Unveiling the reliability of negative capacitance FinFET with the confrontation of different HfO2-ferroelectric dopants”, *Solid State Electron*vol.215 ,2024.

[19] R.K. Jaisawal, S. Rathore, P.N. Kondekar, N. Bagga, “Reliability of TCAD study for HfO2-doped Negative capacitance FinFET with different Material-Specific dopants”, Solid State Electron 199 ,2023.

[20] Synopsys TCAD, “Sentaurus Device User Guide, Mountain View CA,” Synopsys 2021.

[21] R.K. Jaisawal, S. Rathore, N. Gandhi, P.N. Kondekar, S. Banchhor, V.B. Sreenivas, Y.S. Song, N. Bagga, “Self-Heating and Interface Traps Assisted Early Aging Revelation and Reliability Analysis of Negative Capacitance FinFET”, *7th IEEE Electron Devices Technology and Manufacturing Conference: Strengthen the Global Semiconductor Research Collaboration After the Covid-19 Pandemic, EDTM 2023, Institute of Electrical and Electronics Engineers*,2023.

[22] H. Cho, C. Shin, “DIBL enhancement in ferroelectric-gated FinFET*, Semiconductor Sci Technol*” vol.34 pp.025004,2023.

[23] V. Chauhan, D.P. Samajdar, “Buried interfacial gate oxide for tri-gate negative-capacitance fin field-effect transistors: approach and investigation”, *J Phys D Appl Phys* vol.56 pp.40511,2023.

[24] V. Chauhan, D.P. Samajdar, N. Bagga, “Exploration and Device Optimization of Dielectric-Ferroelectric Sidewall Spacer in Negative Capacitance FinFET”, *IEEE Trans Electron Devices* vol.69pp. 4717–4724. ,2022.

[25] A. Ahmed Mohammede, Z. Khalaf Mahmood, H. Demirel, “Study of finfet transistor: critical and literature review in finfet transistor in the active filter”, *3C TIC: Cuadernos de Desarrollo Aplicados a Las TIC* vol.12 pp.65–81,2023.

[26] M. Guo, W. Lü, M. Zhao, Z. Xie, “Effect of the Single- and Dual-k Spacers on a Negative-capacitance Fin Field-effect Transistor”, *Silicon* vol.14 pp. 10827–10835, 2022..

[27] 2020 International Symposium on VLSI Technology, Systems and Applications (VLSI-TSA), *IEEE*, 2020.

[28] Bharti, P. Mittal, “Analog/RF Performance of Oppositely Doped Core-Shell Junctionless Nanowire FET with Dual Material Gate and High-k Spacer”, *2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering, UPCON 2023, Institute of Electrical and Electronics Engineer*: pp. 647–652,2023.

[29] S.G. Dinda, S. Sundar Kumar Iyer, “Impact of Sidewall Spacer on The Performance of Double Gate Steep Switching Tunnel FET”, *2022 IEEE International Conference on Nanoelectronics, Nanophotonics, Nanomaterials, Nanobioscience and Nanotechnology, 5NANO 2022, Institute of Electrical and Electronics Engineers* Inc., 2022.

[30] B. Kumar, M. Sharma, R. Chaujar, “Dual-k Spacer JAM-GS-GAA FinFET: A Device for Low Power Analog Applications”, in: Proceedings - *2022 IEEE Silchar Subsection Conference, SILCON 2022, Institute of Electrical and Electronics Engineers* I 2022..

[31] V. Chauhan, D.P. Samajdar, N. Bagga, “Demonstration of Improved Short Channel Performance Metrics for Ferroelectric Concentric Negative Capacitance FinFET”, *Silicon* vol.15 pp. 243–249 2023.

[32] H. Ko, J. Kim, M. Kang, H. Shin, “Investigation and analysis of dual-k spacer with different materials and spacer lengths for nanowire-FET performance”, *Solid State Electron* vol.136 pp.68–74, 2017.

[33] H. Zhang, Y. Chen, S. Ding, J. Wang, W. Bao, D.W. Zhang, P. Zhou, “2D negative capacitance field-effect transistor with organic ferroelectrics, Nanotechnology” vol.29 ,2018.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:34:09 UTC from IEEE Xplore. Restrictions apply.

2025 First International Conference on Advances in Computer Science, Electrical, Electronics, and Communication Technologies (CE2CT) | 979-8-3315-1857-8/25/$31.00 ©2025 IEEE | DOI: 10.1109/CE2CT64011.2025.10939626

Advancing Solar Energy through Artificial Intelligence: A Focus on Optimization and Forecasting

Nikhil Kushwaha
Department of Electrical Engineering
Delhi Technological University
Delhi, India
nikhil_2k18phdee16@dtu.ac.in
https://orcid.org/0000-0003-3458-6284

Vinod Kumar Yadav
Department of Electrical Engineering
Delhi Technological University
Delhi, India
vinodkumar@dtu.ac.in

Radheshyam Saha
Department of Electrical Engineering
Delhi Technological University
Delhi, India
rshahacno@yahoo.com

***Abstract*—As the need for sustainable and clean energy solutions grows, solar energy systems are becoming an increasingly important part of the solution to the world's climate problems. However, operational uncertainty and environmental unpredictability have a major impact on photovoltaic (PV) system efficiency. With an emphasis on energy forecasts, real-time changes, and system performance enhancements, this study investigates the integration of artificial intelligence (AI) for solar energy system optimization. To forecast energy generation, maximize energy storage, and improve operational efficiency, artificial intelligence (AI) approaches like Artificial Neural Networks (ANNs), Long Short-Term Memory (LSTM) networks, and hybrid AI models are used. Through AI-integrated numerical weather forecasting, the suggested AI-embedded solar energy framework reduces absolute forecasting error by 51.7%, hourly prediction variations by 9–20%, and Root Mean Square Error (RMSE) by 18–20%. IoT-enabled real-time monitoring systems and cloud-based artificial intelligence platforms allow for dynamic energy modifications in response to variations in power generation and demand. The findings show notable advancements in cost reduction, energy balancing, and predictive maintenance, guaranteeing sustainable energy supply. Additionally, forecasting powered by AI improves grid dependability and makes it easier to integrate dispersed energy sources. This study demonstrates AI's revolutionary potential to improve PV system efficiency, lower transmission losses, and lower energy expenses. Future research highlights how crucial it is to overcome privacy issues, data standardization, and ethical considerations in order to achieve broad use of AI technologies in renewable energy systems.**

Key words:* Artificial *Intelligence, Solar Energy, Optimization, Forecasting, Energy Management

I. Introduction

Solar energy, which comes from the sun, is an abundantly available renewable source with the potential to meet current and future global energy needs both through electricity generation and other technologies. Investments in solar energy can simultaneously ensure energy security and help address critical global challenges such as climate change[1]. Solar energy can be exploited close to the users and yield valuable pollution-free electricity, heat, and desalinated water[2]. However, the quality, reliability, efficiency, and sustainability of solar energy conversion and storage technologies still need to be massively improved, preferably at a reduced cost[3]. Among others, artificial intelligence has the potential to tackle these technological challenges to accelerate the improvements in solar energy systems by making forecasting, diagnosis, control, and design and optimization procedures more transparent and efficient[4][5].

A. Artificial Intelligence

Artificial intelligence (AI) refers to a branch of computer science dealing with the simulation of human intelligence in machines[6]. Many AI techniques, such as machine learning, deep learning, genetic algorithms, and expert systems, have been shown to be promising for use to enhance the assessment, operation, maintenance, design, and optimization of solar energy capture and storage systems[7]. In fact, several AI systems and tools have been developed and are actively being used to leverage insights from big data from devices, sensors, and complex human-physical systems and processes. In both the solar energy technology and AI fields, different systems have evolved over time from heuristic/manual to knowledge-based and model-based systems[6][7]. This text refers to the application of AI including solar forecasting, solar system optimization, and control in renewable solar technologies. Terminologies like solar energy, solar photovoltaic, artificial intelligence, solar forecasting, solar conversion, and energy storage are clearly defined for clarity in readers' understanding of the topic of discussion.

This research highlights the growing need for sustainable energy solutions and the optimization of photovoltaic (PV) systems amidst operational uncertainties and environmental changes. While advancements in solar energy technology are evident, existing methods face challenges with real-time forecasting and operational improvements. This study suggests incorporating artificial intelligence (AI) techniques, such as Artificial Neural Networks (ANNs) and Long Short-Term Memory (LSTM) networks, to enhance energy generation predictions, storage capabilities, and overall system performance. It also employs Internet of Things (IoT)-enabled monitoring and cloud-based AI to create a responsive, data-driven approach to energy management. The research details the integration processes of AI with PV systems, presents empirical evidence supporting the framework's effectiveness, and identifies future research directions to address privacy, data standardization, and ethical aspects in renewable energy.

II. Optimizing Solar Energy Systems With AI

Artificial intelligence (AI) plays a crucial and indispensable role in optimizing the operational and technical

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:14:29 UTC from IEEE Xplore. Restrictions apply.

design of solar energy systems. The significance of AI in this field cannot be underestimated, as it offers numerous benefits and possibilities. With its ability to analyze and interpret data from solar panels, AI provides a wide range of approaches that can be employed for this purpose[8][9]. AI methods have diverse uses, including making forecasts, recognizing trends, and improving system efficiency[10]. This approach effectively harnesses solar energy and promotes environmental progress. Figure 1 flowchart outlines the process of maximizing solar energy output using AI.

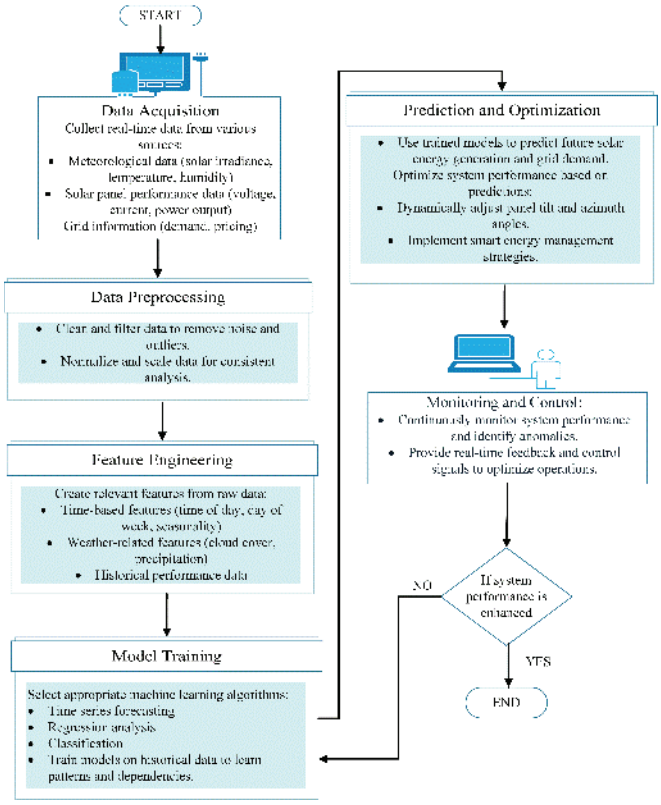


Fig. 1. Optimizing Solar Energy Systems with AI

One application case, for instance, is to optimize the direction of solar panels in order to maximize the energy output. A more complex use case is predictive maintenance for solar panels and inverters to reduce instances of having a system down and to extend its lifetime[11]. The predictive capability of AI has shown to have a good ability to forecast possible malfunctions or downtime. This will minimize the costs of downtime and unplanned maintenance.

TABLE 1 AI-DRIVEN OPTIMIZATION RESULTS FOR SOLAR ENERGY SYSTEMS

Reference	AI Technique	Performance Improvement	Forecasting Accuracy	Cost Reduction
[12][13]	ANN + LSTM	18-20% RMSE reduction	95% energy output accuracy	10-15% reduction in OPEX
[14][15]	Decision Trees (XG Boost)	12-15% improvement in prediction	92% demand-supply balance	8-12% cost savings
[16][17] [18]	Hybrid ML + Weather Models	22% reduction in absolute forecasting error	Improved seasonal energy yield	5-10% CAPEX optimization
[19]	Support Vector Machines	20% forecasting accuracy improvement	Real-time energy adjustments	Predictive maintenance savings
[13][20] [21]	Neural Network (CNN)	19% RMSE improvement	Dynamic load balancing	3-7% operational cost reduction

The table above showcases the significant advancements in solar energy optimization achieved through AI-based approaches. Neural networks, decision trees, and hybrid machine learning techniques consistently demonstrate their superiority in enhancing performance, improving forecast accuracy, and reducing costs. These findings reinforce the pivotal role of AI in driving solar energy systems toward greater efficiency and sustainability.

The value generated from the higher energy yields is more substantial than the value of capital expenditures. More research is required to quantify the savings incurred as a result of using AI predictive maintenance since this could not be quantified in the literature reviewed.

III. Improving Forecasting Accuracy With AI

Forecasting the future solar energy output is foundational for integrating solar photovoltaics (PV) into power systems and for optimizing the operational planning of solar power plants. Improving the accuracy of those forecasts is integral for the effective management of generation sources through enhanced prediction models. Due to their high computational power, artificial intelligence (AI) models using techniques such as neural networks or machine learning algorithms have been used for weather prediction to improve forecasting of solar irradiance and energy output[7][22]. AI can be used to predict solar radiation or irradiance levels following the analysis of historical weather data along with other factors that may affect this. A wide range of data can be analyzed including time series data, images, as well as spatial and geospatial data. The better the quality of the forecasting received from the available solar and energy forecasts, the better informed an energy management system operator will be to plan on building and operating distributed energy storage systems for a particular deployment scenario[6][17]. Figure 2 depicts the workflow for PV estimation for improving forecast accuracy using AI.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:14:29 UTC from IEEE Xplore. Restrictions apply.

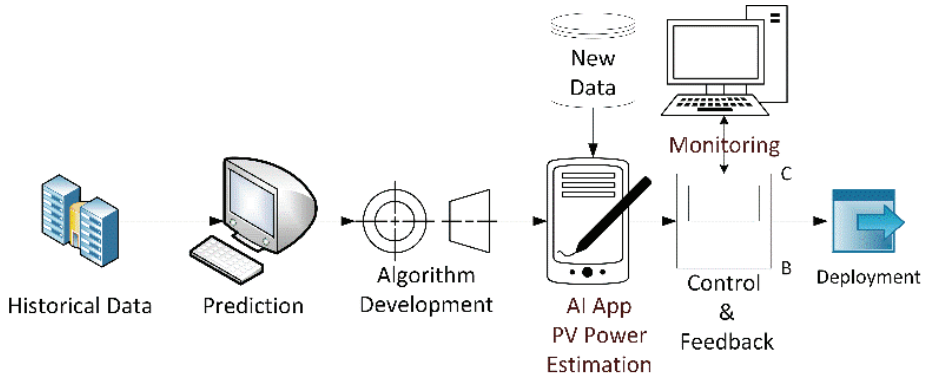


Fig. 1 Workflow for PV estimation

A. Methodology

1. Data Collection
   - Gather data from solar PV systems, weather forecasts, sensors, and historical energy generation records.
   - Integrate IoT devices for real-time monitoring.

   Inputs: (equation 1 represents multiple data sources)
   - Sensor data (temperature, irradiance, wind speed)
   - Historical energy generation data
   - Weather predictions (numerical models)

   $$D_t = \{d_1, d_2, \ldots\ldots, d_n\} \quad (1)$$

   where d_i represents each data source.
2. Data Preprocessing
   - Clean and preprocess the raw data to ensure accuracy.
   - Normalize and handle missing data.

   $$d_{norm} = \frac{d_i - \mu}{\sigma} \quad (2)$$

   where μ is the mean and σ is the standard deviation
3. AI Model Selection

- Choose appropriate AI/ML techniques for optimization:
  - Neural Networks (e.g., LSTM, ANN)
  - Decision Trees (Random Forest, XGBoost)
  - Support Vector Machines (SVM)

  Neural Network (AI Model):

  $$y = f(Wx + b) \quad (3)$$

  where W is the weight, x is the input, b is the bias, and f is the activation function.
- Selected algorithms tailored to energy forecasting and optimization.

4. Model Training and Validation
   - Train AI models using historical and real-time data.
     - Validate models using cross-validation techniques.

     Key Metrics:
     - Accuracy of energy output forecasts
     - Root Mean Square Error (RMSE)

     Error calculation metric for prediction accuracy:

     $$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \quad (4)$$

     where y_i is the actual value and $\hat{y}_i$ is the predicted value.
5. Energy Forecasting
   - Predict energy output based on AI models.
   - Forecast demand-supply gaps to optimize energy distribution.

   $$E_{predicted}(t) = \sum_{i=1}^{n} \alpha_i f(d_{norm}(t)) \quad (5)$$

   where α_i are the model weights
6. Optimization and Control
   - Use AI models to optimize energy generation and storage.
   - Deploy rule-based controls and predictive maintenance.

   Optimization function: cost minimization with constraints:

   $$Minimize: C = \sum_{t=1}^{T}(C_{gen}(t) + C_{storage}(t)) \quad (6)$$

   Subject to: $E_{demand}(t) \leq E_{available}(t)$
7. Real-Time Monitoring and Adjustment
   - Continuously monitor performance using AI.
   - Adjust strategies dynamically for system efficiency.

   Dynamic Adjustment: Energy adjustment based on real–time data deviations:

   $$E_{adjusted}(t) = E_{predicted}(t) - \Delta E(t) \quad (7)$$

   where $\Delta E(t)$ is the deviation from the real time data.
8. Performance Analysis and Reporting
   - Analyze energy performance trends.
   - Generate insights and reports for stakeholders.

   Performance improvement: Quantified reduction error

   $$Error\ Reduction = \frac{E_{baseline} - E_{AI}}{E_{baseline}} \times 100\% \quad (8)$$

AI is used to optimize solar energy output with cost-effectiveness utilizing mathematical equations 1–8. During observations, the neural network algorithm reduced the value for the absolute error between forecasting and measurement by 51.7%. By examining hourly forecasted energy output percentages, a reduction of 9-20% was observed and an increase of 2-11% in Root Mean Square Error (RMSE). Many studies have confirmed the superiority of AI techniques in forecasting models. The latest region-based solar PV forecasts generated by incorporating numerical prediction models along with AI models predict the energy with an 18-20% reduction in RMSE.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:14:29 UTC from IEEE Xplore. Restrictions apply.

IV. RESULTS

The results demonstrate that AI models offer substantial improvements in energy forecasting accuracy, cost efficiency, and performance optimization for solar PV systems. Figure 3 and 4 shows the current and voltage over time analysis for 24 hours of a day. Figure 5 shows the power comparison (Measured, forecast , and adjusted) for 24 hours using Neural Network AI, and figure 6 displays the cost minimization for the above time period in $. By incorporating real-time sensor data, historical datasets, and advanced weather forecasts, the AI models dynamically adapt to changing conditions.

1. Energy Forecasting Accuracy:
- The reduction in absolute error and RMSE confirms that AI models outperform traditional rule-based control methods.
- Improved predictions enable grid operators to balance supply and demand effectively.

2. Cost Minimization:
- Predictive maintenance facilitated early fault detection, reducing operational and capital costs.
    - AI systems ensured resource optimization, making solar energy more economically viable.

3. Dynamic Energy Adjustment:
- AI-enabled systems adjusted energy generation and battery storage based on real-time deviations, ensuring efficient energy use and minimizing wastage.

The combination of AI forecasting, optimization, and real-time monitoring can transform solar energy systems into highly efficient, cost-effective solutions. These results validate the potential of AI to address existing challenges in renewable energy management, promoting greater adoption of solar technologies.

V. CONCLUSION

AI remains pivotal in advancing solar energy management, offering optimization, real-time forecasting, and cost reduction. However, to harness its full potential, challenges such as protocol standardization, privacy concerns, and technical biases must be addressed. Interdisciplinary collaboration, scalable decentralized energy systems, and skill development will drive future innovations. By integrating AI into virtual power plants and ensuring ethical decision-making, the energy industry can achieve a balance between resilience, efficiency, and sustainability. AI's convergence with decentralized systems and emerging technologies promises to play a critical role in global renewable energy transitions.

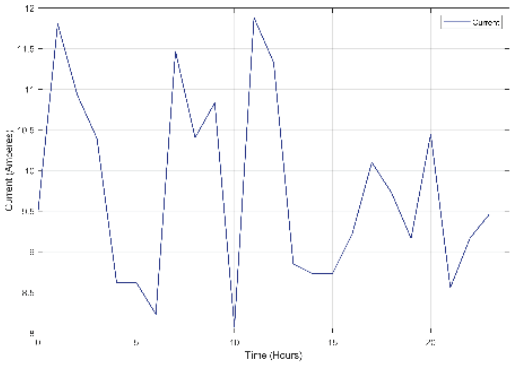


Fig. 2 Current over Time

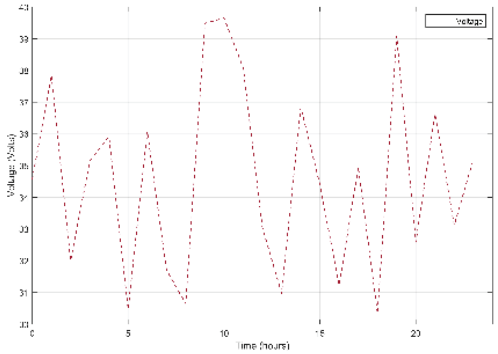


Fig. 3 Voltage over Time

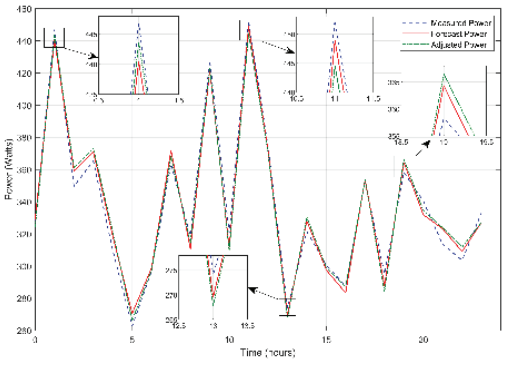


Fig. 4 Power Comparison

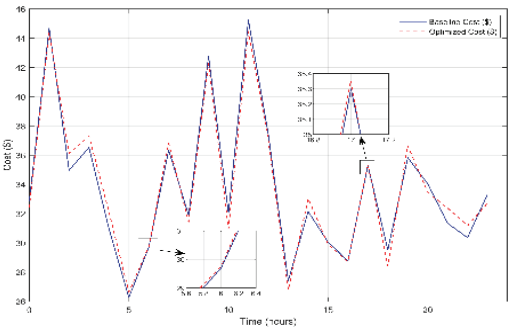


Fig. 5 Cost Minimization

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:14:29 UTC from IEEE Xplore. Restrictions apply.

TABLE II QUANTIFIED IMPROVEMENTS

Performance Metric	Baseline	AI-Optimized	Improvement
Absolute Forecasting Error	High	51.7% Lower	Significant Improvement
RMSE	100%	18-20% Reduction	Accurate Energy Prediction
Operational Cost Savings	Baseline Cost	3-35% Savings	Reduced Maintenance Costs
CAPEX for PV Systems	Baseline Expenditure	0-5% Lower	Improved Resource Planning

This study highlights the critical role of Artificial Intelligence in advancing solar energy management through optimization and forecasting. AI-enabled systems effectively predict energy output, dynamically adjust for deviations in real-time data, and optimize both generation and cost. The results demonstrate quantifiable improvements:

1. Performance Gains: Absolute error reduction by 51.7% and a notable RMSE reduction of 18%-20% in regional forecasts signify enhanced prediction accuracy.
2. Cost Minimization: Real-time optimization minimizes energy-related costs, with demonstrated savings through dynamic adjustments.
3. Dynamic Energy Adjustments: AI models ensure efficient management of deviations, balancing demand and supply seamlessly.

The simulated results revealed that AI-based energy adjustments significantly reduce baseline costs while maintaining reliable power output. Although minor deviations were observed due to simulated uncertainties, the overall performance validates the robustness of AI algorithms for energy optimization. Future research must address challenges like data standardization, ethical concerns, and scaling AI models for larger grids.

By implementing AI-driven strategies, solar plants and smart grids can efficiently manage distributed energy resources, minimize energy losses, and promote sustainable energy practices, thereby accelerating the global transition to clean energy solutions.

VI. VI. CHALLENGES

A. Key Challenges

1. Standardization of Protocols:
   - Lack of standardized communication frameworks and interoperability requirements.
   - Creates inefficiencies in integrating AI with existing energy systems.
2. Data Availability and Privacy Concerns:
   - AI systems require vast datasets for training and optimization.
   - Data acquisition is hindered due to privacy laws and fragmented ownership.
3. Bias in AI Algorithms:
   - AI decision-making processes may inherit biases embedded in training datasets.
   - This raises ethical and legal issues in energy management decisions.
4. Technical Complexity:
   - Existing frameworks for energy system simulations and demand forecasting lack advanced methodologies.
   - Integration challenges due to uncertain and evolving energy structures.
5. Interdisciplinary Research Needs:
   - Collaboration between energy experts, policymakers, and AI researchers is underdeveloped.
   - Limits progress in resilient and scalable energy system design.
6. Resilience and Adaptability:
   - Energy systems require adaptability to uncertain market conditions and external disruptions.
   - AI integration must ensure scalability and dynamic optimization.
7. Decentralized Energy Systems:
   - AI applications in distributed energy systems face incomplete data acquisition and complex operation environments.
   - Virtual technologies for optimizing energy grids are still evolving.
8. Skill and Workforce Development:
   - Energy professionals require training in AI, cloud computing, big data management, and cybersecurity.
   - Shortage of skilled personnel delays practical AI implementation.

TABLE 2 CHALLENGES AND SOLUTIONS

Challenges	Solutions
Standardization of protocols	Develop global frameworks for energy-AI interoperability.
Data availability and privacy concerns	Use federated learning, blockchain-based security, and privacy-preserving AI.
Bias in AI algorithms	Implement fairness-aware AI training methods and bias audits.
Technical complexity	Develop advanced AI simulation frameworks with predictive energy modeling.
Lack of interdisciplinary collaboration	Foster partnerships between AI experts, policymakers, and energy engineers.
Resilience and adaptability of energy systems	Build adaptive AI models for real-time energy and demand optimization.
Incomplete data for decentralized energy grids	Use synthetic datasets and real-time data from IoT sensors for training.
Skill and workforce shortage	Launch professional training programs on AI, cloud computing, and big data.

REFERENCES

[1] M. S. Hossain, A. Wadi Al-Fatlawi, L. Kumar, Y. R. Fang, and M. E. H. Assad, “Solar PV high-penetration scenario: an overview of the global PV power status and future growth,” Energy Syst., Sep. 2024, doi: 10.1007/s12667-024-00692-6.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:14:29 UTC from IEEE Xplore. Restrictions apply.

[2] A. E. Karaca and I. Dincer, A Wind Energy-Based Cogeneration System for Energy and Fresh Water Production, vol. 87. 2022.

[3] A. O. M. Maka and J. M. Alabid, "Solar energy technology and its roles in sustainable development," Clean Energy, vol. 6, no. 3, pp. 476–483, 2022, doi: 10.1093/ce/zkac023.

[4] T. Kamal, M. Karabacak, S. Z. Hassan, H. Li, and L. M. Fernandez-Ramirez, "A Robust Online Adaptive B-Spline MPPT Control of Three-Phase Grid-Coupled Photovoltaic Systems under Real Partial Shading Condition," IEEE Trans. Energy Convers., vol. 34, no. 1, pp. 202–210, 2019, doi: 10.1109/TEC.2018.2878358.

[5] M. Aghaei, F. Grimaccia, C. A. Gonano, and S. Leva, "Innovative Automated Control System for PV Fields Inspection and Remote Control," IEEE Trans. Ind. Electron., vol. 62, no. 11, pp. 7287–7296, 2015, doi: 10.1109/TIE.2015.2475235.

[6] P. Bouquet, I. Jackson, M. Nick, and A. Kaboli, "AI-based forecasting for optimised solar energy management and smart grid efficiency," Int. J. Prod. Res., vol. 62, no. 13, pp. 4623–4644, Jul. 2024, doi: 10.1080/00207543.2023.2269565.

[7] A. N. Abdalla et al., "Integration of energy storage system and renewable energy sources based on artificial intelligence: An overview," J. Energy Storage, vol. 40, no. June 2021, p. 102811, 2021, doi: 10.1016/j.est.2021.102811.

[8] E. Joseph, N. Chukwunweike, C. C. Ekwomadu, and S. A. Adeniyi, "Enhancing Green Energy Systems with Matlab Image Processing: Automatic Tracking of Sun Position for Optimized Solar Panel Efficiency," Int. J. Comput. Appl. Technol. Res., vol. 13, no. 08, pp. 62–72, 2024, doi: 10.7753/ijcatr1308.1007.

[9] S. Chen et al., "Cloud Trajectory Correction Prediction Method Combining Lightweight Tracking and Historical Information Feedback," IEEE Access, vol. 12, pp. 179664–179688, 2024, doi: 10.1109/ACCESS.2024.3506853.

[10] F. Yuan, Y. Lu, Z. Xie, and S. Dai, "Distributed Photovoltaic Distribution Voltage Prediction Based on eXtreme Gradient Boosting and Time Convolutional Networks," IEEE Access, vol. 12, pp. 177576–177588, 2024, doi: 10.1109/ACCESS.2024.3502759.

[11] Y. Zhou, L. Zhou, Z. Yi, D. Shi, and M. Guo, "Leveraging AI for Enhanced Power Systems Control: An Introductory Study of Model-Free DRL Approaches," IEEE Access, vol. 12, pp. 98189–98206, 2024, doi: 10.1109/ACCESS.2024.3422411.

[12] N. Roy et al., "Day-ahead Solar Power Generation Forecasting using LSTM and Random Forest Methods for North Eastern Region of India," in 2022 22nd National Power Systems Conference (NPSC), Dec. 2022, pp. 854–859, doi: 10.1109/NPSC57038.2022.10069833.

[13] D. Liu, Y. Xu, Q. Wei, and X. Liu, "Residential energy scheduling for variable weather solar energy based on adaptive dynamic programming," IEEE/CAA J. Autom. Sin., vol. 5, no. 1, pp. 36–46, Jan. 2018, doi: 10.1109/JAS.2017.7510739.

[14] Jianliang Meng; Yanyan Yang, "The application of improved decision tree algorithm in the electric power marketing," [Online]. Available: https://ieeexplore.ieee.org/document/6321634.

[15] S. M. Taghavinejad, M. Taghavinejad, L. Shahmiri, M. Zavvar, and M. H. Zavvar, "Intrusion Detection in IoT-Based Smart Grid Using Hybrid Decision Tree," in 2020 6th International Conference on Web Research (ICWR), Apr. 2020, pp. 152–156, doi: 10.1109/ICWR49608.2020.9122320.

[16] P. Chandel and L. Roy, "Solar Radiation Prediction Based on Hybrid Machine Learning Technique," in 2023 3rd International Conference on Technological Advancements in Computational Sciences (ICTACS), Nov. 2023, pp. 418–424, doi: 10.1109/ICTACS59847.2023.10390492.

[17] A. Nayak and L. Heistrene, "Hybrid Machine Learning Model for Forecasting Solar Power Generation," in 2020 International Conference on Smart Grids and Energy Systems (SGES), Nov. 2020, pp. 910–915, doi: 10.1109/SGES51519.2020.00167.

[18] E. Sawires, Z. Ismail, F. Amer, and S. Abdellatif, "Experimentally Validated Dynamic Equivalent Circuit Model of Perovskite Solar Cells: Utilizing Machine Learning Algorithms for Parameter Extraction Using I–V and C–V Characteristics," IEEE Trans. Instrum. Meas., vol. 73, pp. 1–11, 2024, doi: 10.1109/TIM.2024.3488138.

[19] B. Hashemi, A.-M. Cretu, and S. Taheri, "Snow Loss Prediction for Photovoltaic Farms Using Computational Intelligence Techniques," IEEE J. Photovoltaics, vol. 10, no. 4, pp. 1044–1052, Jul. 2020, doi: 10.1109/JPHOTOV.2020.2987158.

[20] U. K. Kalla, H. Kaushik, B. Singh, and S. Kumar, "Adaptive Control of Voltage Source Converter Based Scheme for Power Quality Improved Grid-Interactive Solar PV–Battery System," IEEE Trans. Ind. Appl., vol. 56, no. 1, pp. 787–799, Jan. 2020, doi: 10.1109/TIA.2019.2947397.

[21] A. H. Elmetwaly et al., "Improving Power Quality Problems of Isolated MG Based on ANN Under Different Operating Conditions Through PMS and ASSC Integration," IEEE Access, vol. 11, pp. 99822–99835, 2023, doi: 10.1109/ACCESS.2023.3311369.

[22] K. A. Mohammad and S. M. Musa, "Optimization of Solar Energy Using Recurrent Neural Network Controller," Proc. - 2022 14th IEEE Int. Conf. Comput. Intell. Commun. Networks, CICN 2022, vol. 2022-January, no. 02, pp. 681–686, 2022, doi: 10.1109/CICN56167.2022.10041248.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:14:29 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Aggregation of Distributed Energy Resources for Optimal Operation in Local Energy Communities With Grid Constraints

Shubham Gupta, Vinod Kumar Yadav, *Senior Member, IEEE*, and Madhusudan Singh, *Senior Member, IEEE*

***Abstract*—This article presents a framework to efficiently manage a sizable fleet of diverse distributed energy resources (DERs) operating within distribution systems to optimize the operations of local energy communities (LECs) and improve grid services. First, we develop aggregation modeling for DERs to evaluate cumulative flexibility that considers their preferences, spatial placement, and temporal behavior. Subsequently, we employ a hierarchical control framework (HCF) to put these aggregated flexibility of DERs for their effective dispatching. The HCF involves three key entities: an electric utility (EU) operator, community aggregators (CAs), and individual DERs. CAs harness the flexibility obtained from the aggregated DERs within their respective LECs to minimize operational costs while also considering the distribution network constraints. On the other hand, the EU operator coordinates dispatch setpoints received from CAs along with the disaggregated DERs to regulate distribution system node voltages and reduce power losses to enhance grid services. Numerical simulations conducted on a modified IEEE-123 bus radial distribution network demonstrate the efficacy of our approach in effectively managing DERs for cost-efficient operations and improving grid services.**

***Index Terms*—Aggregation model, electric vehicles (EVs), hierarchical control, load modeling, local energy communities (LECs).**

NOMENCLATURE

Variables

$P^{\text{flow}}_{ij,t}, Q^{\text{flow}}_{ij,t}$	Active, reactive branch power flow between buses i and j at time t.
$l^{\text{flow}}_{ij,t}$	Current flow (squared) between buses i and j at time t.
$\upsilon^{\text{node}}_{i,t}$	Node voltage (squared) at bus i at time t.
$E^{\text{CS}}_{i,t}$	Residual energy of CS at bus i at time t.
$P^{\text{ch,CS}}_{i,t}$	Charging power of CS at bus i at time t.
$P^{\text{dch,CS}}_{i,t}$	Discharging power of CS at bus i at time t.
$\zeta^{\text{DL}}_{i,j,t}$	State of DL j at bus i at time t.
$\zeta^{\text{DL}}_{i,j,t_0}$	State of DL j at bus i in current time slot t_0.
$P^{\text{DL}}_{i,j,t}$	Power consumed by DL j at bus i at time t.
$\zeta^{\text{task}}_{i,j,k,t}$	State of task k for DL j at bus i at time t.
$P^{\text{RE}}_{i,g,t}, Q^{\text{RE}}_{i,g,t}$	Active, reactive power supplied by RESs g at bus i at time t.

Parameters

$p^{\text{FL}}_{i,t}, q^{\text{FL}}_{i,t}$	Inflexible (fixed) active, reactive load demand at bus i at time t.
r_{ij}, x_{ij}	Resistance, reactance of a line connecting buses i and j.
$N^{a}_{i,t}, N^{d}_{i,t}$	No. of EVs arrived, departed at bus i at time t.
$E^{a,\text{EV}}_{i,n}$	Residual energy of arrived EV n at bus i.
$E^{d,\text{EV}}_{i,n}$	Desired energy at departure of EV n at bus i.
$\eta^{\text{CS}}_{\text{ch},i}, \eta^{\text{CS}}_{\text{dch},i}$	Efficiencies of the charging, discharging power of CS at bus i.
$\overline{E}^{\text{CS}}_{i,t}, \underline{E}^{\text{CS}}_{i,t}$	Upper, lower bound of the residual energy of CS at bus i at time t.
$\overline{P}^{\text{ch,CS}}_{i,t}$	Maximum charging power of CS at bus i at time t.
$\overline{P}^{\text{dch,CS}}_{i,t}$	Maximum discharging power of CS at bus i at time t.
$\overline{P}^{\text{ch,EV}}_{i,n}$	Maximum charging power of EV n at bus i.
$\overline{P}^{\text{dch,EV}}_{i,n}$	Maximum discharging power of EV n at bus i.
$\overline{E}^{\text{EV}}_{i,n,t}, \underline{E}^{\text{EV}}_{i,n,t}$	Upper, lower bound of the residual energy of EV n at bus i at time t.
$\eta^{\text{dch,EV}}_{i,n}$	Discharging power efficiency of EV n at bus i.
$\eta^{\text{ch,EV}}_{i,n}$	Charging power efficiency of EV n at bus i.
$t^{\text{EV}}_{a,n}, t^{\text{EV}}_{d,n}$	Arrival, departure time of EV n.
$\underline{E}^{\text{EV}}_{i,n}$	Minimum of residual energy of EV n at bus i.
$\overline{E}^{\text{EV}}_{i,n}$	Maximum of residual energy of EV n at bus i.
$\alpha_{i,j}, \beta_{i,j}$	Starting, ending time (operational) of DL j at bus i.
$\hat{\zeta}^{\text{DL}}_{i,j,t}$	State of DL j at bus i in previous time slot t.
$\text{Z}^{\text{DL}}_{i,j}$	No. of time slots (operational) required by DL j at bus i.
$E^{\text{DL}}_{i,j}$	Energy required by DL j at bus i.
$P^{\text{task}}_{i,j,k}$	Rated power of task k for DL j at bus i.
$\text{Z}^{\text{task}}_{i,j,k}$	No. of time slots (operational) required by DL j at bus i to finish task k.
$S^{\text{RE}}_{i,g,t}$	Apparent power capacity of RES g at bus i at time t.

Received 11 May 2024; revised 12 February 2025; accepted 28 March 2025. Paper no. TII-24-2271. *(Corresponding author: Shubham Gupta.)*

The authors are with the Department of Electrical Engineering, Delhi Technological University, New Delhi 110042, India (e-mail: shubham.gupta6691@gmail.com; vinodkumar@dtu.ac.in; madusudan@dce.ac.in).

Digital Object Identifier 10.1109/TII.2025.3558313

1941-0050 © 2025 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:36:46 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

$\phi_{i,g}$ Angle defining power factor of RES g at bus i.
$P^{\text{RE,fore}}_{i,g,t}$ Active power forecast for RES g at bus i at time t.

I. Introduction

A. Background

THE paradigm shift toward clean energy fundamentally relies upon electrification, leading to an escalation in power consumption and a concurrent surge in the quantity and diversity of electrical devices capable of adjusting their demand. For instance, heat pumps are gradually replacing traditional gas boilers. In parallel, the grid has witnessed a notable upswing in the adoption of electric vehicles (EVs) and renewable energy sources (RESs), prominently photovoltaic (PV) panels. Distributed energy resources (DERs) equipped with diverse technologies, such as deferrable loads (DLs), RESs, and EVs, are dynamically transforming the energy framework, steering it toward a more de-carbonized and secure configuration. Thus, recognizing and harnessing the full potential of DERs becomes indisputably paramount. However, managing a substantial array of DERs to facilitate seamless system-wide functioning and control requires a permissive regulatory framework, adaptable infrastructure, and suitable software platforms, which costs a fortune to electric utilities (EUs) [1]. As a result, many EUs are still grappling with sorting out DERs to blend into the grid, seeking to enhance both reliability and energy cost savings.

Recently, there has been a notable increase in the contribution of weather-dependent and variable RESs to the energy generation mix. This trend highlights the growing necessity for enhanced flexibility to ensure a consistent power supply that meets evolving demand patterns. Installing energy storage systems (ESs) alongside RESs has been proposed as a viable and effective strategy for enhancing the adept localization of RESs within distribution networks. However, the practical feasibility and associated costs of implementing such integrated RESs with ESs present challenges that impede their widespread adoption [2]. Consequently, EVs, which share analogous electrical characteristics with ESs, have emerged as a promising resource to facilitate the greater penetration of RESs. Moreover, the strategic temporal scheduling of spatially dispersed DLs can enhance RESs spread through significant demand flexibility. However, a collective surge in the charging of EVs and activation of DLs could lead to sudden spikes in power demand, potentially overwhelming the grid's capacity, unless measures are employed to smooth and distribute peak demand across off-peak periods [3].

As the challenges in the energy landscape are constantly evolving, it is of the utmost importance for EU operators to take a proactive approach to managing distribution networks and DERs. However, DERs can often operate invisibly and lie beyond EU operator control, which presents a challenge due to their intricate behavior. Integrating them into the broader grid operations requires digital technologies that can actively engage individuals as participants. Further, the transition to cleaner energy must be not only to gain political support for clean energy projects but also to foster local acceptance of new developments and infrastructure. Therefore, community-based initiatives are gaining traction as effective ways to create DER-enabled energy systems that are more people-centric, equitable, and resilient [4].

B. Related Literature Review and Research Gap

The growing integration of DERs undeniably enhances the flexibility of the distribution system. However, due to the relatively low individual power capacity and sheer quantities of DERs, a challenge known as the "curse of dimensionality" emerges. Thus, effectively managing each unit of DER becomes impractical within the operational framework of the distribution system [5]. As a solution, existing research suggests employing aggregated operational models at the distribution level for efficiently managing multiple small-capacity DERs [6]. Essentially, these approaches utilize the feasible limits of crucial parameters of DER to exploit flexibility. In [7], the aggregation model has been developed by an arithmetic sum of the individual feasible limits of DERs to determine cumulative flexibility, providing an outer boundary of the feasible region. However, the method has had issues producing viable solutions during the disaggregation process as it enlarged the feasible operation regions [8]. Thus, it further requires an additional problem formulation to ensure a workable solution, which results in higher information communication exchange [9]. Furthermore, methods have aimed to produce feasible disaggregation solutions, albeit increasing the computing burden [10], [11].

Moreover, [12] considers the inner-box approximation to define the feasible aggregate upper and lower power trajectories. In [13], a convex inner approximation is introduced to quantify a feeder's aggregate DER hosting capacity. However, inner aggregation methods substantially compromise the dispatching potential of DERs. As a result, it has become pivotal for future works to aggregate the viable parameter boundaries to enhance precision and practicality while still unleashing the untapped potential of DERs [14].

Recent research has explored integrating aggregators with distribution network operators to harness the collective potential of EVs and delayable loads. In [15], a methodology is outlined for efficient and secure dispatching of DERs. However, it assumes that the DERs can consume power from the grid, neglecting the possibility of bidirectional power flow. In [16], a hierarchical control framework (HCF) is introduced to optimize power flexibility within smart homes equipped with DLs, but its centralized approach raises privacy and data vulnerability concerns. In [17] and [18], an analytical expression is derived to define the dispatchable region for a large fleet of EVs. However, while these methods expand the feasible dispatchable region, they rely on relaxed complementarity constraints for EV battery modeling, which may not hold under practical conditions [19]. Moreover, the methods to assess the exact power flexibility of EVs and delayable loads are put forth in [11] and[20]. However, these models fail to account for the diverse parameters inherent to DERs. In this article, the constraints for aggregating EVs and DLs impose stringent conditions to improve accuracy.

Moreover, to fully exploit the benefits of flexibility derived from DERs, it is necessary to have a suitable control scheme in place. In the literature, various control methods [21], [22], [23], [24], [25], [26], including centralized, decentralized, hybrid, and hierarchical controls, are often applied for optimizing problems, such as voltage regulation, transmission congestion, demand response, uncertainty modeling, and reactive power dispatch, by utilizing the flexibility offered by DERs. Centralized control is frequently employed to coordinate multiple DERs effectively, owing to the increased visibility of the network [21], [22]. On the other hand, decentralized control enables for more quick

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:36:46 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

response times [23], [24]. Consequently, a combination of centralized and decentralized control methods is often adopted to leverage the benefits of both approaches [25]. However, the existing strategies for regulating flexibility overlook crucial aspects, such as the physical distribution of flexibility within the network, the operational limits of individual flexible resources, and grid constraints, which significantly impact the performance of the distribution system [27].

C. Contribution and Article Layout

This article puts forth the existing gap in the literature related to the effective aggregation and control of a diverse array of DERs dispersed throughout a distribution system into three key aspects: first, a lack of EVs and DLs aggregation modeling with stringent constraints to accurately assess cumulative and practical feasible flexibility. Second, the failure to consider distribution network constraints in EVs and DLs flexibility evaluation. Finally, a control scheme is absent that not only effectively coordinates among stakeholders but also ensures the preservation of their data privacy concerns and respects their preferences for self-optimization. While comparing with the existing literature, the main contributions can be summarized as follows:

1) Proposed an aggregation model aimed at assessing the collective feasible operational range for a substantial EV fleet at a charging station (CS). In contrast to [11], [17], [18], and [20], the presented approach notably enhances precision and computational efficiency in effectively managing extensive EVs with considerations for spatial placement and temporal behavior.
2) Presented an approach for DLs modeling represented by unique operational tasks, with the primary objective of facilitating time of use (ToU) price-based demand response. Unlike the existing literature [7], [11], [15], [16], and [20], the proposed model assesses the influence of a large population of DLs on the distribution system, considering their integration into the network.
3) In contrast with prior literature [8] and [20], this article proposed an HCF to facilitate the coordination of the EU operator, community aggregators (CAs), and individual DERs, thereby optimizing the deployment of DERs for cost-optimal operations within local energy communities (LECs) and improved grid services, while considering the distribution network constraints.

The rest of this article is organized as follows. Section II briefs the proposed aggregation and control framework and the branch power flow model for distribution network modeling. The model for EVs aggregation at CS and the modeling of DLs for ToU price-based demand response mechanism are discussed in Section III. The HCF optimal operation strategy for LECs and the EU operator is formulated in Section IV. Numerical results are presented in Section V. Finally, Section VI concludes this article.

II. Proposed Framework Overview and Network Modeling

This article considers the multiple LECs in the distribution system that consist of fixed (nondeferrable) loads along with

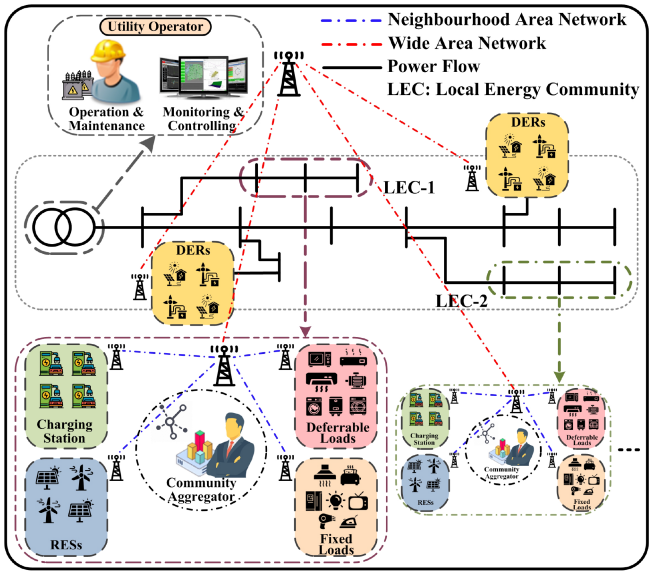


Fig. 1. Schematic of the proposed aggregation and control framework.

different DERs, such as CSs, DLs, and RESs, in the form of PV plants and wind turbine (WT) farms, as shown in Fig. 1. The proposed work assumed that the CAs manage DERs dispersed within their communities. In addition, an EU operator oversees LECs and disaggregated DERs, such as ESs integrated with PV plants and WT farms.

For simplicity, a radial distribution network (RDN) with a single feeder connected to EU modeled by a tree graph $\mathcal{G} := (\mathcal{N}, \mathcal{E})$, where $i \in \mathcal{N} := 1, 2, \ldots, n$ represents a node and $(i, j) \in \mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ represents a distribution line connected ith to jth nodes. The relaxed branch flow model presented in [28] is employed for power flows for this network, as expressed in (1a)–(1d); where, for each node $i \in \mathcal{N}$ at tth time interval, denote $\upsilon_{i,t}^{\text{node}} = |V_{i,t}^{\text{node}}|^2$ refers to magnitude squared of node voltage. Let $s_{i,t}^{\text{FL}} = p_{i,t}^{\text{FL}} + jq_{i,t}^{\text{FL}}$ be node i net complex fixed load demand. Similarly, for each line $(i, j) \in \mathcal{E}$ at tth time interval, denote $l_{ij,t}^{\text{flow}} := |I_{ij,t}^{\text{flow}}|^2$ refers to magnitude squared of branch current flows from i node to j node. Let $S_{ij,t}^{\text{flow}} = P_{ij,t}^{\text{flow}} + jQ_{ij,t}^{\text{flow}}$ be the *sending-end* branch flow from i node to j node. A solution of (1a)–(1d) to a given s^{FL} is denoted by $x(s^{\text{FL}}) := (S^{\text{flow}}, I^{\text{flow}}, V^{\text{node}}, s_0^{\text{FL}})$, is unique for practical radial network where $|V_0| \simeq 1$ and r_{ij}, x_{ij} are small p.u. values

$$p_{j,t}^{\text{FL}} = \sum_{i:i\to j} \left(P_{ij,t}^{\text{flow}} - r_{ij} l_{ij,t}^{\text{flow}}\right) - \sum_{k:j\to k} P_{jk,t}^{\text{flow}} \tag{1a}$$

$$q_{j,t}^{\text{FL}} = \sum_{i:i\to j} \left(Q_{ij,t}^{\text{flow}} - x_{ij} l_{ij,t}^{\text{flow}}\right) - \sum_{k:j\to k} Q_{jk,t}^{\text{flow}} \tag{1b}$$

$$\upsilon_{j,t}^{\text{node}} = \upsilon_{i,t}^{\text{node}} - 2\left(r_{ij} P_{ij,t}^{\text{flow}} + x_{ij} Q_{ij,t}^{\text{flow}}\right) + \left(r_{ij}^2 + x_{ij}^2\right) l_{ij,t}^{\text{flow}} \tag{1c}$$

$$\upsilon_{i,t}^{\text{node}} l_{ij,t}^{\text{flow}} = \left(P_{ij,t}^{\text{flow}}\right)^2 + \left(Q_{ij,t}^{\text{flow}}\right)^2. \tag{1d}$$

Furthermore, the proposed control framework considers the LEC a *subnetwork* comprising DERs linked to the distribution feeder line via a single point of common coupling. Initially, the CA gathers energy demands and behavioral data from DERs

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:36:46 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

dispersed within its operational region. It then evaluates the feasible energy flexibility to facilitate effective dispatching, aiming for the optimal utilization of DERs to achieve cost-optimal operation within the LEC while considering distribution network constraints. As seen in Fig. 1, the information flow is depicted by dotted lines. Subsequently, the CA communicates these dispatch setpoints to the EU operator, who is responsible for coordination among disaggregated DERs and dispatch setpoints received from CAs, aiming to improve grid services while accounting for distribution network constraints.

III. DERs Aggregation Model

To effectively incorporate a substantial number of DERs into the overall system operation, it is imperative to establish an aggregate model that can encompass and account for their inherent flexibility. This article focuses primarily on two crucial categories of DERs, namely, EVs and DLs. This section elucidates the modeling approach employed for aggregating EV fleets at CS and for aggregating DLs to facilitate ToU price-based demand response.

A. Model for EVs Aggregation at CS

In practice, the CS can gather critical data regarding EVs upon connection, including arrival and departure time, the initial and desired energy, the maximum charging rate, and the user-defined charging duration [17]. The ebb and flow of EV arrivals and departures contribute to a dynamic fluctuation in the number of EVs present at the CS over time. Consequently, the EV fleets at the CS are aggregated into a unified entity with time-dependent power and energy constraints, effectively characterizing the operational parameters of the CS.

1) The power capacity constraints of EV aggregation

$$P_{i,t}^{\text{dch,CS}} \le \overline{P}_{i,t}^{\text{dch,CS}} - \left(\overline{P}_{i,t}^{\text{dch,CS}} / \overline{P}_{i,t}^{\text{ch,CS}}\right) P_{i,t}^{\text{ch,CS}} \tag{2a}$$

$$0 \le P_{i,t}^{\text{ch,CS}} \le \overline{P}_{i,t}^{\text{ch,CS}} \tag{2b}$$

$$0 \le P_{i,t}^{\text{dch,CS}} \le \overline{P}_{i,t}^{\text{dch,CS}} \tag{2c}$$

$$\overline{P}_{i,t}^{\text{ch,CS}} = \overline{P}_{i,t}^{\text{ch,CS}} + \overline{P}_{i,n\in N}^{\text{ch,EV}} \left(N_{i,t}^{a} - N_{i,t}^{d}\right) \tag{2d}$$

$$\overline{P}_{i,t}^{\text{dch,CS}} = \overline{P}_{i,t}^{\text{dch,CS}} + \overline{P}_{i,n\in N}^{\text{dch,EV}} \left(N_{i,t}^{a} - N_{i,t}^{d}\right) \tag{2e}$$

where (2a) is from the **Extn-LP** formulation for the battery modeling proposed in [29] and utilized for the reduction in the feasible search space. Constraints in (2b) and (2c) delineate the upper and lower bounds on the charging and discharging powers at CS at bus i at time slot t, respectively. Further, (2d) and (2e) quantify the upper boundaries of charging and discharging power at CS at bus i at time slot t, respectively.

2) The energy capacity constraints of EV aggregation

$$E_{i,t}^{\text{EV}(a,d)} = \sum_{n=1}^{N_{i,t}^{a}} E_{i,n}^{a,\text{EV}} - \sum_{n=1}^{N_{i,t}^{d}} E_{i,n}^{d,\text{EV}} \tag{3a}$$

$$E_{i,t}^{\text{CS}} = E_{i,t-1}^{\text{CS}} + \left[\eta_{\text{ch},i}^{\text{CS}} P_{i,t}^{\text{ch,CS}} - \frac{P_{i,t}^{\text{dch,CS}}}{\eta_{\text{dch},i}^{\text{CS}}}\right] \times \Delta t + E_{i,t}^{\text{EV}(a,d)} \tag{3b}$$

$$\left(\eta_{\text{ch},i}^{\text{CS}} P_{i,t}^{\text{ch,CS}}\right) \Delta t \le \left(\overline{E}_{i,t}^{\text{CS}} - E_{i,t-1}^{\text{CS}} - E_{i,t}^{\text{EV}(a,d)}\right) \tag{3c}$$

$$\left(P_{i,t}^{\text{dch,CS}}\right) \Delta t \le \left(E_{i,t-1}^{\text{CS}} + E_{i,t}^{\text{EV}(a,d)} - \underline{E}_{i,t}^{\text{CS}}\right) \eta_{\text{dch},i}^{\text{CS}} \tag{3d}$$

$$\underline{E}_{i,t}^{\text{CS}} \le E_{i,t}^{\text{CS}} \le \overline{E}_{i,t}^{\text{CS}} \tag{3e}$$

$$\underline{E}_{i,t}^{\text{CS}} = \underline{E}_{i,t-1}^{\text{CS}} + \sum_{n=1}^{N_{i,t}^{a}} \underline{E}_{i,n,t}^{\text{EV}} - \sum_{n=1}^{N_{i,t}^{d}} E_{i,n}^{d,\text{EV}} \tag{3f}$$

$$\overline{E}_{i,t}^{\text{CS}} = \overline{E}_{i,t-1}^{\text{CS}} + \sum_{n=1}^{N_{i,t}^{a}} \overline{E}_{i,n,t}^{\text{EV}} - \sum_{n=1}^{N_{i,t}^{d}} E_{i,n}^{d,\text{EV}} \tag{3g}$$

In (3a), $E_{i,t}^{\text{EV}(a,d)}$ denotes the amount of energy received from arriving EVs and subtracting the desired energy of departing EVs at bus i at time slot t. (3b) outlines the energy update rule for CS at bus i at time slot t. Notably, (3c) and (2b) results in a more stringent condition: $P_{i,t}^{\text{ch,CS}} \le \min\{\overline{P}_{i,t}^{\text{ch,CS}}, (\overline{E}_{i,t}^{\text{CS}} - E_{i,t-1}^{\text{CS}} - E_{i,t}^{\text{EV}(a,d)})\}$, that enforces a tighter upper bound on the charging power capacity, while (3d) and (2c) yields a more rigorous constraint: $P_{i,t}^{\text{dch,CS}} \le \min\{\overline{P}_{i,t}^{\text{dch,CS}}, (E_{i,t-1}^{\text{CS}} + E_{i,t}^{\text{EV}(a,d)} - \underline{E}_{i,t}^{\text{CS}})\}$, that establishes a tighter upper bound on the discharging power capacity, as in [29]. Furthermore, (3e) imposes restrictions on the energy at CS at bus i at time slot t. Constraints (3f) and (3g) calculate the lower and upper boundaries, respectively, for the net energy at CS at bus i at time slot t. Finally, constraints (3h) and (3i) shown at the bottom of this page, determine the permissible operational range for an EV n at CS at bus i at time slot t.

1) Illustrative Example: To better illustrate the proposed model and conduct a comparative analysis with previously reported aggregation models in [17] and [18], we have employed a case study involving two specific EVs. The parameters and underlying assumptions are drawn from [17] and presented in Table I. The notations $t_{a,n}^{\text{EV}}$ and $t_{d,n}^{\text{EV}}$ denote the arrival and

$$\underline{E}_{i,n,t}^{\text{EV}} = \begin{cases} 0, & t \in T \setminus \left[t_{a,n}^{\text{EV}}, t_{d,n}^{\text{EV}}\right] \\ \max\left\{\underline{E}_{i,n}^{\text{EV}}, E_{i,n}^{a,\text{EV}} - \frac{\overline{P}_{i,n,t}^{\text{dch,EV}}}{\eta_{i,n}^{\text{dch,EV}}}\left(t - t_{a,n}^{\text{EV}}\right), E_{i,n}^{d,\text{EV}} - \eta_{i,n}^{\text{ch,EV}} \overline{P}_{i,n,t}^{\text{ch,EV}}\left(t_{d,n}^{\text{EV}} - t\right)\right\}, & t_{a,n}^{\text{EV}} \le t \le t_{d,n}^{\text{EV}} \end{cases} \tag{3h}$$

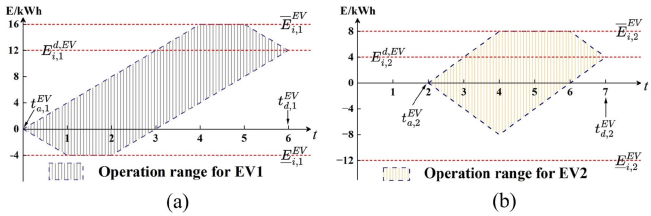
$$\overline{E}_{i,n,t}^{\text{EV}} = \begin{cases} 0, & t \in T \setminus \left[t_{a,n}^{\text{EV}}, t_{d,n}^{\text{EV}}\right] \\ \min\left\{\overline{E}_{i,n}^{\text{EV}}, E_{i,n}^{a,\text{EV}} + \eta_{i,n}^{\text{ch,EV}} \overline{P}_{i,n,t}^{\text{ch,EV}}\left(t - t_{a,n}^{\text{EV}}\right), E_{i,n}^{d,\text{EV}} + \frac{\overline{P}_{i,n,t}^{\text{dch,EV}}}{\eta_{i,n}^{\text{dch,EV}}}\left(t_{d,n}^{\text{EV}} - t\right)\right\}, & t_{a,n}^{\text{EV}} \le t \le t_{d,n}^{\text{EV}}. \end{cases} \tag{3i}$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:36:46 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE I
PARAMETERS OF EVS [17]

EV	$t^{\mathrm{EV}}_{a,n}$	$t^{\mathrm{EV}}_{d,n}$	$E^{a,\mathrm{EV}}_{i,n}$ (kWh)	$\overline{P}^{\mathrm{ch,CS}}_{i,t}$ / $\overline{P}^{\mathrm{dch,CS}}_{i,t}$ (kW)	$B^{\mathrm{EV}}_{i,n}$ (kWh)
1	0	6	4	4	20
2	2	7	12	4	20



(a) (b)

Fig. 2. Operation ranges of the individual EV1 and EV2.

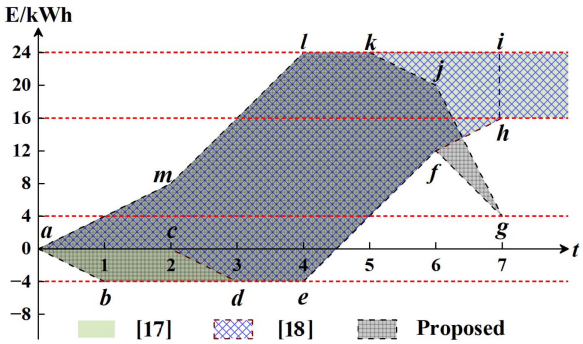


Fig. 3. Operation range for EVs aggregation (EV1 and EV2) models.

departure times of the nth EV respectively, while $E^{a,\mathrm{EV}}_{i,n}$ indicates the residual energy of EV n at bus i upon arrival, exemplified by 4 kWh for EV1 and 12 kWh for EV2. Moreover, $\overline{P}^{\mathrm{ch,CS}}_{i,t}/\overline{P}^{\mathrm{dch,CS}}_{i,t}$ signifies the maximum charging or discharging powers of the CS i at time slot t, and $B^{\mathrm{EV}}_{i,n}$ represents the maximum energy storage capacity of the EV n at bus i . Furthermore, it is assumed that the desired energy level for both EVs upon departure is set at 16 kWh.

The operational regions for EV1 and EV2 are computed utilizing (3h) and (3i), and the results are depicted in Fig. 2(a) and (b), respectively. Further, employing (3f) and (3g) from the proposed model, the energy aggregation of EV1 and EV2 at CS and a comparative evaluation against the models in [17] and [18] is shown in Fig. 3. Corresponding to the 4th interval, EV2 can discharge up to 8 kWh from its initial energy state, as depicted in Fig. 2(b). On the other hand, for the same interval, EV1 requires charging up to 4 kWh from its initial energy state, as shown in Fig. 2(a). Hence, Fig. 3, which represents the aggregation of Fig. 2(a) and (b) gives the net discharging of the CS up to 4 kWh. Furthermore, the upper aggregate energy limit of the CS starts to decline in the 5th time slot to obligate the energy desired by EV1 before it departs from CS. The reduction is further underlined in the next time slots, coinciding with the departures of EV1 from the CS in the sixth and EV2 in the seventh, consequently rendering their energy unavailable for dispatch. In contrast, prior methods exemplified in [17] and [18] exhibit different behavior, where the energy limits of EVs accumulate in subsequent time slots even after they depart from the CS.

In addition, the earlier studies, as in [17] and [18], assume that EVs may charge up to or above the stipulated desired energy, as shown in Fig. 3. This study delves into a critical inquiry: why would an EV be willing to pay for charging over their desired energy? Therefore, the proposed method restrains the upper bound to the desired energy level at the departure time of an EV, refer to Fig. 2.

B. DLs Aggregation for ToU Price-Based Demand Response

Effectively managing a range of DLs is essential in proactively mitigating the risk of sudden surges in power demand and reducing overall electricity payments in ToU pricing environments, thereby optimizing energy efficiency and economic outcomes within an LEC. To achieve this, CAs are responsible for gathering information from consumers and the electricity market, including ToU electricity prices, spatial and temporal operational characteristics, and the energy demands linked to individual household loads.

Mathematically, the modeling for DL aggregation to provide demand response is as follows:

$$\zeta^{\mathrm{DL}}_{i,j,t} = 0, \quad t \in T \setminus [\alpha_{i,j}, \beta_{i,j}] \tag{4a}$$

$$\sum_{t=\alpha_{i,j}}^{t_0-1} \hat{\zeta}^{\mathrm{DL}}_{i,j,t} + \zeta^{\mathrm{DL}}_{i,j,t_0} + \sum_{t=t_0+1}^{\beta_{i,j}} \zeta^{\mathrm{DL}}_{i,j,t} = \mathrm{Z}^{\mathrm{DL}}_{i,j} \tag{4b}$$

$$\sum_{t=\alpha_{i,j}}^{t_0-1} \hat{P}^{\mathrm{DL}}_{i,j,t} + P^{\mathrm{DL}}_{i,j,t_0} + \sum_{t=t_0+1}^{\beta_{i,j}} P^{\mathrm{DL}}_{i,j,t} = E^{\mathrm{DL}}_{i,j} \tag{4c}$$

where (4a) makes sure that an DL j at bus i must function within the predetermined operational time-frame $[\alpha_{i,j}, \beta_{i,j}]$ as specified by the consumer. Constraint (4b) describes that an DL j at bus i requires total $\mathrm{Z}^{\mathrm{DL}}_{i,j}$ time slots to complete all tasks, consequently, DL j needs to start no later than $\beta_{i,j} - \mathrm{Z}^{\mathrm{DL}}_{i,j} + 1$ and be finished no earlier than $\alpha_{i,j} + \mathrm{Z}^{\mathrm{DL}}_{i,j}$. In addition, a DL j at bus i has prespecified energy requirement $E^{\mathrm{DL}}_{i,j}$ to complete all tasks, as shown in (4c).

For a DL j at bus i that has multiple operation tasks $k = 1, 2, \ldots, N_{j,k}$ with distinct power levels and are mutually exclusive to each other are described in (5a)–(5b). However, these tasks must operate in a sequence, where the first task must finished before the second task starts, which is ensured in constraint (5c)

$$P^{\mathrm{DL}}_{i,j,t} = \sum_{k=1}^{N_{j,k}} P^{\mathrm{task}}_{i,j,k} \xi^{\mathrm{task}}_{i,j,k,t} \tag{5a}$$

$$\zeta^{\mathrm{DL}}_{i,j,t} = \sum_{k=1}^{N_{j,k}} \xi^{\mathrm{task}}_{i,j,k,t} \tag{5b}$$

$$\sum_{x=1}^{k-1} \sum_{t=\alpha_{i,j,k}}^{t_0-1} \xi^{\mathrm{task}}_{i,j,x,t} \geq \left(\sum_{x=1}^{k-1} \mathrm{Z}^{\mathrm{task}}_{i,j,k} \right) \xi^{\mathrm{task}}_{,i,j,x,t_0}. \tag{5c}$$

This article categorizes the tasks of a DL based on their operational characteristics, notably distinguishing between interruptible and deferrable tasks and noninterruptible and deferrable tasks. A graphical example in Fig. 4 demonstrates the distinction between task types, where the overall time frame for a sequential

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:36:46 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

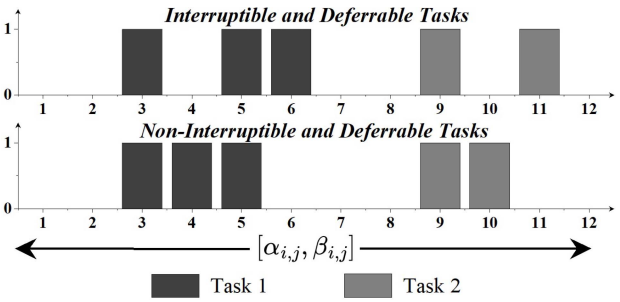


Fig. 4. Graphical example of residential DL with distinct types of tasks.

task is between 1 and 12; the total time slots required to complete task 1 is three, and task 2 is two. Moreover, it is shown that task 1 finishes before the start of task 2.

1) Interruptible and Deferrable Task: For an interruptible and deferrable task, the starting time may vary within the designated operational time frame, denoted as $[\alpha_{i,j,k}, \beta_{i,j,k}]$. Equation (6) assures that the operation of a DL can defer, with the constraint that the delay must not exceed a predetermined time threshold $[\alpha_{i,j,k} + \varphi_{i,j,k}]$, where $\varphi_{i,j,k}$ represents the acceptable duration of delay as perceived by consumers

$$\sum_{t=\alpha_{i,j,k}}^{t_0-1} \hat{\xi}^{\text{task}}_{i,j,k,t} + \xi^{\text{task}}_{i,j,k,t_0} + \sum_{t=t_0+1}^{\alpha_{i,j,k}+\varphi_{i,j,k}} \xi^{\text{task}}_{,i,j,k,t} \geq 1. \tag{6}$$

2) Noninterruptible and Deferrable Task: Deferrable characterizes the rescheduling of a task during the course of the day. Conversely, once it starts, to mitigate any potential reductions in thermal efficiency, it is imperative to complete it without interruption. Constraints (6) and (7) are employed to delineate the attributes of noninterruptible and deferrable properties

$$\sum_{t=t_0+1}^{t_0+\text{Z}^{\text{task}}_{i,j,k}} \xi^{\text{task}}_{i,j,k,t} \geq \text{Z}^{\text{task}}_{i,j,k} \left(\xi^{\text{task}}_{i,j,k,t_0+1} - \xi^{\text{task}}_{i,j,k,t_0}\right). \tag{7}$$

IV. Optimal Operation Strategy of LECs and EU Operator

This study incorporates an HCF optimal operation strategy for executing controls of DERs flexibility at both the LEC and EU levels. The primary objective of the CA is to enable the activation of DERs within an LEC based on signals from the EU operator and grid constraint considerations for optimal cost operations. In turn, the EU operator assumes to coordinate all CAs and provide dispatch setpoints to disaggregated DERs on the feeder while considering network constraints to improve the performance of the electrical grid.

A. Cost Optimal Dispatch Problem Formulation

CA solves the following problem formulation to optimize its operation cost and determine the dispatch setpoints for the available DERs within the LEC

$$\textbf{min} \quad \sum_{t\in T}\sum_{i\in\mathcal{N}} \left[\lambda^{\text{ToU}}_{\text{RL},t}\left(p^{\text{FL}}_{i,t} + \sum_{j\in\mathcal{N}^{\text{SL}}_i} P^{\text{SL}}_{i,j,t}\right)\right.$$
$$\left. + \lambda_{\text{RE},t}\left(\sum_{g\in\mathcal{G}_i} P^{\text{RE}}_{i,g,t}\right) + \lambda^{\text{ToU}}_{\text{CS},t}\left(P^{\text{ch,CS}}_{i,t} - P^{\text{dch,CS}}_{i,t}\right)\right] \tag{8a}$$

$$\text{s.t. } (1)-(7), \upsilon^{\text{node}}_i \in \left\{\underline{V}_i, \overline{V}_i\right\} \tag{8b}$$

$$\left(P^{\text{RE}}_{i,g,t}\right)^2 + \left(Q^{\text{RE}}_{i,g,t}\right)^2 \leq \left(S^{\text{RE}}_{i,g,t}\right)^2 \tag{8c}$$

$$-\phi_{i,g} P^{\text{RE}}_{i,g,t} \leq Q^{\text{RE}}_{i,g,t} \leq \phi_{i,g} P^{\text{RE}}_{i,g,t} \tag{8d}$$

$$0 \leq P^{\text{RE}}_{i,g,t} \leq P^{\text{RE,fore}}_{i,g,t}. \tag{8e}$$

The objective (8a) aims to minimize the overall operation cost, where the first term denotes the electricity payments from CA to the EU operator to supply the power required for fixed $(p^{\text{FL}}_{i,t})$ and DL $(P^{\text{SL}}_{i,j,t})$ demands. The $\lambda^{\text{ToU}}_{\text{RL},t}$ represents the standard ToU electricity rate that applies to power purchased from the EU. The second term in (8a) denotes the payments from CA to RESs for procuring power at a fixed electricity rate, $\lambda_{\text{RE},t}$. The third term in (8a), denoted as $\lambda^{\text{ToU}}_{\text{CS},t}(P^{\text{ch,CS}}_{i,t} - P^{\text{dch,CS}}_{i,t})$, refers to the cost of electricity for the net charging power demand of a CS. The constraints in (8b)–(8e) ensure the feasibility of power flow in the network and DERs power bounds. The constraint (8b) bounds voltages magnitude within specified limits, which implicitly characterize the use of voltage regulators/shunt capacitor banks. Constraints (8c) and (8d) ensure balance in active and reactive power sourced from RESs integrated inside an LEC and keep the power factor constant.

B. Problem Formulation for Coordination Between CAs and EU Operator

In this article, the EU operator aims to improve grid services by reducing voltage regulation and power loss in the distribution network while aligning with the desired dispatch setpoints from the CAs and optimizing the dispatch of disaggregated DERs. Mathematically, the problem formulation for the proposed approach is as follows:

$$\textbf{min} \quad \sum_{t\in T}\left[\sum_{i\in\mathcal{N}}\left[\gamma_{\text{vd}}\left(V^{\text{node}}_{i,t} - V_{\text{nom}}\right)^2\right]\right.$$
$$\left. + \sum_{(i,j)\in\mathcal{E}}\left[\gamma_{\text{loss}}\left(r_{ij} l^{\text{flow}}_{ij,t}\right)\right]\right] \tag{9a}$$

$$\text{s.t. } (1a)-(1d) \text{ for } \upsilon^{\text{node}}_{i,t} \in \left\{\underline{V}_i, \overline{V}_i\right\}, p^{\text{agg}}_{i,t} \tag{9b}$$

$$p^{\text{agg}}_{i,t} = -\sum_{g\in\mathcal{G}_i} P^{\text{RE}}_{i,g,t} + P^{\text{ch,CS}}_{i,t} + \sum_{j\in\mathcal{N}^{SL}_i} P^{\text{SL}}_{i,j,t} \tag{9c}$$

$$\left(P^{\text{RE,DA}}_{i,g,t}\right)^2 + \left(Q^{\text{RE,DA}}_{i,g,t}\right)^2 \leq \left(S^{\text{RE,DA}}_{i,g,t}\right)^2 \tag{9d}$$

$$-\phi_{i,g} P^{\text{RE,DA}}_{i,g,t} \leq Q^{\text{RE,DA}}_{i,g,t} \leq \phi_{i,g} P^{\text{RE,DA}}_{i,g,t} \tag{9e}$$

$$0 \leq P^{\text{RE,DA}}_{i,g,t} \leq P^{\text{RE,fore}}_{i,g,t} \tag{9f}$$

$$E^{\text{ES}}_{i,t} = E^{\text{ES}}_{i,t-1} + \left[\eta^{\text{ES}}_{\text{ch},i} P^{\text{ch,ES}}_{i,t} - \frac{P^{\text{dch,ES}}_{i,t}}{\eta^{\text{ES}}_{\text{dch},i}}\right]\Delta t \tag{9g}$$

$$(2a)-(2c), (3c)-(3e) \text{ for } E^{\text{ES}}_{i,t}, P^{\text{ch,ES}}_{i,t}, P^{\text{dch,ES}}_{i,t}. \tag{9h}$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:36:46 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

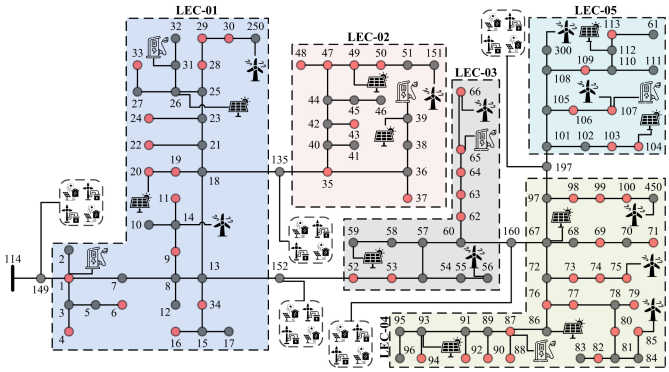


Fig. 5. Modified IEEE 123 bus RDN.

In (9a), the first term aims to enhance node voltages, and the second term minimizes the network power losses. The γ in (9a) denotes for appropriate weights given to the indices in the objective function. The constraints in (9b)–(9h) establish the feasible operations of the DERs and viable network power flow. The constraint (9b) ensures the voltage regulation within specified limits, and (9d)–(9e) regulate the reactive power of the disaggregated RESs. Notably, the EU operator refrains from altering the demand requirements specified by the CAs, thereby eliminating the need for a separate disaggregation problem formulation. Consequently, this approach offers a sequential solution that is computationally efficient and maintains optimality. Unlike during the disaggregation, which often necessitates deviations or curtailments in the preferred energy demand of CA, the proposed method effectively mitigates such suboptimality concerns.

V. Simulation Results and Discussion

The proposed problem is tested on a modified IEEE-123 bus RDN, as shown in Fig. 5. All simulations are performed in A Mathematical Programming Language environment using a KNITRO solver [30] running on an 11th Gen Intel(R) Core(TM) i5-1135G7 @ 2.40 GHz CPU processor with 8 GB RAM.

A. Description of Test System

The test system has been subdivided into five LECs, as illustrated in Fig. 5. Each LEC is equipped with CSs, and RESs in the form of PV plants and WT farms, as well as residential DLs (denoted by nodes in red) along with the fixed load demand (nodes in black). Gathering all relevant parameters for various types of DERs from a single source is considerably challenging. Therefore, the parameters of EV types and charging infrastructure, such as charging power, battery capacity, and maximum energy level, as well as the optimal configuration of CSs, including bus locations, power levels, and charging facilities for different EV types, are obtained from [31]. The power generation profiles for PV plants and WT farms during the peak summer season are adopted from [32], as they reflect real-world scenarios. Since the capacities in [32] are originally provided in megawatts, they are converted to kilowatts to align with the load demand data of the IEEE 123-bus system. For DLs, [33] is referred to define their types and operating parameters, including power level, energy requirements, and total operating time. The net demand for DLs is assumed to vary between 10% and 40% of the power demand

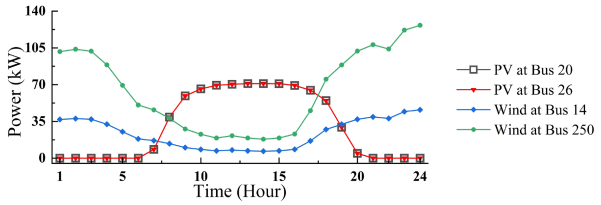


Fig. 6. Hourly power supplied by RESs in LEC-1.

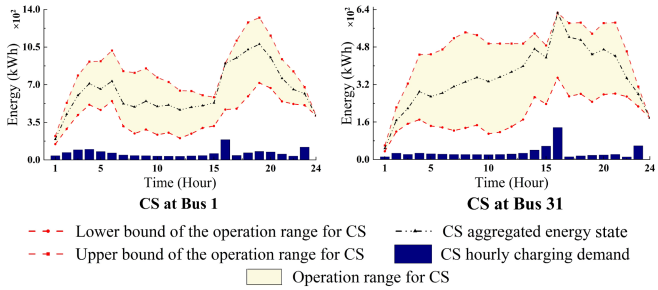


Fig. 7. Operation ranges and the aggregated energy states for CSs at Bus 1 and Bus 31 in LEC-1.

at the red nodes, which is then used to determine the number of DLs connected to each node. For the sake of simplicity, the operational status is presented on an hourly time interval. Moreover, a few DERs, i.e., RESs with integrated ESs, have been strategically placed at bus locations 135, 149, 152, 160, and 197, which are controlled by the EU operator to improve grid services.

B. Community Power Dispatch

To demonstrate the power dispatch within LECs solved by CAs, the numerical findings for the LEC-1 are provided. The LEC-1 configuration comprises PV plants, assumed to be installed at buses 20 and 26, each with an 80 kW capacity, as well as WT farms at bus 14 (80 kW) and bus 250 (220 kW). Furthermore, CSs are placed at buses 1 and 31 to facilitate 140 and 40 EVs with heterogeneous parameters, respectively, as referenced in [31]. In addition, it is assumed that a total of 252 DLs, with a net power demand of 1918.15 kW, are connected across fifteen buses, enabling them to adjust their power demands for ToU price-based demand response. Moreover, the total fixed load demand for LEC-1 is 9435.85 kW, with the hourly demand pattern considered the same for all the nodes.

Fig. 6 depicts the hourly power generation from RESs, showing a peak power output of 212.03 kW during the 18th h period. The net power supplied by the RESs over 24 h amounts to 3596.28 kW. Fig. 7 depicts the feasible operational regions, aggregate energy states, and hourly charging demand for the two CSs at buses 1 and 31. The shaded area between the upper and lower trajectories in Fig. 7 represents the available aggregate energy flexibility for the CSs to effectively meet the charging requirements of EVs. Furthermore, Fig. 8 presents the hourly maximum charging power, aggregate power dispatch for the CSs at buses 1 and 31, and the ToU electricity price derived from [34]. Notably, an observation is made regarding the reduced charging demand of the CSs during the 17th to 22nd time intervals, coinciding with elevated ToU prices, resulting in a significant reduction in operational costs.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:36:46 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

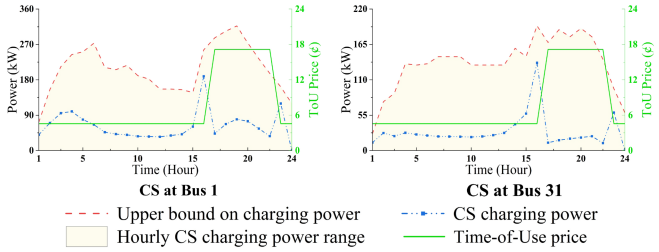


Fig. 8. Aggregate charging power of CSs at Bus 1 and Bus 31 in LEC-1.

TABLE II
DAILY OPERATIONAL DATA FOR CONSIDERED DLS

Term	Interruptible DL			Noninturruptible DL		
DL type	I	II	III	IV	V	VI
Power level (kW)	5.475	4.875	3.5	0.65/1.26	1.875	2
Required time (h)	2	2	4	1/2	2	2
Starting time (h)	8	1	1	4	8	1
Ending time (h)	20	24	24	16	20	12
Energy usage (kWh)	10.95	9.75	14	0.65/2.52	3.75	4

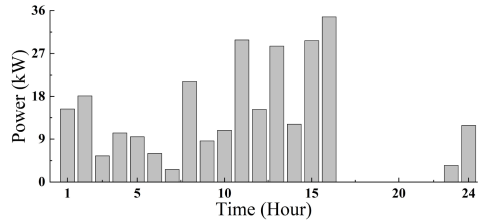


Fig. 9. Net hourly power demand by DLs at Bus 28 in LEC-1.

TABLE III
TOTAL COST OF OPERATION AND POWER DISPATCH IN LECS

Term	LEC-01	LEC-02	LEC-03	LEC-04	LEC-05
Operation cost ($)	1323.16	1105.30	743.29	1397.43	60.82
Power supplied by EU (kW)	10240.29	8390.97	5436.16	11251.65	-171.65
RES generation (kW)	3596.28	3924.36	3876.75	6912.71	5885.36
Power demand from CSs (kW)	2181.05	830.75	857.47	1216.80	811.04

This article investigates six distinct categories of DLs, three following interruptible task characteristics and the other three noninterruptible task characteristics. Table II provides detailed parametric data for each DL, where all DLs except Type IV entail a single operational task, and Type IV involves two operational tasks. In Table II, the second row outlines the power level (kW) for the task, and the third row indicates the required time (h) to complete a task. The fourth and fifth rows collectively outline the operational time frame, indicating the prescribed starting and ending times for each task. Finally, the last row indicates the total energy (kWh) needed to complete a task. The net hourly demand of DLs integrated at bus 28 is depicted in Fig. 9. Notably, there are a total of 90 DLs integrated on bus 28, evenly distributed among the six categories detailed in Table II. It is worth noting that during the high ToU electricity prices from 17th to 22nd time intervals, the power demand of DLs has been shifted to time intervals where ToU electricity prices are comparatively less.

TABLE IV
OVERALL COST AND EVS FEASIBLE ENERGY LIMITS VIOLATION COUNTS FOR THE THREE CASES

Cases	Overall cost ($)	Overall cost in %age of base: Case 1	EVs feasible energy limit violation count	Computation time (s)
Case 1	1323.16	100%	0	0.125
Case 2	1380.56	104.33%	36	0.172
Case 3	1427.14	107.86%	39	0.188

In addition, Table III summarizes the operational costs in all five LECs, along with other significant outcomes such as power supplied by the EU operator in LECs, power generation from RES situated within the LEC, and power demand from the CS in a LEC over the 24 h time horizon. It is observed that due to significant RES generation and low power demand, LEC-5 exports power to the utility, as indicated in Table III. Despite LEC-4 having the highest RES generation capacity among all LECs, it imports a substantial amount of power from the utility. This is attributed to the presence of a large number of fixed and DLs within LEC-4, resulting in 45% higher power demand than LEC-1. However, the operational cost for LEC-4 is 5.6% greater than that of LEC-1, primarily due to having 2.7 times more DL demands in LEC-4 compared to LEC-1. Similarly, although LEC-2 and LEC-3 produce almost the same amount of RES energy and have similar power demands from CSs, LEC-2 ends up operational cost of 49% more than LEC-3. The difference in operational cost comes from LEC-2 needing 4.3 times more fixed load power demand compared to DL demand, where the demand for fixed loads is lower (i.e., 3.6 times) than for DLs in LEC-3.

1) Comparative Analysis: The advantages of the proposed EVs aggregation at the CS framework are presented through the analysis of three cases for quantitative comparison. In each case, the cost-optimal dispatch problem proposed in Section IV is addressed.

1) *Case 1:* This case considers the proposed EVs aggregation framework as outlined in (2)–(3).
2) *Case 2:* For comparison with the proposed EV aggregation framework of case 1, this case does not consider EV aggregation, and instead, models EV at the individual level, as outlined in (2)–(3) except (3f)–(3i).
3) *Case 3:* For comparison with cases 1 and 2, this case does not consider EV aggregation as well as Extn-LP equations utilized to impose strict conditions on EV charging and discharging power capacity. The equations modeling EVs in this case are (2)–(3), excluding (3f)–(3i), (2a), (3c), and (3d).

Table IV reports the comparison between the three cases. The overall cost outlined in Table IV includes operational costs, and penalty costs resulting from the violation of the feasible energy limit constraints of EVs. It is assumed that the penalty cost is set at a rate of 10 ¢/kWh more than the ToU rate ($\lambda^{\text{ToU}}_{\text{CS},t}$) applied for CS during normal conditions [16]. Table IV shows that the overall cost in case 1 amounts to $1323.16 while adopting the proposed EV aggregation framework. Moreover, there is a significant increase in overall cost observed in cases 2 and 3, showing a 4.33% and 7.86% rise, respectively. As shown in the third column of Table IV, the feasible energy limits are violated 36 and 39 times in cases 2 and 3, respectively, mainly because

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:36:46 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

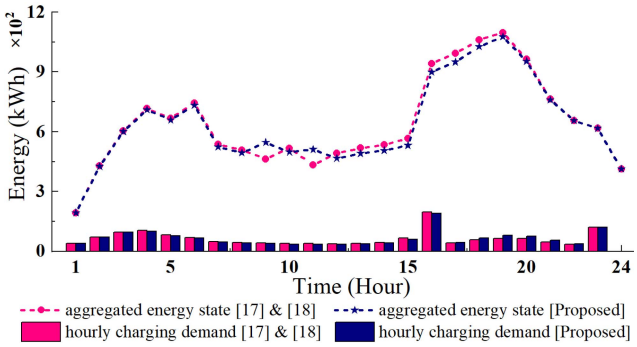


Fig. 10. Comparison of aggregated energy states and hourly charging demand at CS at Bus 1 for EV aggregation methods in [17] and [18] and the proposed method.

of not considering the EVs aggregation in case 2 and EVs aggregation as well as Ext-LP equations in case 3. Computational time (s) of the proposed EVs aggregation framework in case 1 is much less as compared to cases 2 and 3, as shown in Table IV.

2) Comparison of Proposed and Existing EV Aggregation Methods: This section presents a comparative analysis of the proposed approach and the methodologies outlined in [17] and [18] for EVs aggregation at CS. The simulation results indicate that, for both the proposed method and those in [17] and [18], the overall energy state of the CSs remains within their upper and lower flexibility boundaries, derived from the individual EV flexibility limits in each approach, as illustrated in Fig. 3. However, a closer examination of the individual energy states of departing EVs reveals that the proposed method effectively manages the desired energy levels at departure while respecting individual EV charging preferences and optimizing economic efficiency. In contrast, methods in [17] and [18] show that EVs tend to depart with energy states higher than their desired levels during low-price intervals, whereas they often leave with a lower-than-desired energy state during high-price intervals. This occurs because CSs adopt a strategy of maintaining higher energy levels during low-price periods to optimize cost-efficiency when prices rise. Consequently, EVs leaving during high-price intervals experience reduced charging, balancing out the surplus energy provided to EVs departing during low-price intervals.

For instance, during the 16th-h interval, five EVs are scheduled to depart from the CS at Bus 1. Under the proposed method, the simulated energy state is 898.72 kWh, as shown in Fig. 10. In contrast, the approaches in [17] and [18] resulted in a higher energy state of 942.2 kWh, corresponding to the upper bound of the energy state. This difference indicates that, under these strategies, departing EVs collectively received an additional 43.48 kWh beyond their intended energy levels. Over a 24-h period, EVs departing between time intervals 1 and 16 stored a total of 48.13 kWh more energy under the methods in [17] and [18] than under the proposed approach. Conversely, for EVs departing between time intervals 17 and 22, the stored energy is 48.13 kWh lower than their intended levels. This trend follows electricity pricing variations, where these methods reduce energy intake during high-price periods to maintain economic efficiency. As a result, EVs may either incur additional charging costs due to excess energy or receive a charge level lower than their desired state when electricity prices are high, meaning individual EV preferences are not fully accounted for. The proposed method, however, not only optimizes costs at a system-wide level but also integrates individual charging preferences, improving both user satisfaction and operational efficiency.

TABLE V
COMPARISON OF DIFFERENT VaR MEASURES FOR UNCERTAINTY ANALYSIS WITH CHANGE IN CONFIDENCE INTERVAL

Confidence interval	VaR (%)	CVaR (%)	MVaR (%)
0.01	0.10	0.10	0.091
0.025	0.091	0.096	0.088
0.05	0.084	0.092	0.083
0.1	0.075	0.088	0.077

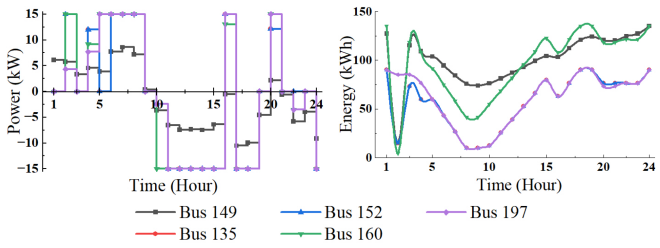


Fig. 11. Discharging power and energy states of disaggregated ESs.

3) Impact of Uncertainty on Cost Optimal Community Power Dispatch: This section investigates the impact of uncertainty in the departure times of EVs on the cost-optimal operations of CAs. The impact of this uncertainty is analyzed using risk measurement tools [35], such as value-at-risk (VaR), conditional VaR, and modified VaR. For the analysis, it is assumed that on any given day, 20% of the total EVs arriving at CSs may depart earlier than scheduled, with a standard deviation of 4 h. Based on this assumption, 100 scenarios are generated addressing the cost-optimal operation in the energy community. The mean outcome of these scenarios is 1323.75 $, with a standard deviation of 0.32 $. The VaR outcomes of the analysis are detailed in Table V, presenting measures for various values of the risk-tolerance level (i.e., 0.01, 0.025, 0.05, and 0.1). The percentage values in Table V represent the worst-case loss relative to the expected value within the given confidence interval, establishing a probabilistic threshold beyond which the operational cost is unlikely to exceed, given a specified confidence level. Notably, a lower confidence interval may lead to higher operational costs but enhances the robustness against the operational cost variation. For example, the simulation results show that the worst-case operational cost is $1323.56 for a confidence interval of 0.99. As the confidence interval decreases to 0.1, 0.01, and 0.001, the worst-case cost estimates rise to $1324.15, $1324.48, and $1324.75, respectively.

C. Dispatch Set-Points From EU Operator

This section presents the numerical results pertaining to the dispatch set-points evaluated for the disaggregated DERs within the network, aiming to improve grid services. Fig. 11 illustrates the discharge power and energy state patterns over 24 h for disaggregated ESs located at buses 149, 135, 152, 160, and 197. Notably, the results indicate that during periods of high power demand from LECs, the ESs discharge, whereas during low-demand intervals, the ESs charge to ensure demand fulfillment locally. For instance, during the 15th h, when demand is low, all ES units engage in charging, while in the 16th h, when LECs experience high demand, all ES units except the one at

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:36:46 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE VI
COMPARISON OF POWER LOSSES (KW) AND VOLTAGE DEVIATION (V^2) UNDER VARIOUS SCENARIOS

	γ_{loss} × power loss (kW)	γ_{vd} × voltage deviation (V^2)
Without DERs (No RESs and ESs)	1003.15	552.10
With RESs (No ESs)	855.49	485.46
With DERs (RESs+ESs) [Proposed]	846.68	483.87

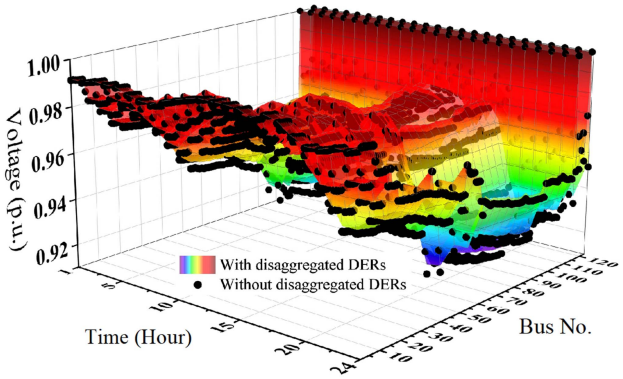


Fig. 12. Voltage profile for the system considered, with and without the dispatch of disaggregated DERs.

bus 149 switch to maximum discharging mode. Furthermore, a comparative analysis of power losses and voltage deviations in various scenarios is detailed in Table VI. It is worth mentioning that when the EU operator does not dispatch the disaggregated DERs to fulfill LECs' energy demand and when the EU operator utilizes the full potential of DERs, the net power losses are notably improved by 15.60%. Fig. 12 illustrates the bus voltage magnitudes (in pu) for the system considered with (colormap surface) and without (scatter) the dispatching disaggregated DERs, revealing a significant improvement in voltage profile. The findings reveal that effective management of disaggregated DERs in coordination with the dispatch set-points received from LECs significantly enhances grid performance.

VI. CONCLUSION AND FUTURE WORK

In conclusion, this article has focused on the development of an efficient framework for the management of a substantial fleet of DERs within the distribution system aimed at optimizing operational costs in LECs and improving grid services at the EU level. The approach first developed aggregation models to determine the cumulative energy flexibility of DERs namely EVs and DLs, followed by a hierarchical control approach for solving the optimal dispatch set-points for these DERs. In addition, a more computationally tractable linear formulation was adopted for EV and ES battery modeling. Through numerical simulations conducted on a modified IEEE-123 bus RDN, the effectiveness of the proposed method was demonstrated.

The quantitative analysis reveals that the proposed EV aggregation framework provides a more accurate and computationally efficient solution duly incorporating network constraints, which makes the framework more practically achievable, especially while managing a large fleet of EVs having spatial placement and temporal behavior. From the numerical analysis, it can be concluded that the presence of a large number of DLs within the distribution network significantly reduces operational costs in energy communities by facilitating the proposed DLs aggregation model for price-based demand response. Moreover, the effective coordination among disaggregated DERs and CAs exerts great influence on power losses and bus voltages in the distribution network. In conclusion, the proposed HCF successfully enhances the overall performance of the electrical grid.

The results of these simulations highlight the capacity of the approach to effectively manage the sizable fleet of DERs in the distribution system, delivering satisfactory outcomes in terms of economic cost and secure system operations, all while maintaining a high level of computational efficiency. The future work may be extended to analyze the proposed work with an unbalanced three-phase network model.

REFERENCES

[1] G. E. Asimakopoulou and N. D. Hatziargyriou, "Evaluation of economic benefits of DER aggregation," *IEEE Trans. Sustain. Energy*, vol. 9, no. 2, pp. 499–510, Apr. 2018.
[2] C. Wei, J. Xu, S. Liao, and Y. Sun, "Aggregation and scheduling models for electric vehicles in distribution networks considering power fluctuations and load rebound," *IEEE Trans. Sustain. Energy*, vol. 11, no. 4, pp. 2755–2764, Oct. 2020.
[3] "Distributed energy resources for net zero: An asset or a hassle to the electricity grid?," International Energy Agency, Paris, Tech. Rep., Sep. 2021. [Online]. Available: https://www.iea.org/commentaries/distributed-energy-resources-for-net-zero-an-asset-or-a-hassle-to-the-electricity-grid
[4] "Empowering people–the role of local energy communities in clean energy transitions," International Energy Agency, Paris, Tech. Rep., Aug. 2023. [Online]. Available: https://www.iea.org/commentaries/empowering-people-the-role-of-local-energy-communities-in-clean-energy-transitions
[5] L. Šikšnys, E. Valsomatzis, K. Hose, and T. B. Pedersen, "Aggregating and disaggregating flexibility objects," *IEEE Trans. Knowl. Data Eng.*, vol. 27, no. 11, pp. 2893–2906, Nov. 2015.
[6] F. Plaum, R. Ahmadiahangar, A. Rosin, and J. Kilter, "Aggregated demand-side energy flexibility: A comprehensive review on characterization, forecasting and market prospects," *Energy Rep.*, vol. 8, pp. 9344–9362, 2022.
[7] S. Barot and J. A. Taylor, "A concise, approximate representation of a collection of loads described by polytopes," *Int. J. Elect. Power Energy Syst.*, vol. 84, pp. 55–63, 2017.
[8] Z. Yi, Y. Xu, W. Gu, L. Yang, and H. Sun, "Aggregate operation model for numerous small-capacity distributed energy resources considering uncertainty," *IEEE Trans. Smart Grid*, vol. 12, no. 5, pp. 4208–4224, Sep. 2021.
[9] Z. Yi, Y. Xu, W. Gu, and W. Wu, "A multi-time-scale economic scheduling strategy for virtual power plant based on deferrable loads aggregation and disaggregation," *IEEE Trans. Sustain. Energy*, vol. 11, no. 3, pp. 1332–1346, Jul. 2020.
[10] R. R. Appino, V. Hagenmeyer, and T. Faulwasser, "Towards optimality preserving aggregation for scheduling distributed energy resources," *IEEE Trans. Control Netw. Syst.*, vol. 8, no. 3, pp. 1477–1488, Sep. 2021.
[11] Y. Wen, Z. Hu, and L. Liu, "Aggregate temporally coupled power flexibility of DERs considering distribution system security constraints," *IEEE Trans. Power Syst.*, vol. 38, no. 4, pp. 3884–3896, Jul. 2023.
[12] X. Chen, E. Dall'Anese, C. Zhao, and N. Li, "Aggregate power flexibility in unbalanced distribution systems," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 258–269, Jan. 2020.
[13] N. Nazir and M. Almassalkhi, "Grid-aware aggregation and realtime disaggregation of distributed energy resources in radial networks," *IEEE Trans. Power Syst.*, vol. 37, no. 3, pp. 1706–1717, May 2022.
[14] D. Yan, C. Ma, and Y. Chen, "Distributed coordination of charging stations considering aggregate EV power flexibility," *IEEE Trans. Sustain. Energy*, vol. 14, no. 1, pp. 356–370, Jan. 2023.
[15] V. Rigoni, D. Flynn, and A. Keane, "Coordinating demand response aggregation with LV network operational constraints," *IEEE Trans. Power Syst.*, vol. 36, no. 2, pp. 979–990, Mar. 2021.
[16] K. Utkarsh, F. Ding, X. Jin, M. Blonsky, H. Padullaparti, and S. P. Balamurugan, "A network-aware distributed energy resource aggregation framework for flexible, cost-optimal, and resilient operation," *IEEE Trans. Smart Grid*, vol. 13, no. 2, pp. 1213–1224, Mar. 2022.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:36:46 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

[17] M. Zhou, Z. Wu, J. Wang, and G. Li, "Forming dispatchable region of electric vehicle aggregation in microgrid bidding," *IEEE Trans. Ind. Inform.*, vol. 17, no. 7, pp. 4755–4765, Jul. 2021.
[18] X. Shi, Y. Xu, G. Chen, and Y. Guo, "An augmented Lagrangian-based safe reinforcement learning algorithm for carbon-oriented optimal scheduling of EV aggregators," *IEEE Trans. Smart Grid*, vol. 15, no. 1, pp. 795–809, Jan. 2024.
[19] N. Nazir and M. Almassalkhi, "Guaranteeing a physically realizable battery dispatch without charge-discharge complementarity constraints," *IEEE Trans. Smart Grid*, vol. 14, no. 3, pp. 2473–2476, May 2023.
[20] Y. Wen, Z. Hu, S. You, and X. Duan, "Aggregate feasible region of DERs: Exact formulation and approximate models," *IEEE Trans. Smart Grid*, vol. 13, no. 6, pp. 4405–4423, Nov. 2022.
[21] M. R. Islam, H. Lu, J. Hossain, M. R. Islam, and L. Li, "Multiobjective optimization technique for mitigating unbalance and improving voltage considering higher penetration of electric vehicles and distributed generation," *IEEE Syst. J.*, vol. 14, no. 3, pp. 3676–3686, Sep. 2020.
[22] J. M. Morales, S. Pineda, and Y. Dvorkin, "Learning the price response of active distribution networks for TSO-DSO coordination," *IEEE Trans. Power Syst.*, vol. 37, no. 4, pp. 2858–2868, Jul. 2022.
[23] L. Wang, A. Dubey, A. H. Gebremedhin, A.K. Srivastava, and N. Schulz, "MPC-based decentralized voltage control in power distribution systems with EV and PV coordination," *IEEE Trans. Smart Grid*, vol. 13, no. 4, pp. 2908–2919, Jul. 2022.
[24] S. Zafar, A. Blavette, G. Camilleri, H. Ben Ahmed, and J.-J. Prince Agbodjan, "Decentralized optimal management of a large-scale EV fleet: Optimality and computational complexity comparison between an adaptive MAS and MILP," *Int. J. Electr. Power Energy Syst.*, vol. 147, 2023, Art. no. 108861, doi: 10.1016/j.ijepes.2022.108861.
[25] M. Ostadijafari, J. C. Bedoya, W. Wang, A. Dubey, C.-C. Liu, and N. Yu, "Proactive demand-side participation: Centralized versus transactive demand-supply coordination," *Electric Power Syst. Res.*, vol. 206, 2022, Art. no. 107789, doi: 10.1016/j.epsr.2022.107789.
[26] A. Singla, K. Singh, and V. K. Yadav, "Optimization of distributed solar photovoltaic power generation in day-ahead electricity market incorporating irradiance uncertainty," *J. Mod. Power Syst. Clean Energy*, vol. 9, no. 3, pp. 545–560, 2021.
[27] H. Patil and V. N. Kalkhambkar, "Grid integration of electric vehicles for economic benefits: A review," *J. Mod. Power Syst. Clean Energy*, vol. 9, no. 1, pp. 13–26, 2021.
[28] M. Farivar and S. H. Low, "Branch flow model: Relaxations and convexification—Part I & II," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2554–2564, Aug. 2013.
[29] D. Pozo, "Linear battery models for power systems analysis," *Electric Power Syst. Res.*, vol. 212, 2022, Art. no. 108565, doi: 10.1016/j.epsr.2022.108565.
[30] R. A. Waltz and T. D. Plantenga, "Knitro user's manual: Version 6.0," Ziena Optimization, Inc., Mar. 2009. [Online]. Available: https://www.yumpu.com/en/document/view/37639225/knitro-users-manual-version-60-ziena-optimization-inc
[31] S. Gao, J. Wu, and B. Xu, "Controllability evaluation of EV charging infrastructure transformed from gas stations in distribution networks with renewables," *Energies*, vol. 12, no. 8, 2019, Art. no. 1577.
[32] "Renewables trend," 2023. [Online]. Available: https://www.caiso.com/TodaysOutlook/Pages/supply.html
[33] Z. Chen, L. Wu, and Y. Fu, "Real-time price-based demand response management for residential appliances via stochastic optimization and robust optimization," *IEEE Trans. Smart Grid*, vol. 3, no. 4, pp. 1822–1831, Dec. 2012.
[34] "City of fort collins utilities residential rates," 2023. [Online]. Available: https://www.fcgov.com/utilities/residential/rates/electric
[35] R. Rockafellar and S. Uryasev, "Conditional value-at-risk for general loss distributions," *J. Bank. Finance*, vol. 26, no. 7, pp. 1443–1471, 2002.



Shubham Gupta received the B.E. degree in electrical engineering from the College of Technology and Engineering, Udaipur, India, in 2012, the M.Tech. degree in power engineering from Guru Nanak Dev Engineering College, Ludhiana, India, in 2015. He is currently working toward the Ph.D. degree in electrical engineering with a research focus on the techno-economic analysis and optimization of distributed energy resources in modern distribution systems with the Delhi Technological University, New Delhi, India.

His research interests include electricity pricing, distributed energy resources, and energy system modeling.



Vinod Kumar Yadav (Senior Member, IEEE) received the B.Tech. degree in electrical engineering from the Institute of Engineering and Technology, Bareilly, India, in 2003, the M.Tech. degree in power system engineering from the National Institute of Technology, Jamshedpur, India, in 2005, and the Ph.D. degree in power system engineering from the Indian Institute of Technology, Roorkee, India, in 2011.

His research interests include renewable energy systems, power system planning and optimization, distributed generation, and smart grid.



Madhusudan Singh (Senior Member, IEEE) was born in Ghazipur, India, in 1968. He received the B.Sc. degree in electrical engineering from the Faculty of Technology, Dayalbagh Educational Institute, Dayalbagh, India, in 1990, the M.E. degree in control and instrumentation from the University of Allahabad, Allahabad, India, in 1992, and the Ph.D. degree in electrical engineering from the Faculty of Technology, University of Delhi, New Delhi, India, in 2006.

His research interests include modeling and analysis of electrical machines, power quality in power systems, renewable energy systems, voltage control aspects of self-excited induction generators, power electronics, and drives.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:36:46 UTC from IEEE Xplore. Restrictions apply.

Posted on 25 Mar 2025 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.174291179.92552739/v1 — This is a preprint and has not been peer-reviewed. Data may be preliminary.

AI-Driven Mental Healthcare 5.0: A Survey of Opportunities and Challenges in Leveraging Large Language Models and Generative AI

Akshi Kumar[1], Karan Gupta[2], Kartik Verma[2], and Sanjay Kumar[2]

[1]Goldsmiths University of London Department of Computing
[2]Delhi Technological University

March 25, 2025

Abstract

The integration of Large Language Models (LLMs) and Generative AI into mental healthcare is reshaping diagnostic, therapeutic, and support systems, marking a transition toward AI-driven mental healthcare. This paper explores various methodologies for incorporating LLMs into mental health applications, categorizing research into key strategies such as prompt engineering, fine-tuning, emotional theory integration, AI-powered chatbots, and long-term memory architectures. Through case studies, we evaluate their effectiveness in areas like depression detection, suicide risk assessment, and AI-assisted psychological counseling, highlighting both their potential benefits and associated challenges. While LLMs enable more empathetic interactions, personalized interventions, and early risk identification, concerns remain regarding bias, misinformation, safety risks, and ethical issues related to privacy, informed consent, and AI's role in clinical decision-making. We outline critical research gaps, stressing the importance of real-world validation, standardized evaluation methods, and strong risk mitigation frameworks. This paper serves as a guide for the responsible, scalable, and ethical integration of AI in mental healthcare, ensuring these technologies enhance patient well-being while adhering to clinical and ethical standards.

ORIGINAL ARTICLE

AI-Driven Mental Healthcare 5.0: A Survey of Opportunities and Challenges in Leveraging Large Language Models and Generative AI

Akshi Kumar[1] | **Karan Gupta**[2] | **Kartik Verma**[2] | **Sanjay Kumar**[2]

[1]Department of Computing, Goldsmiths, University of London, United Kingdom

[2]Department of Computer Science and Engineering, Delhi Technological University, New Delhi, Delhi, 110042, India

Correspondence
Sanjay Kumar
Email: sanjay.kumar@dtu.ac.in

Funding information
-NA

The integration of Large Language Models (LLMs) and Generative AI into mental healthcare is reshaping diagnostic, therapeutic, and support systems, marking a transition toward AI-driven mental healthcare. This paper explores various methodologies for incorporating LLMs into mental health applications, categorizing research into key strategies such as prompt engineering, fine-tuning, emotional theory integration, AI-powered chatbots, and long-term memory architectures. Through case studies, we evaluate their effectiveness in areas like depression detection, suicide risk assessment, and AI-assisted psychological counseling, highlighting both their potential benefits and associated challenges. While LLMs enable more empathetic interactions, personalized interventions, and early risk identification, concerns remain regarding bias, misinformation, safety risks, and ethical issues related to privacy, informed consent, and AI's role in clinical decision-making. We outline critical research gaps, stressing the importance of real-world validation, standardized evaluation methods, and strong risk mitigation frameworks. This paper serves as a guide for the responsible, scalable, and ethical integration of AI in mental healthcare, ensuring these technologies enhance patient well-being while adhering to clinical and ethical standards.

KEYWORDS
Large Language Models (LLM), Mental Health, Healthcare 5.0, Generative AI, Chatbots

1 | INTRODUCTION

Millions of people around the world suffer from mental health-related issues, which underscores the urgent need for accessible, efficient, and compassionate mental healthcare. The increasing prevalence of stress, depression, suicide, and other mental health conditions poses a significant public health challenge globally Ji et al. (2022a); Sagastibeltza et al. (2023); Sharma et al. (2023b). Mental health extends beyond the absence of illness, encompassing emotional resilience, psychological well-being, and the ability to navigate life's challenges Saraceno (2002). Recognizing its fundamental role in overall health, societies increasingly prioritize mental health awareness, destigmatization, and investment in robust healthcare infrastructureBhugra et al. (2013). The evolution of Healthcare 5.0 envisions AI-driven personalized and preventive care, making scalable mental health interventions more accessible and adaptive. This has led to a growing interest in Artificial Intelligence (AI)-driven solutions, particularly Large Language Models (LLMs)

and Generative AI, which offer innovative approaches for mental health assessment, support, and intervention James (2020). Traditional psychological state assessments relied on machine learning (ML) techniques such as Support Vector Machines (SVM), Neural Networks, AdaBoost, k-nearest Neighbors (kNN), Deep Learning, Decision Trees (DT), and Clustering Shatte et al. (2019); Benfares et al. (2021). Despite demonstrating efficacy, these strategies frequently require extensive labeled data and manual feature design.

Unlike traditional methods, Generative AI and LLMs leverage a transformative approach by employing extensive pre-trained language models with self-supervised training. LLMs can engage in human-like conversations, detect subtle linguistic cues, and offer personalized emotional support Nowacki et al. (2024). Their ability to analyze textual data—ranging from social media posts to therapy transcripts—enables early detection of stress, depression, and suicide risk Frueh (2015). Figure 1 depicts the distribution of prevalent mental health concerns across diverse populations, with the most frequently discussed issues being anxiety and depression.

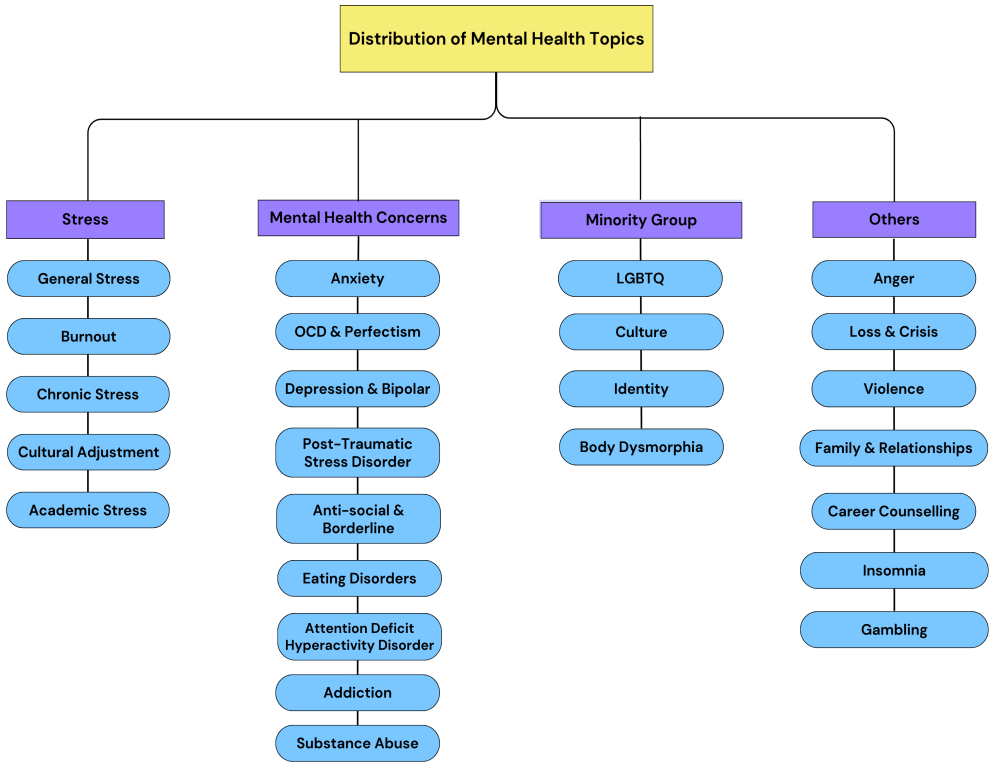


FIGURE 1 The distribution of mental health-related topics.

In addition, Generative AI enhances mental health research by synthesizing data, generating context-aware responses, and simulating therapeutic interactions. These models facilitate the creation of synthetic datasets that reflect diverse mental states, aiding in the fine-tuning of diagnostic algorithms and expanding the scope of AI-driven interventions Mandapuram et al. (2018); Lau et al. (2025). This integration presents a dynamic and adaptive framework for mental health assessment, aligned with the evolving nature of psychological conditions Naveed et al. (2023).

However, the adoption of LLMs and Generative AI in mental healthcare presents ethical, technical, and interpretability challenges Ning et al. (2024); Solaiman (2024). Issues such as bias in training data, privacy concerns, AI hallucinations, and the potential for misinformation must be critically addressed to ensure responsible deployment. This paper provides a comprehensive overview of state-of-the-art methodologies, emerging challenges, and future research directions at the intersection of LLMs, Generative AI, and mental health detection Fan et al. (2023).

The following selection criteria were adopted for the research papers:

- **Time Period**:This study examines advancements in the application of Large Language Models (LLMs) in mental health from 2019 to 2025, capturing recent breakthroughs while incorporating earlier literature to provide a comprehensive contextual overview.
- **Keywords**: A targeted keyword search was employed, including terms such as 'Large Language Models,' 'Generative AI,' 'Mental Health,' 'ChatGPT,' 'Ethics in AI,' and 'Context Learning.' This approach ensures a focused and exhaustive review of the most relevant developments in the field.
- **Inclusion Criteria**: Only peer-reviewed conference and journal papers written in English were considered. The study prioritizes the exploration of LLMs in mental health, highlighting their applications, methodologies, and ethical challenges.
- **Exclusion Criteria**: In order to maintain the study's validity, works that were published before 2019, or lacking direct relevance to the research focus were not included.

In order to have a comprehensive understanding of LLMs in the context of mental health, this research focuses on conducting a detailed analysis of their advantages, limitations, and ethical implications. This study will not only gives us insights of LLMs application in the field of mental health but will also enable us to discover techniques through which they can be implemented. The key contributions of this article in mental health are as follows.

- Categorizes research on LLM applications in mental health, covering prompt engineering, fine-tuning, emotional theory integration, AI chatbots, and long-term memory mechanisms.
- Examines depression detection, suicide risk assessment, and AI-driven psychological counseling, evaluating model performance across diverse methodologies.
- Assesses the strengths and limitations of integrating LLMs in mental health, highlighting concerns such as bias, safety risks, and reasoning failures.
- Explores privacy, informed consent, algorithmic bias, and risks of misuse, emphasizing the need for transparent and fair AI deployment.
- Identifies areas that need further exploration, including real-world clinical validation, standardized evaluation frameworks, and risk mitigation strategies.
- Provides actionable insights and recommendations to advance the ethical and effective integration of LLMs in mental healthcare.

The flow of information in this survey is structured as follows. Section 2 outlines the fundamental concepts of Large Language Models (LLMs) and their benchmarks. Section 3 pans across the current state of LLMs applied in mental health contexts. A taxonomy is proposed to categorize existing research approaches based on techniques used, including prompt engineering, fine-tuning strategies, integration of psychological theory, and conversational agents. Section 4 covers the datasets which most appear in the literature. Section 5 closely examines the ethical dimensions and practical challenges accompanying the use of LLMs for mental health purposes. Protecting personal data privacy, ensuring proper consent, avoiding misuse, and confronting algorithmic biases represent some of the complex considerations discussed. Section 6 discusses research limitations and opportunities to responsibly advance the integration of LLMs and mental healthcare. Section 7 concludes the article by summarizing the comprehensive integration of LLMs and mental health.

2 | BACKGROUND

2.1 | Large Language Models

LLMs have now become powerful AI systems, able to process and create coherent text while showing generalization across many tasks Radford et al. (2019); Brown et al. (2020). They are trained through text-based data, which helps them improve their ability to predict the next word or phrase in a sentence. Over the years, natural language processing (NLP) has advanced a lot, moving from statistical methods to neural language modeling and then evolving further from pre-trained language models (PLMs) to large language models (LLMs).

Early work on LLMs, like T5 Raffel et al. (2019) and mT5 Xue et al. (2021), mainly used transfer learning. It wasn't until GPT-3 Brown et al. (2020) came along that LLMs showed zero-shot transferability to downstream tasks without needing fine-tuning. Even though pre-trained LLMs can answer task queries based on examples and descriptions, their performance in zero-shot scenarios is still lower compared to few-shot learning situations. To improve generalization for unseen tasks, fine-tuning with task-specific data is suggested Chung et al. (2022a); Wang et al. (2022b,a). This leads to better zero-shot performance and helps minimize misaligned behavior.

LLMs demonstrate these competencies due to their extensive scale, even when no explicit training is provided to instill them. The LLM's capability to tackle diverse tasks with human-level performance is accompanied by drawbacks such as slow training and inference, substantial hardware demands, and increased operational expenses. These requirements have constrained their widespread adoption and created opportunities for the development of more efficient architectures Zeng et al. (2023); Wang et al. (2023b). Table 1 lists some of the major Large Language Models that can be employed for intervention in mental health.

TABLE 1 Basic details of Large Language Models released in recent years with varying sizes. Here parameters column corresponds to the number of trainable parameters in each LLM where M, B, and T stand for millions, Billions, and Trillions, respectively

LLM	Parameters	Focus Areas	Availability	Architecture
T5 Raffel et al. (2019)	11B, 110B	Text-to-text tasks, summarization, translation, question answering	Publicly available (API)	Encoder-decoder architecture
GPT-3 Brown et al. (2020)	175B	Text generation, translation, question answering, creative text formats	OpenAI API (paid access)	Generative Pre-trained Transformer model
mT5 Xue et al. (2021)	10B, 100B	Text-to-text tasks, multilingual translation	Publicly available (code)	Multimodal variant (mT5-XXL)
PanGu Zeng et al. (2021)	200B	Chinese language processing, text generation, translation	Limited access for research	Transformer-based architecture
ERNIE 3.0 Sun et al. (2021b)	10B	Knowledge-enhanced tasks, semantic understanding, reasoning	Publicly available (code)	Knowledge-infused model architecture
Jurassic-1 Lieber et al. (2021)	178B	Text-to-text tasks, question answering, summarization	AI21 Labs API (paid access)	Jurassic-1 Jumbo variant (355B)

Continued on next page

Table 1 – Continued from previous page

LLM	Parameters	Focus Areas	Availability	Architecture
Yuan 1.01 Wu et al. (2021)	245B	Chinese language processing, text generation, translation	Limited access for research	Transformer-based architecture
Gopher Rae et al. (2021)	280B	Text-to-text tasks, question answering, summarization	Publicly available (code)	Dense retrieval capabilities
ERNIE 3.0 Titan Wang et al. (2021)	10B	Knowledge-enhanced tasks, semantic understanding, reasoning	Publicly available (code)	Knowledge-infused model architecture
GPT-NeoX-20B Black et al. (2022)	20B	Text generation, translation, question answering	Publicly available (code)	Open-source alternative to GPT-3
OPT Zhang et al. (2022)	125M	Text-to-text tasks, efficient training	Publicly available (code)	Optimized model architecture for efficiency
AlphaCode Li et al. (2022)	530B	Code generation, competitive programming	DeepMind (limited access)	Transformer-based architecture
AlexaTM Soltan et al. (2022)	137B	Text-to-text tasks, question answering, conversational AI	Amazon Research (limited access)	Multimodal capabilities (text, audio, visual)
GPT-4 OpenAI et al. (2023)	100T / 800T (estimated)	Text generation, translation, reasoning, multimodality	Limited access (research & select partners)	Likely uses novel architectures & training techniques
FLAN-T5 Chung et al. (2022b)	137B / 540B	Multilingual tasks, translation, summarization	Google Research (limited access)	Dense retrieval capabilities
Alpaca Taori et al. (2023)	137B	Chinese language processing, text generation, translation	Limited access for research	Transformer-based architecture with domain-specific optimizations
Alpaca-LoRA Alpaca-LoRA (2023)	137B	Chinese language reasoning, factual language understanding	Limited access for research	Combines Alpaca with Long-Range Arena reasoning techniques
ChatGLM-6B Zeng et al. (2022)	6B	Conversational AI, dialogue, open-ended tasks	Tsinghua University (limited access)	Emphasis on long-term conversation coherence
BELLE-LLAMA-7B-2M Ji et al. (2023)	7B	Chinese dialogue tasks, factual language understanding	Lianjia Technology (limited access)	Focuses on improving factual accuracy and consistency in dialogue

Continued on next page

Table 1 – Continued from previous page

LLM	Parameters	Focus Areas	Availability	Architecture
Vicuna-7B-v1.1 Chiang et al. (2023)	7B	Chinese language processing, text generation, translation	Hugging Face AI (limited access)	Optimized for efficiency and runtime performance
Baichuan 2 **?**	7B, 13B	Multilingual tasks, domain-specific applications	Open-source	Trained on 2.6 trillion tokens
Command R+ Cohere (2024)	104B	Real-world challenges, retrieval augmented generation	API and direct download	Supports multiple tools for complex tasks
OpenAI o1 OpenAI (2024)	300B	Scientific problem-solving, coding, complex mathematics	API access only	Emphasizes reasoning with extensive thought process
Gemini 2.0 DeepMind (2024)	Not disclosed	Multimodal processing (text, images, audio, video)	API access only	Designed for the "agentic era" of AI
Llama 3.3 70B Instruct AI (2024)	70B	Instruction tuning, multilingual support	Open-source	Advanced language model with diverse applications
DeepSeek R1 AI (2025)	671B	Text generation, reasoning, efficient processing	Open-source adaptation available	Reinforcement learning with mixture of experts approach

2.2 | Benchmarks for LLMs

Several benchmarks have been developed to evluate the performance of Large Language Models (LLMs) across diverse tasks. Hendrycks et al. (2021) created a through evaluation test for assessing text models. Liang et al. (2023) offered a comprehensive assessment, evaluating large language models across various aspects such as language understanding and common-sense reasoning. Srivastava et al. (2023) introduced 204 challenging tasks covering various domains, with the goal of evaluating tasks beyond the capabilities of existing language models. C-EVAL estimates the developed knowledge and reasoning capabilities of foundation models in the context of the Chinese language Huang et al. (2023). Zhong et al. (2023) served as an evaluation framework named AGI-Eval to assess the performance of foundation models in human-centric standardized exams. Although these benchmarks are diverse and comprehensive, they focus majorly on general capabilities and do not specifically account for the nuances of mental health.

2.3 | Mental Health Benchmarks for LLMs

MentalLLM is designed to evaluate the capability of LLMs to assess mental health outputs through the examination of online text data Xu et al. (2023). PsyEval acknowledged that symptoms of mental disorders are nuanced, subjective, and highly personalized, it recognizes the need for expertise, empathy, and awareness of emergency response, aspects not covered in other benchmarks. This entails grasping subtle emotional states, identifying nuanced signs of mental distress, and delivering safe and empathetic interactions Jin et al. (2023). Dialogue safety focuses on understanding the safety of replies rendered by LLMs in mental health support Qiu et al. (2023). MultiMedQA Singhal et al. (2022) is specifically tailored for medical question-answering, assessing LLMs based on their clinical knowledge and

question-answering abilities. These benchmarks are devised for certain downstream tasks, offering a more specialized evaluation of LLMs.

3 | CURRENT STATE OF LLMS IN MENTAL HEALTH

This section delves into the LLMs' role in understanding, diagnosing, and providing support for mental health concerns. We provide a glance at the different model customization processes employed and developed to tackle mental healthcare problems. We provide an overview of the proposed taxonomy scheme, viz., Prompt Engineering, Chatbots, Fine-Tuning, Pre-training, Emotional & Linguistic Context, and Long-Term Memory, of segregating the current research landscape based on the method and technique used. The proposed taxonomy seeks to categorize the primary applications of LLMs into overarching groups, allowing for a comprehensive examination of each category and its potential in greater detail.

3.1 | Prompt Engineering

Prompt engineering is designing and tweaking prompts to guide LLMs in producing the right kind of responses. In mental healthcare, this process is especially important because well-crafted prompts can provide the model with the right context, helping it generate more accurate and relevant responses. To make sure the model performs well, prompts need to be carefully designed, taking into account not just technical aspects but also human interaction and ethical considerations Kamalloo et al. (2023); Liu et al. (2023a).

Xu et al. (2023) conducted a comprehensive evaluation of five Large Language Models (GPT-3.5, GPT-4, FLAN-T5, Alpaca, and Alpaca-LoRA) for mental healthcare applications, specifically evaluating their performance on six mental health prediction tasks using social media data. The authors proposed a general zero-shot prompt template that can be used to predict a variety of mental health outcomes from online text data. The proposed zero-shot prompt template was as follows:

$$Prompt_{ZS} = TextData + Prompt_{Part1-S} + Prompt_{Part2-Q} + OutputConstraint$$

where "TextData is the online text data generated by end-users. $Prompt_{Part1-S}$ provides specifications for a mental health prediction target. $Prompt_{Part2-Q}$ poses the question for LLMs to answer, and OutputConstraint controls the output of models (e.g. Only return yes or no for a binary classification task)"

The authors proposed several design strategies for $Prompt_{Part1-S}$, including:

(i) Basic: Leaving it as blank.
(ii) Context Enhancement: Providing more social media context about the TextData.
(iii) Mental Health Enhancement: Inserting a mental health concept by asking the model to act as an expert.
(iv) Context and Mental Health Enhancement: Combining both enhancement strategies by asking the model to act as a mental health expert under the social media context.

For $Prompt_{Part2-Q}$, the authors focus on two categories of mental health prediction targets:

1. Predicting critical mental states: Such as stress or depression.
2. Predicting high-stake risk actions: Such as suicide.

The authors tailored the question description for each category and investigated binary and multiclass classification

tasks. They also performed minor modifications based on specific tasks to confirm suitable questions were answered. They also suggested employing few-shot prompting across various LLMs. This technique involves furnishing the LLM with a limited set of prompt-label examples relevant to the specific task. These examples are then compiled into a few-shot prompt, which serves to guide the LLM's response generation.

The design of the prompt is straightforward as:

$$Prompt_{FS} = [SamplePrompt_{ZS} - label]_M + Prompt_{ZS}$$

where M is the number of prompt-label pairs and is capped by the input length limit of a model.

The results indicate that the success of different strategies depended on the specific LLM and the task at hand. For example, including additional context or mental health-related information in the prompt appeared to lower the performance of task-focused models such as Alpaca-LoRA and FLAN-T5. However, both smaller and larger LLMs have shown encouraging performance in mental health tasks. Notably, FLAN-T5 and GPT-4 have demonstrated performance levels that are approaching those of task-specific natural language processing (NLP) models. In the realm of improving prompt design strategies for LLMs, it is evident that these approaches generally yield effective results for dialogue-focused models. However, the effectiveness of such strategies varies based on the model's primary focus, particularly between dialogue-focused and task-solving-focused models. In particular, these strategies tend to work better for critical action prediction tasks, such as suicide prediction. Interestingly, a few-shot prompting is more effective for the smaller language models Alpaca and FLAN-T5 than for the larger language models GPT-3.5 and GPT-4. This may be because smaller models are easier to quickly adapt to a new task by seeing just a few examples. Furthermore, the provision of additional contextual information about the task and input consistently leads to performance improvements across a wide range of cases. This emphasis on contextual information is a key factor in optimizing LLM performance in various applications within the mental health domain. The influence of model architecture and trainable parameters is also notable. Dialogue-focused models equipped with larger trainable parameters, exemplified by the comparison between Alpaca and Alpaca-LoRA, as well as GPT-3.5 and Alpaca, demonstrate a heightened ability to leverage contextual or domain-specific information contained within prompts. Interestingly, GPT-4 exhibits a less pronounced effect in response to different prompts, indicating variability in the impact of model architecture on prompt utilization within the mental health context.

Similarly, in a recent study by Wang et al. (2023a), an innovative approach known as the "Dialogue Chain-of-Thought (CoT) prompting" is proposed to explore the realms of personality, emotion, and psychology within user queries in the context of mental health. This approach involves the utilization of a benchmark that encompasses Chinese and English datasets, demonstrating the consistent improvement of response quality by breaking down the response generation process into distinct reasoning steps. Furthermore, the authors introduced a detailed framework that facilitates structured and transparent reasoning in dynamic conversations. They presented two distinct modes of implementing CoT Prompting:

1. Standard CoT:
   - The authors built on conventional prompting and Chain-of-Thought (CoT) techniques by proposing 'Explicit CoT,' a method designed to improve LLM reasoning. This approach broke down the reasoning process into N-1 successive phases, with each phase producing intermediate reasoning results. In the final stage (N), the model combined the initial dialogue context with all previous intermediate results to generate the final response.
2. Explicit CoT:
   - The "Explicit CoT" method aimed to improve the reasoning abilities of LLMs by breaking down response generation into multiple sequential steps. It divided the reasoning process into N-1 consecutive steps, with each step producing intermediate reasoning results. The Nth step then combined the original dialogue context with all intermediate outputs to generate the final response.

The "Explicit CoT" approach offered several advantages. Firstly, processing intermediate outputs provided convenience, allowing for actions such as generating personalized responses by incorporating user profiles or filtering out erroneous reasoning results. Furthermore, these intermediate outputs could be retained for future use, making them valuable for various purposes. These intermediate results can also serve as criteria for selecting demonstrations under few-shot settings, a concept elaborated on in the subsequent section. In summary, the Explicit CoT technique fosters a more transparent and systematic progression of reasoning within LLMs, leading to improvements in transparency and interpretability within the dialogue generation process. The authors experiment on the following Large Language Models: for Chinese: ChatGLM-6B, BELLE-LLAMA-7B-2M, ChatGPT; for English: Flan-T5-XXL, Alpaca-7B, Vicuna-7B-v1.1, ChatGPT.

Firstly, the models demonstrated the ability to generate empathetic and supportive responses tailored to the user's psychological state which holds great promise for enhancing the quality of mental health support. By taking into account the user's emotional and psychological condition, the model can contribute to a more effective and compassionate counseling experience. This capacity to provide personalized responses can be a valuable asset in improving the overall well-being of individuals seeking mental health assistance. Secondly, the model's performance on psychology datasets can be regarded as accurate for evaluating the efficacy of language models in generating counseling responses.

3.2 | Fine-tuning

Fine-tuning of LLMs deals with customizing pre-trained models to improve their performance on specific tasks. This process starts with a model that has already been trained on a large, diverse dataset and then refines it using a more targeted dataset relevant to the specific task. By adjusting the model's parameters during fine-tuning, it becomes better at recognizing the specific nuances of that domain, thereby improving its ability to understand and generate content that aligns with the intended application Wu et al. (2023).

Researchers used fine-tuned language models to better understand and respond to mental health questions, enabling it to grasp the nuances of mental health language. Liu et al. (2023b) created a method to make AI generate longer, more relevant answers. They do this by: i) building a vocabulary specific to mental health, ii) combining this vocabulary with the AI's existing knowledge, and iii) teaching the AI how to use this combined vocabulary. By fine-tuning, the model gets better at understanding the details and nuances of mental health language and context.

To create a vocabulary that captures the subtleties of mental health data, Liu et al. (2023b) used a method that allows words to be split more flexibly. This made sure the vocabulary was tailored to the specific needs of the task. After that, the newly created vocabulary was combined with the language model's existing one while keeping the token sequence unchanged. This approach enabled the model to retain what it had already learned while also adapting to the new task. To generate meaningful text with this updated vocabulary, new token representations were formed by connecting the new tokens to sub-word units from the original vocabulary and averaging their embeddings. The performance of this adaptive vocabulary approach was assessed on the CN PsyQA and MHP Reddit datasets, comparing it against conventional tokenizers. For the analysis, reduced versions of the GPT2 and BART models were utilized as the base architecture. For the tokenization process, the authors compare different tokenizers, including gpt2-chinese-cluecorpussmall, bart-base-chinese, and chinese-LLaMA-plus-7B for the CN PsyQA dataset, and GPT2 and bart-base for the MHP Reddit dataset. These tokenizers use different vocabulary sizes and tokenization methods, such as WordPiece and BPE-dropout. To evaluate the effectiveness of the task-adaptive tokenization approach, the authors conduct both automatic evaluation and human evaluation. The automatic evaluation includes metrics such as BLEU scores and RougeL. The human evaluation involves recruiting native professionals who assess the generated responses based on fluency, coherence, and professional expression. The study further examined how varying the size of the task-specific vocabulary affected the quality of a generated text. Generation efficiency was analyzed by measuring the average time taken in centiseconds and the number of tokens produced per generation, along with the response length and token utilization rate. The experimental results demonstrated that the task-adaptive tokeniza-

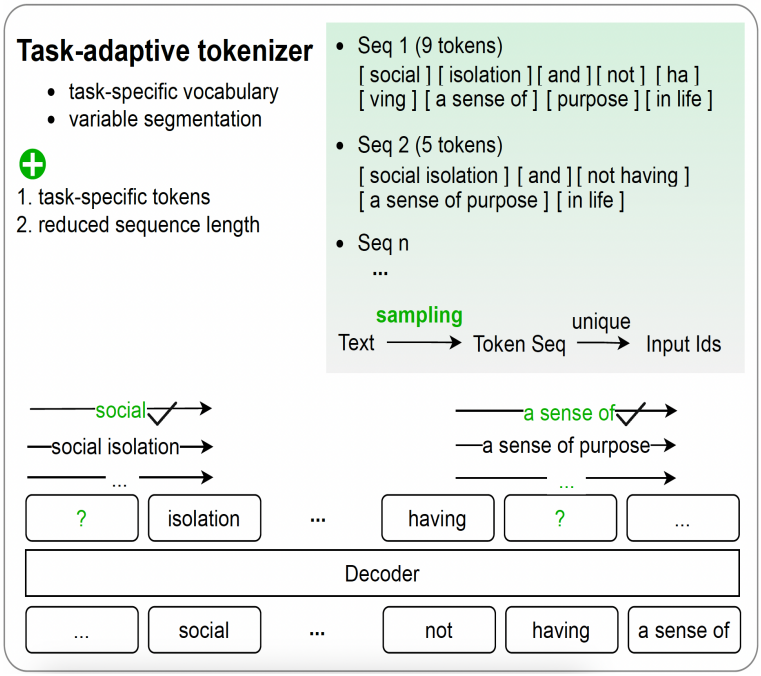


FIGURE 2 With task-adaptive tokenization, the same phrase appearing in different training instances may result in different tokenizations. For instance, Seq 1 and 2 illustrate two different token sequences of the same input using a task-adaptive tokenizer as per Liu et al.Liu et al. (2023b)

tion method enhanced the generative performance of the models when applied to mental health question-answering tasks.

Figure 2 uses an example to display the potential of task-adaptive tokenization. In terms of effectiveness, the task-adaptive tokenization approach achieves higher BLEU scores and RougeL scores compared to tokenizers from off-the-shelf pre-trained models. This indicates that the generated responses are more similar to the reference responses and have better overall quality. The human evaluation results also support the effectiveness of the task-adaptive tokenization approach.

Xu et. al Xu et al. (2023) also introduced a strategy that aligns more closely with traditional few-shot transfer learning, where the objective is to further train Large Language Models using a limited amount of domain-specific data, which bears similarity to prior research in the field of mental health Naveed et al. (2023). Our study investigates various fine-tuning techniques to enhance the performance of LLMs in mental healthcare.

- Single-dataset Fine-tuning: Consistent with previous studies in the mental health domain, they initially perform foundational fine-tuning on a single dataset, specifically the training set. This process involves adapting the LLM to the nuances of the specific dataset. Subsequently, they evaluate the fine-tuned model on both the same dataset (referred to as the test set) to assess its performance and on different datasets to determine its generalizability.
- Multi-dataset Fine-tuning: This investigation delves into the process of fine-tuning LLMs across multiple datasets simultaneously. This is accomplished by employing instruction fine-tuning, which empowers Large Language Models to handle a variety of tasks originating from different datasets. The previous models were fine-tuned for specific tasks, such as depression prediction or suicidal ideation prediction, and they became task-specific after training. In contrast, this approach fine-tunes LLMs across multiple mental health datasets with diverse instructions for different tasks within a single round of fine-tuning, equipping them to address multiple tasks

without requiring additional task-specific fine-tuning.

The analysis demonstrates that instruction-based fine-tuning enhances large language model performance in mental health prediction applications. The customized versions, Mental-Alpaca and Mental-FLAN-T5, demonstrate superior results compared to their base models, GPT variants, and the specialized MentalRoBERTa across most evaluation metrics. While both Mental-Alpaca and Mental-FLAN-T5 focus on dialogue interactions, they exhibit distinct advantages: Mental-Alpaca excels in processing natural human language, whereas Mental-FLAN-T5 shows stronger task-specific performance. Organizations with limited resources might benefit from task-oriented LLMs, while those with adequate data and computational capacity could achieve better outcomes through fine-tuning dialogue models. Models with dialogue capabilities, such as Alpaca, are particularly well-suited to mental health support applications. These findings indicate a significant potential for the creation of effective mental health assessment systems based on big language models, which could improve mental healthcare accessibility.

In a similar innovative approach, Zheng et al. (2023) leveraged the generative power of Large Language Models (LLMs) to create a vast dataset called ExTES (Extensible Emotional Support Dialogue dataset). This dataset is specifically designed to develop robust and adaptable emotional support dialogues. ExTES was developed meticulously, using ideas from psychology, past research on emotional support, and even components from creative literature to enrich and diversify the interactions. The researchers identified 36 common scenarios and arranged the dataset with psychological research as the foundation.

3.3 | Emotional & Linguistic Context

Emotions play an important role in our human experience, frequently serving as signs of problems we may be facing. By recognizing and dealing with these emotions, mental health counselors can assist clients in processing and making sense of their feelings, which can be extremely helpful. Ratican and Hutson (2023) introduced the 6DE model, a multidimensional approach to analyzing emotions. The 6DE model offers a more nuanced understanding of emotional states within text-based communication, enabling AI systems to generate appropriate emotional responses and enhance user experiences Sharma et al. (2023a). The model has six dimensions: arousal, valence, dominance, agency, fidelity, and novelty; for understanding and quantifying emotional states. By incorporating sub-dimensions within each dimension, this approach can potentially enhance the comprehension and generation of appropriate emotional responses by AI systems, leading to more effective human-AI interactions.

Garg (2024) formulated the task of determining wellness dimensions in Reddit posts as a multivariate concept extraction and classification problem. By considering multiple wellness dimensions simultaneously, the proposed approach enables a more comprehensive understanding of an individual's mental health state and the underlying causes and consequences of mental disturbance as depicted in Figure 3. Recent literature presents six different dimensions of wellness because of mental and social well-being, namely, (i) physical aspect, (ii) intellectual aspect, (iii) vocational aspect, (iv) social aspect, (v) emotional aspect, and (vi) spiritual aspect Strout and Howard (2012).

Social media content presents unique challenges for wellness dimension identification due to linguistic ambiguity and informal communication styles. The unstructured nature of social media posts, including colloquialisms, abbreviations, and non-standard expressions, creates obstacles for computational analysis. The inherent subjectivity and contextual dependence of wellness-related topics further complicate accurate identification.

To overcome these challenges, researchers suggest refining pre-trained transformer models by fine-tuning them with specialized datasets like WELLXPLAIN, which helps in better understanding the way people communicate on social media Sharma et al. (2023a). Another key approach is to develop models specifically designed to handle informal language and interpret contextual details more effectively, leading to improved accuracy. The Concept Extraction Framework (CEF) has been introduced as an automated method to identify and classify wellness-related ideas in text, making it easier to analyze mental health discussions in greater depth. LLMs, including BERT, RoBERTa, ALBERT, DeBERTa, PsychBERT, MentalBERT, and ClinicalBERT, are employed within the CEF to process and analyze text, iden-

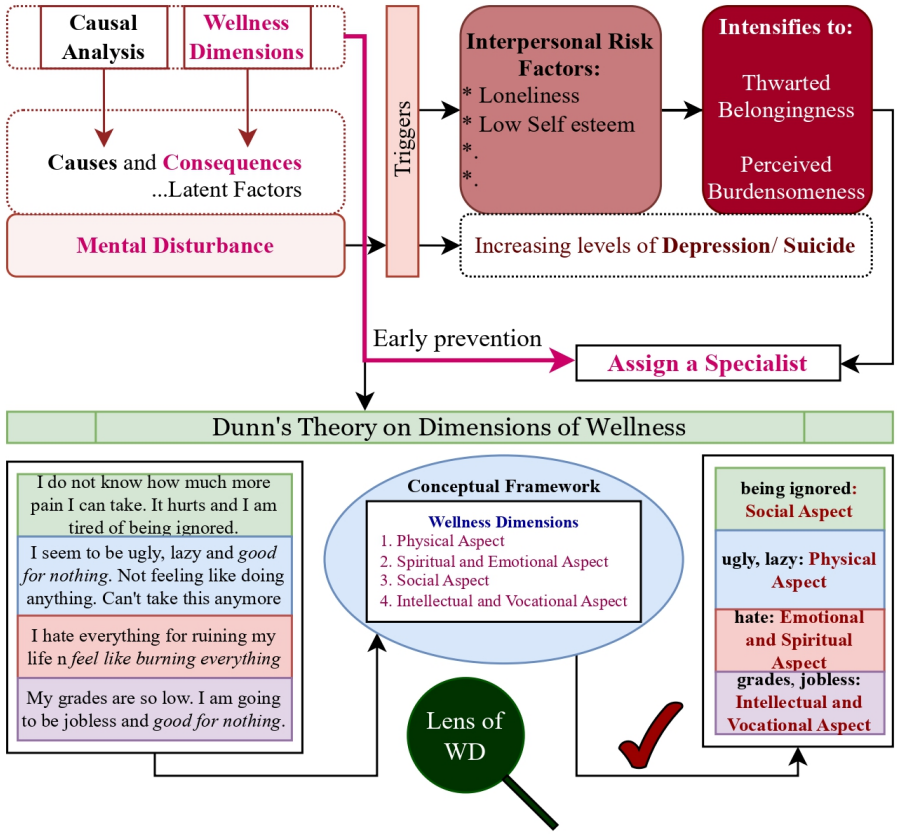


FIGURE 3 An illustration of the utility of WellXplain. The wellness dimensions in their work are grounded in Dunn's theory of wellness Garg (2024).

tifying relevant wellness dimensions and their contextual embeddings. These models, pre-trained on vast amounts of text data and fine-tuned for specific tasks, enable effective linguistic comprehension for concept extraction.

3.4 | Pre-training

Pre-training is crucial for improving LLM performance. It allows these models to get a complete understanding of language, which is particularly useful in mental health. Pre-training allows LLMs to better understand and respond to individual problems with greater accuracy and empathy. Ma et al. (2023) introduced the Psy-LLM framework that uses LLMs to answer questions in online psychological consultations. This framework aims to improve mental health support by combining pre-trained LLMs with professional QA datasets and selected psychological content. Its performance is evaluated using both intrinsic and extrinsic metrics to check how coherent and relevant its responses are. The Psy-LLM framework is built on two pre-trained models: the PanGu model and the WenZhong model.

The PanGu model is a language model with around 200 billion parameters that uses vanilla transformer layers Zeng et al. (2021). It is the first Chinese pre-training model, trained on a huge amount of unlabeled data. PanGu provides better language representations, improved initialization parameters, and effective regularization, which makes it ideal for use in the Psy-LLM framework for question-answering Zeng et al. (2021). On the other hand, the WenZhong model is built on a large pre-trained model trained using a Chinese corpus Zhang et al. (2023). Its architecture focuses on a decoder-based design. The model implemented with the WenZhong framework has 110 million parameters spread

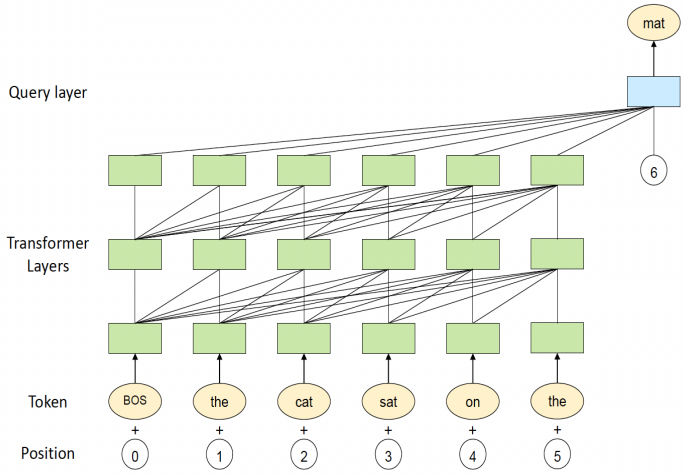


FIGURE 4 The model layers and architecture of the PanGu model as proposed in Zeng et al. (2021)

over 12 layers Zhang et al. (2023). It also benefits from being pre-trained on the Wudao Corpus, which enhances its ability to understand and generate language. This framework used two main datasets: the PsyQA dataset, which includes 22,000 mental health-related question-answer pairs and 56,000 detailed responses, and a custom dataset compiled from Chinese social media platforms such as Tianya, Zhihu, and Yixinli.

The model's performance is assessed through intrinsic and extrinsic evaluation methods. For intrinsic evaluation, metrics like perplexity, Rouge-L, distinct-1, and distinct-2 are used Zeng et al. (2021). Perplexity checks how well the model predicts the next word in a sequence, while Rouge-L measures the similarity between the generated response and the reference answer. Distinct-1 and distinct-2 focus on the diversity of the generated responses by looking at unique unigrams and bigrams. Extrinsic evaluation, on the other hand, looks at how effectively the model provides meaningful and relevant responses to users seeking psychological support. Additionally, human evaluation is carried out to gather deeper insights into the model's overall performance.

$$\text{Perplexity} = 2^{-\frac{1}{N}\sum_{i=1}^{N}\log_2 P(w_i)} \tag{1}$$

$$ROUGE\text{-}L = \frac{\text{LCS}(\text{machine-generated text, reference text})}{\text{Total number of words in reference text}} \tag{2}$$

$$\text{Distinct-1} = \frac{\text{Number of Unique Unigrams}}{\text{Total Number of Unigrams}} \tag{3}$$

$$\text{Distinct-2} = \frac{\text{Number of Unique Bigrams}}{\text{Total Number of Bigrams}} \tag{4}$$

The evaluation results show that the PanGu model performs better than the WenZhong model when it comes

to helpfulness, fluency, relevance, and logical coherence. Both intrinsic and human evaluation metrics confirm this Ma et al. (2023). Human evaluators generally find PanGu's responses to be more useful, fluent, relevant, and logically sound compared to those generated by WenZhong. One major drawback of the WenZhong model is its vulnerability to exposure bias during training Zhang et al. (2023). Exposure bias appears when samples come from the target language distribution, which leads to errors that still lack effective solutions. Moreover, repeated training of the WenZhong model seems to result in a noticeable drop in output quality.

In their study, Prystawski et. al. Prystawski et al. (2023) tested whether introducing rationales that reason about the latent variables posited by probabilistic models of language understanding could improve the capacity of large language models to choose appropriate paraphrases for metaphor. Prompts were designed with an instructional sentence followed by 10 examples. In addition, metaphors can be used in psycho-education to explain complex psychological concepts to individuals who may have limited knowledge or understanding of mental health. For example, the metaphor of a "broken record" can be used to explain rumination or repetitive negative thinking patterns. This metaphor helps individuals grasp the concept and recognize the need to break free from unhelpful thought patterns **?**. The LLMs are prompted to infer the QUD and reason about the literal state of the world to choose an appropriate paraphrase for the metaphor. Similarity prompts involve identifying the subject and object of the metaphor and reasoning about the property transferred from the object to the subject. The LLMs use this information to select an appropriate paraphrase for the metaphor.

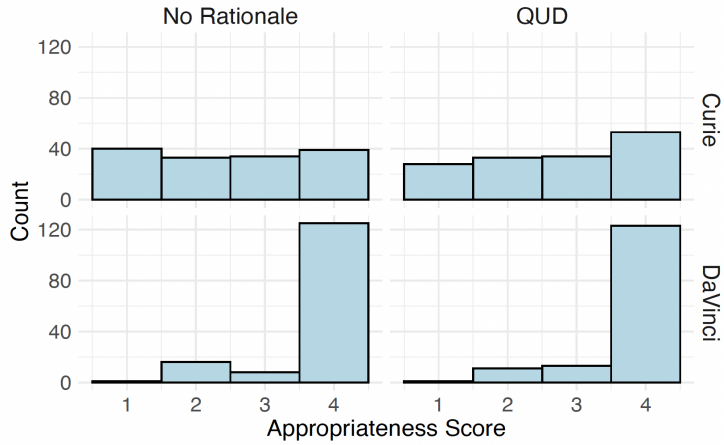


FIGURE 5 Histograms of paragraph appropriateness distributions for GPT-3 Curie (top) and DaVinci (bottom) with both no rationales (left) and QUD rationales (right) evaluated by Prystawski et al. (2023)

The dataset used for evaluation is the Katz corpus, which consists of both non-literary and literary metaphors in English. The metaphors in the corpus are of the form "[subject] is [object]" or "[subjects] are [object(s)]." The data set includes norms related to aspects such as comprehensibility, metaphor goodness, and familiarity for each metaphor, based on human ratings. Four non-metaphorical paraphrases were created for each metaphor to evaluate appropriateness. The best interpretations transferred a relevant property from the object to the subject, while the second-best transferred a less appropriate property. The third-best expression expressed a fact about the subject and the worst paraphrase expressed the opposite interpretation of the best. The main metric used for evaluating model responses was the appropriateness score of the chosen paraphrases (ranging from 1 to 4). Bayesian cumulative logistic regression and Pearson correlations were used for analysis. Ratings of psycholinguistic features of the metaphors were considered in the evaluation.

In the upper section of Figure 5, the appropriateness distributions for GPT-3 Curie are presented, comparing scenarios with and without QUD prompts. The incorporation of QUD prompts is observed to shift the appropriateness

distribution, moving it away from a rating of 1 and toward 4. Moving to the lower section of Figure 5, the complete appropriateness distributions for GPT-3 DaVinci are depicted under two conditions: without rationale prompts and with QUD rationale prompts. In both cases, the distributions predominantly favor a rating of 4, yet when QUD prompts are employed, there is a slight increase in responses rated 3 compared to those rated 2. However, it is important to note that the study primarily focuses on paraphrasing metaphors and does not extensively explore other aspects of metaphor comprehension, such as understanding the figurative meaning or interpreting the intended message. Additionally, the evaluation metrics used, such as appropriateness scores, may not capture the full complexity of metaphor understanding. Furthermore, while the study provides insights into the effectiveness of rationales in improving LLM performance, it does not thoroughly investigate the underlying mechanisms or provide a comprehensive analysis of the limitations of the approach. Future research could delve deeper into the specific cognitive processes involved in metaphor comprehension and explore alternative methods to enhance LLMs' metaphor understanding capabilities.

3.5 | Chatbots

Another approach is using chatbots in mental healthcare by implementing artificial intelligence-driven chat-based systems to provide support, guidance, and interventions for individuals facing mental health challenges. These chatbots engage with users via interactive chats to provide mental health support. This study focuses on improving the performance of this framework in the evaluation of dialogue models involved in mental support conversations. The dialogue model serves as a filter in order to decide whether to selectively utilize or ignore the expert's advice based on the context and dialogue history. Dialogue models trained include:

- Vanilla BlenderBot (2.7B): baseline BlenderBot model fine-tuned on EmpatheticDialogues, ConvAI, WizardofWiki, and BlendedSkillTalks in a multi-task style.
- BlenderBot for Mental Health: It was created by fine-tuning the BlenderBot (BB) on the original ESConv dataset. Furthermore, it integrates multi-task fine-tuning approach using both BlendedSkillTalks and ESConv with equal training weight.
- BlenderBot for Mental Health with Reasoning: This model adopts the Ask an Expert approach for mental health support systems and further, based on reasoning processes which are majorly obtained through prompts, it is fine-tuned.

Evaluation results depict that the "Ask an Expert" approach has improved the performance of dialogue models in mental support conversation. In the automatic evaluation of reasoning processes, there is a clear improvement in similarity and entailment scores from GPT-1 to DaVinci, with DaVinci outperforming other models. In the human evaluation of reasoning processes, the annotations containing reasoning information provided by the expert models are found to be valid and informative. In the evaluation of dialogue models, the models with reasoning processes outperform the baseline model in terms of providing constructive suggestions, sharing similar experiences, and being engaging and empathetic. The human evaluation results also show that davinci performed better than other models in generating high-quality dialogues.

4 | AVAILABLE DATASETS

Most of the datasets that are apt for LLMs are collected from social media posts based on the nature of the posts. Dealing with unstructured and semi-structured data from social media is one of the trickiest tasks. It involves managing missing values and cohesively presenting multi-modal information. Details of some of the widely used datasets and experimental setups are given in Table 2. The available datasets are discussed in this section.

- *CLPsych 2015 Shared task dataset*: The CLPsych dataset Coppersmith et al. (2015) consists of three modules, Depression vs. Control (DvC), PTSD vs. Control (PvC), and Depression vs. PTSD (DvP) (Post-traumatic stress disorder or PTSD, is an anxiety disorder that develops in reaction to physical injury or severe mental or emotional distress, such as military combat, violent assault, natural disaster, or other life-threatening events).
- *Reddit Self-reported Depression Diagnosis (RSDD) dataset* Yates et al. (2017): The RSDD dataset comprises Reddit posts from around 9,000 users who have self-reported a diagnosis of depression, referred to as "diagnosed users." Additionally, the dataset includes approximately 107,000 matched control users.

- *Multimodal Dictionary Learning (MDDL) dataset* Shen et al. (2017): The MDDL dataset for depression detection consists of three modules. The modules are created using tweets from 2009 to 2016, where users are labeled as depressed if their anchor tweets meet the strict pattern "(I'm/Iwas/Iam/I'vebeen) diagnosed depression."

- *Self-Reported Mental Health Diagnoses(SMHD) dataset* Cohan et al. (2018): SMHD dataset comprises Reddit posts from users who have self-reported being diagnosed with one or several of nine mental health conditions, referred to as "diagnosed users," along with matched control users. Several studies in the literature have utilized this dataset, and it focuses on multiple mental health conditions rather than solely the depression dataset.

- *eRISK* Martın-Rodilla et al. (2022): This dataset has been accessible online for experiments and analysis for a few years to support the goals of a shared task. This dataset, dedicated to early risk detection, has been provided by CLEF Lab to address the challenges associated with identifying depression, anorexia, and self-harm for a considerable period.

- *Sina Weibo* Cao et al. (2019): This dataset is extracted from the Chinese social media platform Sina Weibo comprises 3652 users exhibiting suicidal tendencies and 3677 users without any suicidal risk.

- *Suicide Risk Assessment using Reddit(SRAR)* Gaur et al. (2019): The SRAR dataset is openly accessible in the public domain and consists of 500 Redditors (anonymized), their posts, and expert-annotated labels. The research community anticipates utilizing this dataset soon to improve and advance proposed techniques.

- *Dreaddit* Turcan and McKeown (2019): Dreaddit introduces a novel text corpus designed for identifying stress in lengthy, multi-domain social media data. This dataset comprises 190,000 posts gathered from five different categories of Reddit communities.

- *GoEmotion* Demszky et al. (2020) : The GoEmotion dataset includes 58,000 meticulously curated comments sourced from Reddit, each human-annotated with one of 27 emotion categories or labeled as Neutral. It also features a filtered version derived from an inter-rater agreement, comprising a train/test/validation split.

- *The University of Maryland Reddit Suicidality Dataset (UMD-RD)* Shing et al. (2020) : The UMD-Reddit Dataset consists of two sub-directories, one containing data from 11,129 users who posted on SuicideWatch and another for 11,129 users who did not. Each user in the dataset has complete longitudinal data from the 2015 Full Reddit Submission Corpus. Academic researchers have actively utilized the UMD-Reddit dataset since 2019, accessing it through a signed agreement.

- *SDCNL dataset*: The SDCNL dataset was obtained through the Reddit API and scraped from two subreddits, namely r/SuicideWatch and r/Depression, comprising a total of 1,895 posts. Two fields were extracted from the scraped data: the original text of the post as inputs and the subreddit to which it belongs as labels. Posts originating from r/SuicideWatch are labeled as suicidal, while those from r/Depression are labeled as depressed.

- *SWMH dataset* Ji et al. (2022b): The dataset used in this study is available upon request. It comprises 5,326 suicidal posts out of 20,000 extracted posts and 594 suicidal tweets out of 10,000 extracted tweets, serving for experiments and the evaluation of the proposed classification approach for suicidal risk detection.

- *PsyQA dataset* Sun et al. (2021a): PsyQA is a Chinese dataset designed for psychological Question-Answer (QA) support. It features answers that have been provided by well-trained volunteers and professional counselors. This dataset encompasses a total of 22,346 questions and 56,063 answers, with each question being associated with one or more ground-truth responses.

5 | ETHICAL CONSIDERATIONS AND CHALLENGES

Advances in language models have led to more effective AI-driven mental health support tools. Due to its transformative capabilities, it has ensured the wide availability of mental health support for those in need. However, there is a need of a comprehensive evaluation of its ethical implications and addressing the challenges that accompany this. This section uncovers the underlying issues which generally occurs in the deployment of LLMs in the mental health field. In its initial phase, the conventional methods generally comprised of text-based services or face-to-face interactions. Hence, certain issues like long waiting times and limited access were prime concerns. With the advent of LLMs, these issues got resolved easily since it provided quick, affordable, and scalable mental health support.

In the majority of cases, to ensure fairness, it becomes necessary to address the biases that AI can generally learn from input data Bender and Friedman (2018). Thus, it becomes an issue that falls under ethical consideration. Moreover, it also raises concerns regarding professional competence and accountability.

Still, AI has limited capability While it comes to comprehending and conveying human emotions. LLMs may produce unethical and improper information, such as racist and ageist aspects, as well as other forms of bias, such as suicidal responses because they learn from uncensored online sources Bender et al. (2021). Therefore, it raises fundamental doubts about the effectiveness and fairness of LLMs. Dealing with sensitive topics like gender, religion, or other divisive concerns, the models may unwittingly feed preconceptions or come across as hostile in their responses. Hence, it is essential to carefully design and monitor AI systems to handle delicate subjects with sensitivity and fairness Ferrara (2023).

6 | LIMITATIONS AND OPPORTUNITIES

6.1 | Limitations

LLMs trained on extensive datasets carry inherent biases and errors, rendering them susceptible to generating inaccurate and misleading information. This limitation is particularly salient when addressing intricate subjects like mental health Chow et al. (2023); Holzinger et al. (2023); Rao et al. (2023); Wang et al. (2023c). Furthermore, the capability of LLMs to grasp the subtleties of human language lags behind that of humans, giving rise to potential misunderstandings and misinterpretations, which can prove detrimental in clinical contexts. Despite of their transformative capabilities, LLMs are limited by their lack of in-depth human linguistic understanding which adversely impact medical decision-making. While still in its early stages, existing studies have identified several key gaps. A key problem is the significant need to create more robust and dependable LLM systems, as current models suffer from uneven performance Wang et al. (2023c). The lack of established methods makes evaluating LLMs for mental healthcare particularly difficult. The biggest challenge in evaluating LLMs for mental health is the absence of standard methods Wang et al. (2023c). Without suitable evaluation methodologies, it is difficult to determine the effectiveness of these models. A clear, uni-

TABLE 2 Different mental health datasets and their details.

Reference	Dataset	Details	Nature
Muskan Garg Garg (2024)	WellXplain	Wellness concept extraction in Reddit posts	Custom
Glen et al. Coppersmith et al. (2015)	CLPsych	Data from Twitter users who state a diagnosis along with demographically-matched community controls	Depression, Control, and PTSD
Andrew et al. Yates et al. (2017)	RSDD	Reddit posts of users who have claimed to have been diagnosed with depression; 9210 users in depression and 1,07,274 users in control group	Depression, Control
Shen et al. Shen et al. (2017)	MDDL	300 million users and 10 billion tweets in Depression, Non-depression, and Depression candidate	Depression
Cohan et al. Cohan et al. (2018)	SMHD	Reddit dataset for multi-task mental health illness	Mental Health
Parapar et al. Martın-Rodilla et al. (2022)	eRISK	Early risk detection by CLEF lab about problems of detecting depression, anorexia and self-harm	Depression
Cao et al. Cao et al. (2019)	Sina Weibo	3,652 (3,677) users with (without) suicide risk from Sina Weibo	Suicidal Id.
Gaur et al. Gaur et al. (2019)	SRAR	Posts from 500 Redditors (anonymized) & annotated by domain expert	Suicide Risk
Turcan et al. Turcan and McKeown (2019)	Dreaddit	190K Reddit posts of 5 different categories	Stress
Demszky et al. Demszky et al. (2020)	GoEmotion	Manually annotated 58k Reddit comments for 27 emotion categories	Emotion
Shing et al. Shing et al. (2020)	UMD-RD	11,129 users who posted on r/SuicideWatch and 11,129 users who did not	Suicidal Id.
Haque et al. Haque et al. (2021)	SDCNL	Reddit dataset of 1895 posts of depression and suicide	Suicidal and Depression
Garg et al. Garg et al. (2022)	CAMS	Interpretable causal analysis of mental illness in social media (Reddit) posts	Mental Health
Zheng et al. Zheng et al. (2023)	ExTES	36 emotionally impactful scenarios that encompass a wide array of everyday life situations	Emotion
Ji et al. Ji et al. (2022b)	SWMH	Collection contains a total of 54,412 posts comprising suicide-related intention and mental disorders	Suicidal Id.
Sun et al. Sun et al. (2021a)	PsyQA	Chinese dataset of psychological health support in the form of question (22K) and answer (56K) pair	Mental Health

form method for evaluating LLMs would make it easy to grasp their strengths as well as their weaknesses. However, existing evaluation procedures remain restricted, and more accurate ways must be developed to ensure that LLMs are actually suited for clinical usage Campion et al. (2010).

Integration of LLMs into existing mental healthcare systems is a dimension that needs in-depth exploration. More research is required to understand this for seamless integration Wang et al. (2023c). The current literature focuses primarily on demonstrating LLM capabilities but lags in evaluating the associated risks and harms. Risk management should also be included, followed by techniques for content filtering, output verification, and safety constraints. Additionally, there hasn't been enough testing with actual people who need mental health support.. Studies should prioritize real-world interactions to assess the practical effectiveness of LLMs Li et al. (2023); Campion et al. (2010). Not enough studies have been conducted comparing LLMs to human mental health professionals. It's crucial to thoroughly assess the performance of LLMs in detection, counseling, and treatment Li et al. (2023). The usage of data of public online forums, which may not generalize well to clinical settings, is a critical limitation. The requirement of developing diverse and representative datasets to enhance the reliability of LLMs is also important Chhikara et al. (2023); Li et al. (2023); Xu et al. (2023). The potential of LLMs in explainable mental health analysis needs improvement. In order to promote better evaluation, the research needs to focus on enhancing transparency in LLM reasoning Ji et al. (2021). Efforts should be made to handle LLM content generation so that unethical outputs can be avoided. This involves researching techniques for content filtering, safety constraints, and output verification. The potential for overreliance on and addiction issues stemming from prolonged LLM use in mental health support requires further exploration. The mitigation strategies are also an essential domain in order to ensure responsible LLM utilization. Anthropomorphizing LLMs for mental health chatbots can lead to people becoming overly attached, which may pose risks. This warrants investigation to understand potential consequences. Evaluating LLM mental health detection on more diverse datasets, including multilingual assessments, is imperative. This expands the applicability of LLMs and improves their robustness. In conclusion, while LLMs offer promise in enhancing mental healthcare, addressing these research gaps and challenges is pivotal to ensure their safe and effective integration into clinical practice. Continued research efforts in these areas will ultimately contribute to better mental health outcomes for all.

6.2 | Opportunities

The applications of LLMs in the context of mental health encompass several key techniques which still have a lot of room for improvement. We identify a few areas where there is an opportunity to improve our LLMs.These are given below:

Real-world testing and evaluation: One fundamental aspect of LLM integration into mental healthcare involves conducting clinical trials or field studies to evaluate the effectiveness of LLM mental health chatbots when interacting with patients in real healthcare settings. Gathering feedback from both patients and clinicians is crucial for assessing the strengths and weaknesses of LLM systems. Furthermore, testing LLM detection systems retrospectively on clinical records is essential for determining the real-world accuracy of these systems.

Bias and Fairness: The potential biases in existing mental health LLM systems and data are areas of concern. Investigating these biases and developing techniques to mitigate them is critical to promote fairness in LLM training data and outputs.

Safety and Ethics: Ensuring the safety and ethical use of LLMs in mental healthcare is imperative. Tools to automatically detect and filter harmful LLM-generated text are under development, and interventions are being designed to prevent overreliance and addiction when using LLM chatbots for mental health support Ji et al. (2021).

Model Diversity: Diversifying LLM mental health systems by testing them on various cultural texts and languages is

being pursued. Adaptations to LLM architectures and training methods are being made to enhance performance on non-English data.

Clinical Efficacy: Conducting studies to determine the outcomes of LLM interventions in comparison to traditional treatment is crucial. LLMs are also being evaluated as assistive tools for clinicians to improve mental health workflows and processes.

User Assessment: Developing strategies to allow users of mental healthcare-oriented generative AI solutions to assess the performance of responses based on performance metrics such as accuracy, precision, recall, F1-score, and user satisfaction scores to help evaluate the AI's performance in specific tasks. Also, customized user feedback using a long-term memory approach can help to establish a continuous monitoring system to track the AI's performance and iteratively improve its algorithms based on user feedback and new data.

Furthermore, several specific research directions and considerations have been identified:

- Studying algorithmic biases in mental health LLMs and developing mitigation strategies Harrer (2023).
- Enhancing the interpretability of LLM-generated mental health assessments and recommendations Ji et al. (2021).
- Exploring the multimodal capabilities of LLMs, including their integration with physiological signals for mental health monitoring and companionship Wang et al. (2023c)
- Examining the integration of LLMs with professional mental health workflows Wang et al. (2023c); Srivastava et al. (2022).
- Using LLMs to augment existing therapies, thereby providing patients with personalized feedback and support between therapy sessions Sharma et al. (2021).

7 | CONCLUSION

The integration of Large Language Models (LLMs) into mental healthcare presents transformative opportunities but also introduces complex challenges that must be carefully addressed. This paper has provided a comprehensive examination of the diverse techniques employed to harness LLMs for mental health applications, shedding light on their capabilities, limitations, and ethical implications. By introducing a structured taxonomy, we conceptualized the evolving research landscape, analyzing key methodologies such as prompt engineering, fine-tuning, emotional theory integration, AI-driven chatbots, and long-term memory mechanisms. Case studies on depression detection, suicide risk prediction, and AI-assisted psychological counseling demonstrated LLMs' potential in risk assessment, emotional support, and personalized interventions. Techniques like instruction fine-tuning and domain-specific dataset curation significantly enhanced model performance, while psychologically informed prompts improved response accuracy. Conversational AI agents showed promise in fostering intuitive and engaging interactions, paving the way for AI-assisted mental health interventions. However, critical challenges remain. Bias, reasoning failures, safety risks, and inconsistencies in clinical reliability indicate that LLMs are not yet suitable for high-stakes mental health applications without rigorous oversight. The lack of real-world clinical validation, reliance on publicly available data, and long-term impacts of AI-driven mental health tools require urgent attention. Ethical concerns surrounding privacy, informed consent, and algorithmic bias must be addressed to ensure that AI augments rather than compromises mental healthcare.

This study underscores both the promise and the pitfalls of integrating Artificial Intelligence within Healthcare 5.0. To harness LLMs responsibly, developers, practitioners, ethicists, and policymakers must collaborate to establish ethical guidelines, safety mechanisms, and clinically validated AI frameworks. With prudent design, transparency, and ongoing evaluation, LLMs have the potential to enhance accessible, inclusive, and empathetic mental healthcare, shaping a future where AI serves as a responsible ally in supporting mental well-being.

DATA AVAILABILITY STATEMENTS

The data underlying this article will be shared on a reasonable request to the corresponding author.

references

AI, D. (2025) Deepseek-r1: A 671b parameter mixture-of-experts model. Available at `https://docsbot.ai/models/compare/deepseek-r1/llama-3-70b-instruct`.

AI, M. (2024) Llama 3.3 70b instruct: Multilingual instruction-tuned model. Available at `https://docsbot.ai/models/compare/deepseek-r1/llama-3-3-70b-instruct`.

Alpaca-LoRA (2023) Instruct-tune llama on consumer hardware. *URL https://github.com/tloen/alpaca-lora*.

Bender, E. M. and Friedman, B. (2018) Data statements for natural language processing: Toward mitigating system bias and enabling better science. *Transactions of the Association for Computational Linguistics*, **6**, 587–604. URL: `https://aclanthology.org/Q18-1041`.

Bender, E. M., Gebru, T., McMillan-Major, A. and Shmitchell, S. (2021) On the dangers of stochastic parrots. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. New York, NY, USA: ACM.

Benfares, C., Akhrif, O., El Idrissi, Y. E. B. and Hamid, K. (2021) A clinical support system for classification and prediction of depression using machine learning methods. *Computational Intelligence*, **37**, 1619–1632. URL: `https://onlinelibrary.wiley.com/doi/abs/10.1111/coin.12377`.

Bhugra, D., Till, A. and Sartorius, N. (2013) What is mental health? *International Journal of Social Psychiatry*, **59**, 3 – 4. URL: `https://api.semanticscholar.org/CorpusID:206413096`.

Black, S., Biderman, S., Hallahan, E., Anthony, Q., Gao, L., Golding, L., He, H., Leahy, C., McDonell, K., Phang, J., Pieler, M., Prashanth, U. S., Purohit, S., Reynolds, L., Tow, J., Wang, B. and Weinbach, S. (2022) GPT-NeoX-20B: An open-source autoregressive language model. In *Proceedings of BigScience Episode #5 – Workshop on Challenges & Perspectives in Creating Large Language Models*, 95–136. virtual+Dublin: Association for Computational Linguistics. URL: `https://aclanthology.org/2022.bigscience-1.9`.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A. et al. (2020) Language models are few-shot learners. *Advances in neural information processing systems*, **33**, 1877–1901.

Campion, E. W., Scott, L., Bowab, C., Fleming, S., Blizzard, T. J., Lamb, C., Easley, T. J. and Drazen, J. M. (2010) Benefits, limits, and risks of gpt-4 as an ai chatbot for medicine. *N. Engl. J. Med.*, **363**, 677–678.

Cao, L., Zhang, H., Feng, L., Wei, Z., Wang, X., Li, N. and He, X. (2019) Latent suicide risk detection on microblog via suicide-oriented word embeddings and layered attention.

Chhikara, P., Pasupulety, U., Marshall, J., Chaurasia, D. and Kumari, S. (2023) Privacy aware question-answering system for online mental health risk assessment.

Chiang, W.-L., Li, Z., Lin, Z., Sheng, Y., Wu, Z., Zhang, H., Zheng, L., Zhuang, S., Zhuang, Y., Gonzalez, J. E., Stoica, I. and Xing, E. P. (2023) Vicuna: An open-source chatbot impressing gpt-4 with 90%* chatgpt quality. URL: `https://lmsys.org/blog/2023-03-30-vicuna/`.

Chow, J. C. L., Sanders, L. and Li, K. (2023) Impact of ChatGPT on medical chatbots as a disruptive technology. *Front. Artif. Intell.*, **6**, 1166014.

Chung, H. W., Hou, L., Longpre, S., Zoph, B., Tay, Y., Fedus, W., Li, Y., Wang, X., Dehghani, M., Brahma, S., Webson, A., Gu, S. S., Dai, Z., Suzgun, M., Chen, X., Chowdhery, A., Castro-Ros, A., Pellat, M., Robinson, K., Valter, D., Narang, S., Mishra, G., Yu, A., Zhao, V., Huang, Y., Dai, A., Yu, H., Petrov, S., Chi, E. H., Dean, J., Devlin, J., Roberts, A., Zhou, D., Le, Q. V. and Wei, J. (2022a) Scaling instruction-finetuned language models.

— (2022b) Scaling instruction-finetuned language models.

Cohan, A., Desmet, B., Yates, A., Soldaini, L., MacAvaney, S. and Goharian, N. (2018) Smhd: A large-scale resource for exploring online language usage for multiple mental health conditions.

Cohere (2024) Command r+: Open-source llm for real-world challenges. Available at https://console.groq.com/docs/model/deepseek-r1-distill-llama-70b.

Coppersmith, G., Dredze, M., Harman, C., Hollingshead, K. and Mitchell, M. (2015) CLPsych 2015 shared task: Depression and PTSD on Twitter. In *Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, 31–39. Denver, Colorado: Association for Computational Linguistics. URL: https://aclanthology.org/W15-1204.

DeepMind, G. (2024) Gemini 2.0: Multimodal ai for the agentic era. Available at https://github.com/cheahjs/free-llm-api-resources/blob/main/README.md.

Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G. and Ravi, S. (2020) GoEmotions: A dataset of fine-grained emotions. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 4040–4054. Online: Association for Computational Linguistics. URL: https://aclanthology.org/2020.acl-main.372.

Fan, L., Li, L., Ma, Z., Lee, S., Yu, H. and Hemphill, L. (2023) A bibliometric review of large language models research from 2017 to 2023. *ArXiv*, **abs/2304.02020**. URL: https://api.semanticscholar.org/CorpusID:257952516.

Ferrara, E. (2023) Should chatgpt be biased? challenges and risks of bias in large language models.

Frueh, B. (2015) Solving mental healthcare access problems in the twenty-first century. *Australian Psychologist*, **50**.

Garg, M. (2024) Wellxplain: Wellness concept extraction and classification in reddit posts for mental health analysis. *Knowledge-Based Systems*, **284**, 111228.

Garg, M., Saxena, C., Saha, S., Krishnan, V., Joshi, R. and Mago, V. (2022) CAMS: An annotated corpus for causal analysis of mental health issues in social media posts. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, 6387–6396. Marseille, France: European Language Resources Association. URL: https://aclanthology.org/2022.lrec-1.686.

Gaur, M., Alambo, A., Sain, J. P., Kursuncu, U., Thirunarayan, K., Kavuluru, R., Sheth, A., Welton, R. and Pathak, J. (2019) Knowledge-aware assessment of severity of suicide risk for early intervention. In *The World Wide Web Conference*, WWW '19, 514–525. New York, NY, USA: Association for Computing Machinery. URL: https://doi.org/10.1145/3308558.3313698.

Haque, A., Reddi, V. and Giallanza, T. (2021) Deep learning for suicide and depression identification with unsupervised label correction.

Harrer, S. (2023) Attention is not all you need: the complicated case of ethically using large language models in healthcare and medicine. *EBioMedicine*, **90**, 104512.

Hendrycks, D., Burns, C., Basart, S., Zou, A., Mazeika, M., Song, D. and Steinhardt, J. (2021) Measuring massive multitask language understanding.

Holzinger, A., Keiblinger, K., Holub, P., Zatloukal, K. and Müller, H. (2023) AI for life: Trends in artificial intelligence for biotechnology. *N. Biotechnol.*, **74**, 16–24.

Huang, Y., Bai, Y., Zhu, Z., Zhang, J., Zhang, J., Su, T., Liu, J., Lv, C., Zhang, Y., Lei, J., Qi, F., Fu, Y., Sun, M. and He, J. (2023) C-eval: A multi-level multi-discipline chinese evaluation suite for foundation models. *ArXiv*, **abs/2305.08322**. URL: https://api.semanticscholar.org/CorpusID:258685666.

James, R. (2020) Mental health and behavior. URL: https://api.semanticscholar.org/CorpusID:229659501.

Ji, S., Li, X., Huang, Z. and Cambria, E. (2022a) Suicidal ideation and mental disorder detection with attentive relation networks. *Neural Computing and Applications*, **34**, 10309–10319.

— (2022b) Suicidal ideation and mental disorder detection with attentive relation networks. *Neural Computing and Applications*, **34**, 10309–10319.

Ji, S., Zhang, T., Ansari, L., Fu, J., Tiwari, P. and Cambria, E. (2021) Mentalbert: Publicly available pretrained language models for mental healthcare.

Ji, Y., Deng, Y., Gong, Y., Peng, Y., Niu, Q., Ma, B. and Belle, X. L. (2023) Be everyone's large language model engine.

Jin, H., Chen, S., Wu, M. and Zhu, K. Q. (2023) Psyeval: A comprehensive large language model evaluation benchmark for mental health. *arXiv preprint arXiv:2311.09189*.

Kamalloo, E., Dziri, N., Clarke, C. L. A. and Rafiei, D. (2023) Evaluating open-domain question answering in the era of large language models.

Lau, Y., Ang, W. H. D., Ang, W. W., Pang, P. C.-I., Wong, S. H. and Chan, K. S. (2025) Artificial intelligence–based psychotherapeutic intervention on psychological outcomes: A meta-analysis and meta-regression. *Depression and Anxiety*, **2025**, 8930012. URL: https://onlinelibrary.wiley.com/doi/abs/10.1155/da/8930012.

Li, Y., Choi, D., Chung, J., Kushman, N., Schrittwieser, J., Leblond, R., Eccles, T., Keeling, J., Gimeno, F., Dal Lago, A. et al. (2022) Competition-level code generation with alphacode. *Science*, **378**, 1092–1097.

Li, Y., Li, Z., Zhang, K., Dan, R., Jiang, S. and Zhang, Y. (2023) Chatdoctor: A medical chat model fine-tuned on a large language model meta-ai (llama) using medical domain knowledge.

Liang, P., Bommasani, R., Lee, T., Tsipras, D., Soylu, D., Yasunaga, M., Zhang, Y., Narayanan, D., Wu, Y., Kumar, A., Newman, B., Yuan, B., Yan, B., Zhang, C. and Cosgrove, C. A. (2023) Holistic evaluation of language models. *Transactions on Machine Learning Research*. URL: https://openreview.net/forum?id=iO4LZibEqW. Featured Certification, Expert Certification.

Lieber, O., Sharir, O., Lenz, B. and Shoham, Y. (2021) Jurassic-1: Technical details and evaluation. *White Paper. AI21 Labs*, **1**, 9.

Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H. and Neubig, G. (2023a) Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Comput. Surv.*, **55**, 1–35.

Liu, S., Deng, N., Sabour, S., Jia, Y., Huang, M. and Mihalcea, R. (2023b) Task-adaptive tokenization: Enhancing long-form text generation efficacy in mental health and beyond. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 15264–15281.

Ma, Z., Mei, Y. and Su, Z. (2023) Understanding the benefits and challenges of using large language model-based conversational agents for mental well-being support.

Mandapuram, M., Thodupunori, S. R., Bodepudi, A. and Reddy, M. (2018) Investigating the prospects of generative artificial intelligence. *Asian Journal of Humanity, Art and Literature*. URL: https://api.semanticscholar.org/CorpusID:258731430.

Martın-Rodilla, P., Losada, D. E. and Crestani, F. (2022) Overview of erisk 2022: Early risk prediction on the internet. In *Experimental IR Meets Multilinguality, Multimodality, and Interaction: 13th International Conference of the CLEF Association, CLEF 2022, Bologna, Italy, September 5–8, 2022, Proceedings*, vol. 13390, 233. Springer Nature.

Naveed, H., Khan, A., Qiu, S., Anwar, S., Usman, M., Barnes, N. and Mian, A. (2023) A comprehensive overview of large language models.

Ning, Y., Teixayavong, S., Shang, Y., Savulescu, J., Nagaraj, V., Miao, D., Mertens, M., Ting, D. S. W., Ong, J. C. L., Liu, M., Cao, J., Dunn, M., Vaughan, R., Ong, M. E. H., Sung, J. J.-Y., Topol, E. J. and Liu, N. (2024) Generative artificial intelligence and ethical considerations in health care: a scoping review and ethics checklist. *The Lancet Digital Health*, **6**, e848–e856. URL: https://www.sciencedirect.com/science/article/pii/S2589750024001432.

Nowacki, A., Sitek, W. and Rybiński, H. (2024) Llmental: Classification ofnbsp;mental disorders withnbsp;large language models. In *Foundations of Intelligent Systems: 27th International Symposium, ISMIS 2024, Poitiers, France, June 17–19, 2024, Proceedings*, 35–44. Berlin, Heidelberg: Springer-Verlag. URL: https://doi.org/10.1007/978-3-031-62700-2_4.

OpenAI, :, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., Avila, R., Babuschkin, I., Balaji, S., Balcom, V., Baltescu, P., Bao, H., Bavarian, M. et al. (2023) Gpt-4 technical report.

OpenAI (2024) Openai o1: Emphasizing reasoning in ai models. Available at https://www.reddit.com/r/ChatGPTPro/comments/1igl5me/i_built_3_apps_with_deepseek_openai_o1_and_gemini/.

Prystawski, B., Thibodeau, P., Potts, C. and Goodman, N. D. (2023) Psychologically-informed chain-of-thought prompts for metaphor understanding in large language models.

Qiu, H., Zhao, T., Li, A., Zhang, S., He, H. and Lan, Z. (2023) A benchmark for understanding dialogue safety in mental health support. In *Natural Language Processing and Chinese Computing* (eds. F. Liu, N. Duan, Q. Xu and Y. Hong), 1–13. Cham: Springer Nature Switzerland.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., Sutskever, I. et al. (2019) Language models are unsupervised multitask learners. *OpenAI blog*, **1**, 9.

Rae, J. W., Borgeaud, S., Cai, T., Millican, K., Hoffmann, J., Song, F., Aslanides, J., Henderson, S., Ring, R., Young, S. et al. (2021) Scaling language models: Methods, analysis & insights from training gopher. *arXiv preprint arXiv:2112.11446*.

Raffel, C., Shazeer, N. M., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W. and Liu, P. J. (2019) Exploring the limits of transfer learning with a unified text-to-text transformer. *J. Mach. Learn. Res.*, **21**, 140:1–140:67. URL: `https://api.semanticscholar.org/CorpusID:204838007`.

Rao, A., Kim, J., Kamineni, M., Pang, M., Lie, W. and Succi, M. D. (2023) Evaluating ChatGPT as an adjunct for radiologic decision-making. *medRxiv*.

Ratican, J. and Hutson, J. (2023) The six emotional dimension (6de) model: A multidimensional approach to analyzing human emotions and unlocking the potential of emotionally intelligent artificial intelligence (ai) via large language models (llm). *Journal of Artificial Intelligence and Robotics*, **1**.

Sagastibeltza, N., Salazar-Ramirez, A., Martinez, R., Jodra, J. L. and Muguerza, J. (2023) Automatic detection of the mental state in responses towards relaxation. *Neural Computing and Applications*, **35**, 5679–5696.

Saraceno, B. (2002) The who world health report 2001 on mental health. *Epidemiology and Psychiatric Sciences*, **11**, 83–87.

Sharma, A., Lin, I. W., Miner, A. S., Atkins, D. C. and Althoff, T. (2021) Towards facilitating empathic conversations in online mental health support: A reinforcement learning approach. WWW '21, 194–205. New York, NY, USA: Association for Computing Machinery. URL: `https://doi.org/10.1145/3442381.3450097`.

— (2023a) Human–ai collaboration enables more empathic conversations in text-based peer-to-peer mental health support. *Nature Machine Intelligence*, **5**, 46–57.

Sharma, A., Sharma, K. and Kumar, A. (2023b) Real-time emotional health detection using fine-tuned transfer networks with multimodal fusion. *Neural computing and applications*, **35**, 22935–22948.

Shatte, A. B. R., Hutchinson, D. M. and Teague, S. J. (2019) Machine learning in mental health: a scoping review of methods and applications. *Psychological Medicine*, **49**, 1426 – 1448. URL: `https://api.semanticscholar.org/CorpusID:73426701`.

Shen, G., Jia, J., Nie, L., Feng, F., Zhang, C., Hu, T., Chua, T.-S. and Zhu, W. (2017) Depression detection via harvesting social media: A multimodal dictionary learning solution. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, IJCAI'17, 3838–3844. AAAI Press.

Shing, H.-C., Resnik, P. and Oard, D. (2020) A prioritization model for suicidality risk assessment. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 8124–8137. Online: Association for Computational Linguistics. URL: `https://aclanthology.org/2020.acl-main.723`.

Singhal, K., Azizi, S., Tu, T., Mahdavi, S. S., Wei, J., Chung, H. W., Scales, N., Tanwani, A., Cole-Lewis, H., Pfohl, S., Payne, P., Seneviratne, M., Gamble, P., Kelly, C., Scharli, N., Chowdhery, A., Mansfield, P., y Arcas, B. A., Webster, D., Corrado, G. S., Matias, Y., Chou, K., Gottweis, J., Tomasev, N., Liu, Y., Rajkomar, A., Barral, J., Semturs, C., Karthikesalingam, A. and Natarajan, V. (2022) Large language models encode clinical knowledge.

Solaiman, B. (2024) Generative artificial intelligence (genai) and decision-making: Legal ethical hurdles for implementation in mental health. *International Journal of Law and Psychiatry*, **97**, 102028. URL: `https://www.sciencedirect.com/science/article/pii/S0160252724000773`.

Soltan, S., Ananthakrishnan, S., FitzGerald, J., Gupta, R., Hamza, W., Khan, H., Peris, C., Rawls, S., Rosenbaum, A., Rumshisky, A. et al. (2022) Alexatm 20b: Few-shot learning using a large-scale multilingual seq2seq model. *arXiv preprint arXiv:2208.01448*.

Srivastava, A., Rastogi, A., Rao, A., Shoeb, A. A. M., Abid, A., Fisch, A. and Brown, A. R. (2023) Beyond the imitation game: Quantifying and extrapolating the capabilities of language models. *Transactions on Machine Learning Research*. URL: https://openreview.net/forum?id=uyTL5Bvosj.

Srivastava, A., Suresh, T., Lord, S. P., Akhtar, M. S. and Chakraborty, T. (2022) Counseling summarization using mental health knowledge guided utterance filtering. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, KDD '22, 3920–3930. New York, NY, USA: Association for Computing Machinery. URL: https://doi.org/10.1145/3534678.3539187.

Strout, K. A. and Howard, E. P. (2012) The six dimensions of wellness and cognition in aging adults. *Journal of Holistic Nursing*, **30**, 195–204. URL: https://doi.org/10.1177%2F0898010112440883.

Sun, H., Lin, Z., Zheng, C., Liu, S. and Huang, M. (2021a) PsyQA: A Chinese dataset for generating long counseling text for mental health support. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, 1489–1503. Online: Association for Computational Linguistics. URL: https://aclanthology.org/2021.findings-acl.130.

Sun, Y., Wang, S., Feng, S., Ding, S., Pang, C., Shang, J., Liu, J., Chen, X., Zhao, Y., Lu, Y., Liu, W., Wu, Z., Gong, W., Liang, J., Shang, Z., Sun, P., Liu, W., Xuan, O., Yu, D., Tian, H., Wu, H. and Wang, H. (2021b) Ernie 3.0: Large-scale knowledge enhanced pre-training for language understanding and generation. *ArXiv*, **abs/2107.02137**. URL: https://api.semanticscholar.org/CorpusID:235731579.

Taori, R., Gulrajani, I., Zhang, T., Dubois, Y., Li, X., Guestrin, C., Liang, P. and Hashimoto, T. B. (2023) Stanford alpaca: an instruction-following llama model (2023). *URL https://crfm. stanford. edu/2023/03/13/alpaca. html*, **1**, 3.

Turcan, E. and McKeown, K. (2019) Dreaddit: A reddit dataset for stress analysis in social media.

Wang, H., Wang, R., Mi, F., Deng, Y., Wang, Z., Liang, B., Xu, R. and Wong, K.-F. (2023a) Cue-cot: Chain-of-thought prompting for responding to in-depth dialogue questions with llms.

Wang, S., Sun, Y., Xiang, Y., Wu, Z., Ding, S., Gong, W., Feng, S., Shang, J., Zhao, Y., Pang, C. et al. (2021) Ernie 3.0 titan: Exploring larger-scale knowledge enhanced pre-training for language understanding and generation. *arXiv preprint arXiv:2112.12731*.

Wang, Y., Kordi, Y., Mishra, S., Liu, A., Smith, N. A., Khashabi, D. and Hajishirzi, H. (2022a) Self-instruct: Aligning language model with self generated instructions. *arXiv preprint arXiv:2212.10560*.

Wang, Y., Le, H., Gotmare, A. D., Bui, N. D. Q., Li, J. and Hoi, S. C. H. (2023b) Codet5+: Open code large language models for code understanding and generation.

Wang, Y., Mishra, S., Alipoormolabashi, P., Kordi, Y., Mirzaei, A., Naik, A., Ashok, A., Dhanasekaran, A. S., Arunkumar, A. and Stap, D. (2022b) Super-NaturalInstructions: Generalization via declarative instructions on 1600+ NLP tasks. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, 5085–5109. Abu Dhabi, United Arab Emirates: Association for Computational Linguistics. URL: https://aclanthology.org/2022.emnlp-main.340.

Wang, Y., Zhao, Y. and Petzold, L. (2023c) Are large language models ready for healthcare? a comparative study on clinical language understanding.

Wu, C., Zhang, X., Zhang, Y., Wang, Y. and Xie, W. (2023) Pmc-llama: Further finetuning llama on medical papers. *ArXiv*, **abs/2304.14454**. URL: https://api.semanticscholar.org/CorpusID:263888272.

Wu, S., Zhao, X., Yu, T., Zhang, R., Shen, C., Liu, H., Li, F., Zhu, H., Luo, J., Xu, L. et al. (2021) Yuan 1.0: Large-scale pre-trained language model in zero-shot and few-shot learning. *arXiv preprint arXiv:2110.04725*.

Xu, X., Yao, B., Dong, Y., Gabriel, S., Yu, H., Hendler, J., Ghassemi, M., Dey, A. K. and Wang, D. (2023) Mental-llm: Leveraging large language models for mental health prediction via online text data.

Xue, L., Constant, N., Roberts, A., Kale, M., Al-Rfou, R., Siddhant, A., Barua, A. and Raffel, C. (2021) mt5: A massively multilingual pre-trained text-to-text transformer.

Yates, A., Cohan, A. and Goharian, N. (2017) Depression and self-harm risk assessment in online forums. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 2968–2978. Copenhagen, Denmark: Association for Computational Linguistics. URL: https://aclanthology.org/D17-1322.

Zeng, A., Liu, X., Du, Z., Wang, Z., Lai, H., Ding, M., Yang, Z., Xu, Y., Zheng, W., Xia, X., Tam, W. L., Ma, Z., Xue, Y., Zhai, J., Chen, W., Zhang, P., Dong, Y. and Tang, J. (2023) Glm-130b: An open bilingual pre-trained model.

Zeng, A., Liu, X., Du, Z., Wang, Z., Lai, H., Ding, M., Yang, Z., Xu, Y., Zheng, W., Xia, X. et al. (2022) Glm-130b: An open bilingual pre-trained model. *arXiv preprint arXiv:2210.02414*.

Zeng, W., Ren, X., Su, T., Wang, H., Liao, Y., Wang, Z., Jiang, X., Yang, Z. and Wang, K. (2021) Pangu-α: Large-scale autoregressive pretrained chinese language models with auto-parallel computation.

Zhang, J., Gan, R., Wang, J., Zhang, Y., Zhang, L., Yang, P., Gao, X., Wu, Z., Dong, X., He, J., Zhuo, J., Yang, Q., Huang, Y., Li, X., Wu, Y., Lu, J., Zhu, X., Chen, W., Han, T., Pan, K., Wang, R., Wang, H., Wu, X., Zeng, Z. and Chen, C. (2023) Fengshenbang 1.0: Being the foundation of chinese cognitive intelligence.

Zhang, S., Roller, S., Goyal, N., Artetxe, M., Chen, M., Chen, S., Dewan, C., Diab, M. T., Li, X., Lin, X. V., Mihaylov, T., Ott, M., Shleifer, S., Shuster, K., Simig, D., Koura, P. S., Sridhar, A., Wang, T. and Zettlemoyer, L. (2022) Opt: Open pre-trained transformer language models. *ArXiv*, **abs/2205.01068**. URL: `https://api.semanticscholar.org/CorpusID:248496292`.

Zheng, Z., Liao, L., Deng, Y. and Nie, L. (2023) Building emotional support chatbots in the era of llms.

Zhong, W., Cui, R., Guo, Y., Liang, Y., Lu, S., Wang, Y., Saied, A., Chen, W. and Duan, N. (2023) Agieval: A human-centric benchmark for evaluating foundation models.

Alzheimer's Disease Detection Using Bidirectional Empirical Mode Decomposition

Neha Garg
Department of Electronics and Communication Engg.
Ajay Kumar Garg Engineering College,
Ghaziabad, India,
ngcoolneha@gmail.com

Amit Garg
Department of Electronics and Communication Engg.
Ajay Kumar Garg Engineering College, Ghaziabad, India,
amitstp@gmail.com

Mahipal Singh Choudhry
Department of Electronics and Communication Engg.
Delhi Technological University
Delhi, India,
msc_1976@yahoo.com

2025 10th International Conference on Signal Processing and Communication (ICSC) | 979-8-3315-1189-0/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICSC64553.2025.10967815

***Abstract*— Alzheimer's disease is a progressively deteriorated, irreversible brain ailment that impairs memory and cognitive function. Early detection of Alzheimer's disease is crucial as it can reduce the risk factors and can provide the effective therapy to the affected person within stipulated time. In the proposed approach, bidirectional empirical mode decomposition is applied on 2-D Magnetic Resonance Images to decompose them into four Intrinsic Mode Functions. Then, using four Intrinsic Mode Functions, the Gray Level Co-occurrence Matrix has been computed. The Gray Level Co-occurrence Matrix interacts directly with the image intensity, it offers information on patterns that mutually occur. For every Intrinsic Mode Functions, 22 statistical features are obtained from these matrices. In order to improve the classification accuracy, Particle Swarm Optimization (PSO) is further utilized for extracting optimal and relevant features. The classification results for proposed method are 96.4% accuracy, 96.0 % sensitivity and 97.0 % specificity. The proposed work outperforms the existing methods for AD detection in terms of its outstanding classification results.**

Keywords—Alzheimer's disease detection, Bidirectional Empirical Mode Decomposition, Gray Level Co-occurrence Matrix, Magnetic Resonance Imaging, Particle Swarm Optimization, Feature Selection

I. INTRODUCTION

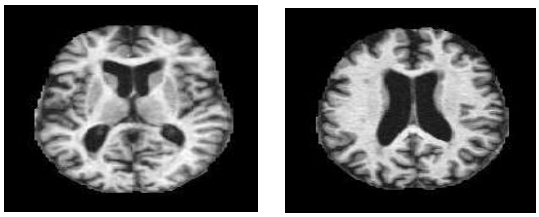
Alzheimer's disease (AD) is a fatal neurological disorder that primarily impairs older adults' memory and cognitive skills. AD is a serious health condition that spreads globally and primarily affects those over the age of 65 years [1]. Since AD cannot be cured completely at present and can only be treated to delay its progression, it is essential to precisely diagnose the patient's illness at right time. The early diagnosis of AD is essential for managing and reducing its effect. AD is triggered by the growth of tangles in the brain. These tangles are due to the accumulation of amyloid plaques and tau proteins. This process results in breaking the synaptic and neuronal connections in the brain [2]. The neuronal loss creates structural changes in the brain. These changes can be visualised through diagnostic tests like Magnetic Resonance Imaging (MRI), Positron emission tomography (PET), Diffusion Tensor Imaging (DTI). These structural changes are altered cortical thickness, enlarged ventricles, reduced grey matter tissue density, and reduced hippocampal size, shape, and volume [3]. The MR technique provides images with good resolution and high contrast. This makes MRI as the best technique to examine structural changes and grey matter density fluctuations in the brain [4]. The features obtained from MRI are more informative and relevant and thus useful for early AD detection. Researchers have presented a variety of feature extraction techniques in the literature such as Voxel Based Morphometry (VBM), Surface Based Morphometry (SBM) and Region of Interest (ROI) based methods [5]. VBM tests the white matter and grey matter density in the brain of different groups of subjects or within a same subject over time. VBM suffers from misalignment and imperfect image registration issues. SBM helps in analysing structural changes in the cortical surface of the brain. This method faces challenges while accomplishing cortical surface reconstruction. ROI based method focussed on the affected part of the brain due to AD. The accuracy of this method depends upon the expertise of the researcher [5]. This paper proposed a method which detects the structural changes and grey matter density fluctuations from the whole brain rather than a specific region. The whole-brain approaches do not need to extract any part from the brain rather they use whole brain for extracting informative features [6].

In this paper, the image is decomposed into four Intrinsic Mode Functions (IMFS) using Bidirectional Empirical Mode Decomposition (BEMD). Next, each IMF's Grey Level Co-occurrence Matrix (GLCM) is obtained, and the 22 statistical features are taken out from each IMF's GLCM. Particle Swarm Optimization (PSO) is applied for selecting significant and appropriate features from these 88 statistical features. The paper contributes in the following way:

- This work introduces a new feature extraction technique for AD classification that can be utilized to identify various AD stages by measuring the spatial pattern of structural and grey level details, which vary with AD severity.
- The proposed method uses BEMD with GLCM for feature extraction. BEMD provides multiscale decomposition of the image into IMFs, allowing for a detailed analysis of signal components at different scales. GLCM helps in finding the grey scale variation between pixels in an image by determining the spatial relationship between pixels

979-8-3315-1189-0/25/$31.00 ©2025 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:34:29 UTC from IEEE Xplore. Restrictions apply.



(i) (ii)

Fig. 1. 2-D MR slice from OASIS dataset (i) Alzheimer's Disease and (ii) Normal Control

separated by a given distance. The combined strength of BEMD and GLCM improves the classification accuracy by providing more informative features.

- The proposed method captures both local and global structural information from different parts of the brain which is very beneficial for detecting AD at different stages.

The efficacy of the proposed approach is evaluated by comparing it with three existing methods using performance measures as accuracy, sensitivity, and specificity. The paper has been divided into following sections: Section II describes the dataset and background information needed to implement the proposed work; Section III describes the results of the implementation work; and Section IV completes this paper with recommendations for future research.

II. Proposed Method And Material

A. Database

The Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. Another dataset is Open Access Series of Imaging Studies (OASIS). From the OASIS [7] and ADNI collections, a total of 28 AD and 28 NC subjects are chosen, and their 3-D MR images are retrieved. The four center slices are taken out of each 3-D MR image. These 2-D MR images, or center slices, are employed in classification. There are 112 2-D MR images overall in each class. OASIS is a data warehouse where scientists may access and gather brain MRIs for their research projects. The MRI sample that is available at http://www.oasis-brains.org is shown in Fig. 1.

B. Method

The proposed approach employs BEMD for image decomposition, GLCM for extracting statistical features and PSO for selecting appropriate features. The selected features are used for AD classification. Three groups have been implemented in this work. The group 1 includes the direct application of GLCM over MR images. Total 22 statistical features have been extracted from these GLCMs matrices and have been used for classification. The second group combines BEMD and GLCM for extracting features. BEMD has been applied on 2-D MR scans to obtain four IMFs with decomposition at first level. Subsequently GLCM has been calculated for these four IMFs. Total 88 statistical features are obtained from these four IMFs with contribution of 22 features from each IMF. These features have been used for AD classification in group 2. The third group includes PSO with BEMD and GLCM. PSO selects the optimum features from these 88 features. The number of features selected by PSO are 34. These features are used for AD classification. GLCM efficiently extracts the key features from an image by giving all the information about the direction, adjacent pixel interval, and grey level changes. PSO is simple to implement and provides globally optimal solutions. PSO supports parallel computing architectures and efficiently explores the search space. All these properties make it a valuable tool for enhancing the classification performance of groups. Fig. 2 illustrates the flow of process in the proposed approach.

C. Bidirectional Empirical Mode Decomposition

BEMD [8-10] is the modified version of EMD. EMD is used for analysing nonlinear and non-stationary signals by breaking down it into Intrinsic Mode Functions (IMFs). These IMFs are obtained by iteratively sifting through the signal to extract oscillatory modes with low to high frequency ranges. Each IMF corresponds to scale or frequency component of the original signal. By decomposing MRI signals into IMFs, BEMD can capture both global and local features present in the images, providing valuable information for image interpretation and analysis. The difference in between EMD and BEMD is that EMD supports unidirectional sifting and BEMD supports bidirectional sifting. The bidirectional sifting helps in capturing temporal changes in the signal potentially and the most favourable case include complex non-stationary signal. BEMD is adaptive and data driven. BMED does not depend on predefined basis functions like wavelets for analysing the signal's underlying structure. Rather, it amends itself according to the local characteristics of the signal and thus handles the signals with varying spectral content and non-stationary dynamics effectively. BEMD offers multiple properties like multi-scale analysis, artifact removal, feature extraction, and dynamic analysis and thus becomes a valuable tool for advancing MRI research.

D. Grey Level Co-occurrence Matrix

The GLCM of an image provides all the information regarding the direction, intensity change in the grey level, and neighbouring interval between the pixels [11, 12]. GLCM is suitable for feature extraction because of this property.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:34:29 UTC from IEEE Xplore. Restrictions apply.

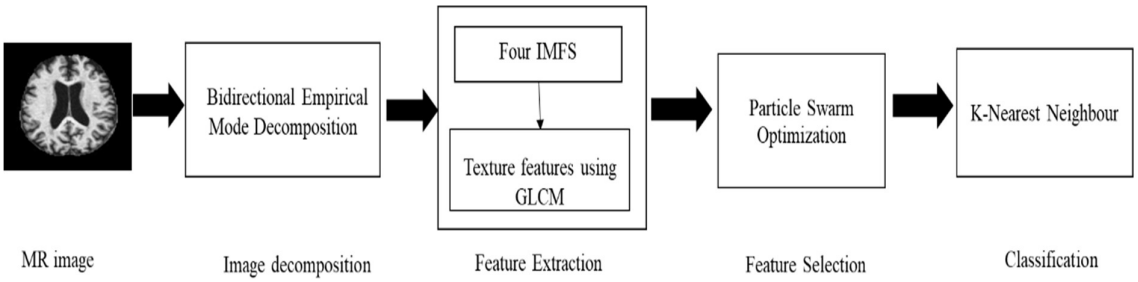


Fig. 2. Proposed Methodology

Let suppose grey scale image I_g contains N_b number of grey levels thus the GLCM ($G_{m,n}$) of image will be of order N_b. The (m,n)th value in $G_{m,n}$ indicates the frequency of distribution of two adjacent pixels with intensities 'm' and 'n'. These pixels may be at distance "d" with d=1,2 3 so on. The orientation of these pixels is at any four angles 0º, 45º, 90º,135º which is shown in Fig. 3. The 22 Haralick texture features can be computed for all four directions [13]. GLCM provides following 22 statistical features.

1. Angular second Moment: It represents uniform patterns in the image.

2 Joint Entropy: It represents the measure of randomness of joint distribution of two variables.

3.**Dissimilarity:** It defines the linear metrics of local changes in an image.

4.Contrast / Inertia: It quantifies the difference in between the highest and lowest intensity value of two adjacent pixels.

5.Homogeneity1/ Inverse difference: It indicates the uniformity of pixel values in a small region of an image.

6.Correlation: It represents how the two pixels are linearly related.

7.Homogeneity 2/ Inverse difference moment: It shows the local homogeneity of an image.

8.Autocorrelation: It measures the roughness of an image.

9.Cluster Shade: It calculates the skewness of GLCM distribution capturing the asymmetry of cluster distribution in images.

10.Cluster Prominence: It gauges variations in the pixel's local intensity. Its high value denotes the image's asymmetry.

11.Maximum probability: It indicates the instances of the most common pair of neighbouring intensity values.

12.Sum of Squares or variance: It calculates the gray intensity level dispersion around the mean value.

13.Sum Average: It calculates the mean of the total of the image's potential grey-level intensities.

14.Sum Variance: It calculates the average dispersion of the total possible grey-level image intensities around the mean value.

15.Sum Entropy: It gauges the amount of clutter in relation to the image's total grey-level intensity.

16.Difference variance: It is a measurement of the image's varying grey level spread dispersion with respect to the mean.

17.Difference entropy: It measures the clutter in relation to the image's distribution of grey-level variation.

18.Information measures of correlation (1): It calculates the texture's complexity and correlation strength.

19.Information measures of correlation (2): It computes the strength of correlation, and the complexity of texture.

20.Maximal correlation coefficient: It provides understanding of complexity of textures.

21.Inverse difference normalized (INN): It indicates the level of smoothness in local regions of an image.

22.Inverse difference moment normalized (IDN): IDN gives details about how rapidly pixel values change within the image.

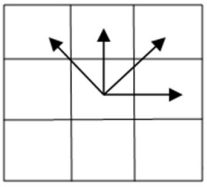


Fig. 3. GLCM with four directions

E. Particle Swarm Optimisation (PSO)

The Particle Swarm Optimization (PSO) originally presented by Kennedy and Eberhart in 1995 is a computational optimization algorithm inspired by the social behaviour of birds [14-15]. PSO shows parallel processing thus offers high convergence rate and low computation time. PSO searches for local and global optima in the search space [16]. It needs less parameters for tuning. The algorithm is inspired by the intelligence of swarm which can be seen when they search for food. Their mutual association, imparting, and learning help them for searching food. The particles in PSO can be considered as swarms. The particles in the search space have to maintain their location and velocity, and they update their location based on their own experience (personal best) and the experiences of neighbouring particles (global best). As a result, PSO effectively explores the search space, with each particle having the ability to investigate nearby areas of the solution space and being impacted by the swarm's overall best solution. After 't^{th}' iteration in search space, the position of particle p is given as

$$\mathrm{Po}_p^t = \left[\mathrm{po}_{0,p}^t, \mathrm{po}_{1,p}^t, \mathrm{po}_{2,p}^t, \mathrm{po}_{3,p}^t \dots\dots, \mathrm{po}_{N,p}^t\right] \quad (1)$$

where N is the dimensional space of the problem. Each particle is moving in search space for searching global minima as an optimum solution. The particle p moving with velocity 'Vt' at t^{th} iteration is given as

$$\mathrm{Vt}_p^t = \left[\mathrm{vt}_{0,p}^t, \mathrm{vt}_{1,p}^t, \mathrm{vt}_{2,p}^t, \mathrm{vt}_{3,p}^t \dots\dots, \mathrm{vt}_{N,p}^t\right] \quad (2)$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:34:29 UTC from IEEE Xplore. Restrictions apply.

After (t+1) [th] iteration, the position of the pth particle is given by $Po_p^{t+1} = Po_p^t + Vt_p^{t+1}$ (3)

And the velocity of particle after (t+1)[th] iteration is given by

$$Vt_p^{t+1} = w1Vt_p^t + c_1r_1\left(pbt_p^t - Po_p^t\right) + c_2r_2\left(gbt_p^t - Po_p^t\right) \quad (4)$$

In equation (4), $'pbt'$ is the best position searched by the pth particle and $'gbt'$ is the best position searched by the complete group of particles. The 'w1' represents the weight inertia constant, c1 and c2 represent the acceleration coefficients and r_1 and r_2 represent the random numbers. The values of c1, c2, w1, N and t used in implementation work are 2,2, 0.9,30 and 100 respectively. Fig. 4 illustrates the PSO convergence graph in proposed work.

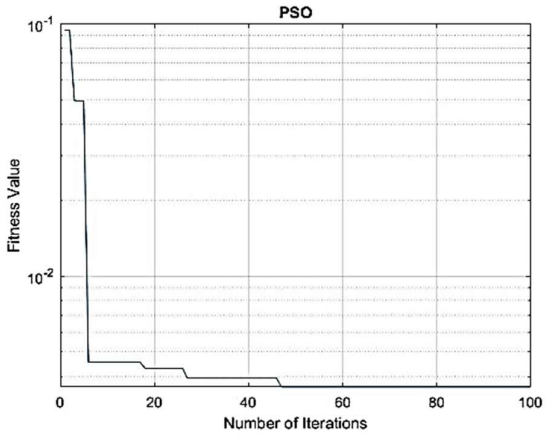


Fig. 4. Convergence Plot

III. SIMULATION RESULTS AND DISCUSSION

A. Quantitative Analysis of proposed method

The performance of the proposed work is evaluated by using parameters like Accuracy, Sensitivity and Specificity Accuracy [17] defines the proportion of rightly classified true positives and true negatives out of the total number of cases in the dataset. Sensitivity [18-19] defines the proportion of rightly classified true positives and specificity defines the proportion of rightly classified true negatives. These are mathematically formulated as

$$\text{Accuracy (Ac)} = \frac{(Tre_{pt}+Tre_{ng})}{(Tre_{ng}+Tre_{pt}+Fs_{ng}+Fs_{pt})} \quad (5)$$

$$\text{Sensitivity (Se)} = \frac{(Tre_{pt})}{(Tre_{pt}+Fs_{ng})} \quad (6)$$

$$\text{Specificity (Sp)} = \frac{(Tre_{ng})}{(Tre_{ng}+Fs_{pt})} \quad (7)$$

Where, Tre_{ng} as true negative, Tre_{pt} as true positive, Fs_{ng} as false negative and Fs_{pt} as false positive. False negative means the positive samples are detected as negative. False positive means the normal samples are detected as positives. True positive means the positive samples are detected as positives and True negative means the negative samples are detected as negatives.

B. Performance Evaluation

The proposed work provides outstanding results for AD classification. Three groups have been implemented and compared for analysing the efficiency of proposed work. The first group includes the direct implementation of GLCM on 2-D MR scans. The total 22 statistical features are obtained from matrices for all four directions at d=1.The results with this group are 92.1% using k-NN classifier. The second group includes the application of BEMD over MR scans first and then GLCM is applied. BEMD provides four IMFs and GLCM is calculated for four IMFs. The 22 statistical features are extracted from each IMF and used for AD classification. This group provides more informative features as it extracts micro and macro details from local and global parts of the brain. The classification accuracy with this group is 94.6 %. The third group includes the proposed approach which adds feature selection stage after BEMD and GLCM. The PSO has been used for feature selection process. This group provides 96.4% accuracy,96% sensitivity and 97% specificity. Table I illustrates the performance comparison between proposed work and different groups implemented in this study using OASIS dataset.

TABLE I: COMPARATIVE ANALYSIS OF DIFFERENT GROUPS

Method	Ac				Se				Sp			
	0°	45°	90°	135°	0°	45°	90°	135°	0°	45°	90°	135°
GLCM	92	92	91.5	91.5	88	91	96	93	96	93	94	90
BEMD+GLCM	94.6	94.6	94.2	94.2	92	95	97	96	97	95	91	98
BEMD+GLCM+PSO [Proposed]	95.1	96.4	96	96.4	93.2	96	97	96	97.2	96	95	96

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:34:29 UTC from IEEE Xplore. Restrictions apply.

Table II: Performance Comparison Between Proposed and Current Algorithms

Methods	OASIS Dataset (Ac/Se/Sp)	ADNI Dataset (Ac/Se/Sp)
BEMD+GLCM+PSO [Proposed]	96.4/96/97	95.1/94/96
Graph Kernel Approach [20]	92/91/93	90.2/94/87
Spiking Neural Network [21]	91/90/92	90.1/94/87
Hessian detector +LBP [22]	85.5/87/84	86.2/87/86
VGG-16[23]	86.7/85.2/87.3	85.4/86.5/84.6
Voxel-wise GM [24]	89.3/87/92	90.8/86.3/94.8

Table II lists the classification results of five existing methods [20-24] and proposed method on MATLAB 2019 with OASIS and ADNI datasets. The proposed method outperforms the current methods. The recent study [20] used graph kernel approach for extracting 3-D texture features. This method provides 92% accuracy, 91% sensitivity and 93% specificity while implementation on MATLAB with OASIS dataset. The study [21] used spiking neural network for AD classification. This study is based on deep learning including the strength of biologically inspired neurons. This method provides an accuracy, sensitivity, and specificity of 91% ,90% and 92% respectively on OASIS dataset. The study [22] combined the strength of hessian detector with Local Binary Pattern. This method gives 85.5% of accuracy,87% of sensitivity and 84% of specificity on implementation. The study [23] used deep learning model i.e. VGG-16 model for AD classification. The study [24] used voxel-based morphometry to extract grey density maps. These maps were used as features for AD classification. The results with the proposed method are 96.4% accuracy, 96% sensitivity, and 97% specificity using k-NN classifier on OASIS dataset. While on ADNI dataset, these figures are 95.1% accuracy, 94% sensitivity and 96% specificity. The proposed approach gives outstanding results for AD classification. The Receiver Operating Characteristic (ROC) graph for proposed and current methods is illustrated in Fig. 5 and Fig. 6 for OASIS and ADNI datasets, respectively.

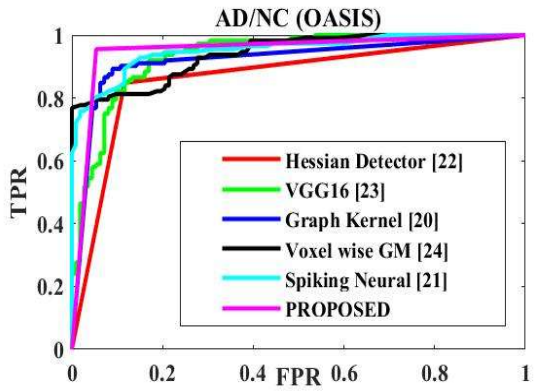


Fig. 5. ROC plot on OASIS dataset

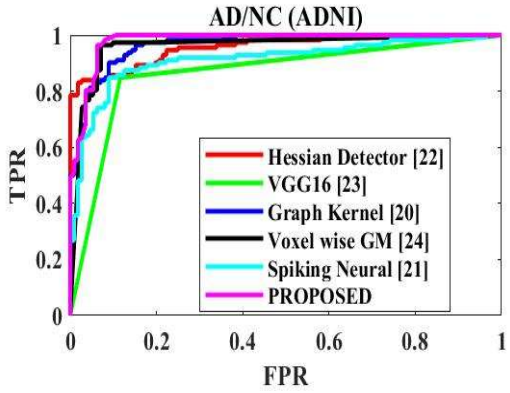


Fig. 6. ROC plot on ADNI dataset

IV. CONCLUSION

The strength of BEMD, GLCM, and PSO are combined in this study to propose a new feature extraction method for AD detection. The BEMD's ability to handle images with non-stationary characteristics and GLCM's strength in texture analysis make proposed approach strong enough to capture subtle changes in brain with different stages of AD. The experiments conducted demonstrate the ability of proposed method for capturing informative features for AD detection. The proposed approach gives 96.4% of accuracy,96.0% of sensitivity and 97% of specificity. In future this approach can be used for investigating the classification at different stages of AD. The proposed approach can be validated with different modes of decomposition and modalities with different feature selection methods.

ACKNOWLEDGEMENT

I am thankful to my guide and colleague for their valuable support. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation;

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:34:29 UTC from IEEE Xplore. Restrictions apply.

Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

REFERENCES

[1] Kuo, Y. C., & Rajesh, R, "Challenges in the treatment of Alzheimer's disease: recent progress and treatment strategies of pharmaceuticals targeting notable pathological factors," *Expert review of neurotherapeutics*, vol. *19, no. 7*, pp. 623–652, 2019.

[2] Ossenkoppele, R., van der Kant, R., & Hansson, O, "Tau biomarkers in Alzheimer's disease: towards implementation in clinical practice and trials," *The Lancet Neurology*, vol. 21, no.8, pp. 726-734, 2022.

[3] Ávila-Villanueva, M., Marcos Dolado, A., Gómez-Ramírez, J., & Fernández-Blázquez, M., "Brain Structural and Functional Changes in Cognitive Impairment Due to Alzheimer's Disease," *Frontiers in psychology*, vol. *13*, p. 886619, 2022.

[4] Knudsen, L. V., Gazerani, P., Duan, Y., Michel, T. M., & Vafaee, M. S., "The role of multimodal MRI in mild cognitive impairment and Alzheimer's disease.," *Journal of neuroimaging official journal of the American Society of Neuroimaging*, vol. *32, no.* 1, pp. 148–157, 2022.

[5] Garg, N., Choudhry, M. S., & Bodade, R. M., "A review on Alzheimer's disease classification from normal controls and mild cognitive impairment using structural MR images," *Journal of neuroscience methods*, vol. *384*, p. 109745, 2023.

[6] Liu, J., Li, M., Lan, W., Wu, F. X., Pan, Y., & Wang, J., "Classification of Alzheimer's Disease Using Whole Brain Hierarchical Network," IEEE/*ACM transactions on computational biology and bioinformatics*, vol. *15, no.* 2, pp. 624–632, 2018.

[7] Ardekani, B. A., Figarsky, K., & Sidtis, J. J., "Sexual dimorphism in the human corpus callosum: an MRI study using the OASIS brain database," *Cerebral cortex (New York, N.Y.: 1991)*, vol. *23*, no. 10, pp. 2514–2520, 2018.

[8] De Souza, U. B., Escola, J. P. L., Vedovatto, T., da Cunha Brito, L., & Lemos, R. P., "Bidirectional EMD-RLS: Performance analysis for denoising in speech signal," *Journal of Computational Science*, vol. *74*, p. 102181, 2023.

[9] Weng, Binwei, and Kenneth E. Barner, "Optimal and bidirectional optimal empirical mode decomposition," *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*. Vol. 3. IEEE, 2007.

[10] Acharya, U. R., Koh, J. E. W., Hagiwara, Y., Tan, J. H., Gertych, A., Vijayananthan, A. &Yeong, C. H., "Automated diagnosis of focal liver lesions using bidirectional empirical mode decomposition features, "*Computers in biology and medicine*, vol. *94*, pp. 11-18, 2018.

[11] Feng, Q., & Ding, Z., "MRI Radiomics Classification and Prediction in Alzheimer's Disease and Mild Cognitive Impairment: A Review, " *Current Alzheimer research*, vol. *17, no.* 3, pp. 297–309, 2020.

[12] Barburiceanu, S., Terebes, R., & Meza, S., "3D texture feature extraction and classification using GLCM and LBP-based descriptors,".*Applied Sciences*, vol. *11, no.* 5, p. 2332, 2021.

[13] Haralick, R. M., Shanmugam, K., & Dinstein, I. H., "Textural features for image classification," *IEEE Transactions on systems, man, and cybernetics*, vol. 6, pp. 610-621, 1973.

[14] Poli, Riccardo, James Kennedy, and Tim Blackwell, "Particle swarm optimization," *Swarm intelligence*, vol. 1, pp. 33-57, 2007.

[15] Piotrowski, Adam P., Jaroslaw J. Napiorkowski, and Agnieszka E. Piotrowska, "Population size in particle swarm optimization," *Swarm and Evolutionary Computation, vol.* 58, p.100718, 2020.

[16] Gu, Qinghua, Yingyin Liu, Lu Chen, and Naixue Xiong, "An improved competitive particle swarm optimization for many-objective optimization problems," *Expert Systems with Applications, vol.* 189, p. 116118, 2022.

[17] Trevethan, R., "Sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice," *Frontiers in public health*, vol. *5*, p. 307, 2017.

[18] Lalkhen, A. G., & McCluskey, A., "Clinical tests: sensitivity and specificity," *Continuing education in anaesthesia critical care & pain*, vol. *8, no.* 6, 221-223, 2008.

[19] Parikh, R., Mathai, A., Parikh, S., Sekhar, G. C., & Thomas, R., "Understanding and using sensitivity, specificity and predictive values," *Indian journal of ophthalmology*, vol. *56, no.* 1, pp. 45-50, 2008.

[20] de Mendonça, Lucas José Cruz, Ricardo José Ferrari, and Alzheimer's Disease Neuroimaging Initiative, "Alzheimer's disease classification based on graph kernel SVMs constructed with 3D texture features extracted from MR images," *Expert Systems with Applications*, vol. 211, p. 118633, 2023.

[21] Turkson, Regina Esi, Hong Qu et al., "Classification of Alzheimer's disease using deep convolutional spiking neural network," *Neural Processing Letters, vol.* 53, no. 4, pp. 2649-2663, 2021.

[22] Francis, Ambily, Immanuel Alex Pandian, and Alzheimer's Disease Neuroimaging Initiative, "Early detection of Alzheimer's disease using local binary pattern and convolutional neural network," *Multimedia Tools and Applications, vol.* 80, no. 19, pp. 29585-29600, 2021.

[23] Sharma, Shagun, Kalpna Guleria, Sunita Tiwari, and Sushil Kumar, "A deep learning based convolutional neural network model with VGG16 feature extractor for the detection of Alzheimer Disease using MRI scans," *Measurement: Sensors, vol.* 24, p. 100506, 2022.

[24] Liu, Manhua, Daoqiang Zhang, Dinggang Shen, and Alzheimer's Disease Neuroimaging Initiative, "Ensemble sparse classification of Alzheimer's disease," *NeuroImage, vol.* 60, no. 2, p. 1106-1116, 2012.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:34:29 UTC from IEEE Xplore. Restrictions apply.

Check for updates

Journal of Neuroscience Research

WILEY

JOURNAL OF Neuroscience Research

RESEARCH ARTICLE

An Interplay Between Hypothalamic Microstructure, Systemic Metabolism and Gut Microbiome Composition in Male Rats at Hyperacute Timepoint Post TBI

Palkin Arora[1,2] | Megha Kumari[1,3] | Kavita Singh[1] | M. Memita Devi[1] | Poonam Rana[1] | Rajat Sandhir[2] | Richa Trivedi[1]

[1]Radiological, Nuclear and Imaging Sciences (RNAIS), Institute of Nuclear Medicine and Allied Sciences (INMAS), DRDO, Delhi, India | [2]Department of Biochemistry, Panjab University, Chandigarh, India | [3]Department of Biotechnology, Delhi Technological University (DTU), Delhi, India

Correspondence: Richa Trivedi (richa.inmas@gov.in; triricha@gmail.com)

Received: 11 April 2024 | **Revised:** 22 February 2025 | **Accepted:** 19 March 2025

Edited by: Lawrence S. Sherman and Martin L. Doughty. Reviewed by John Wu

Funding: This work was supported by the Defense Research & Development Organization (DRDO), Ministry of Defense, India (Grant INM-324). Palkin Arora is supported by a DRDO fellowship (Grant EST/3009/JRF/2020).

Keywords: diffusion tensor imaging | gut microbiome | hypothalamus | inflammation | metabolomics | traumatic brain injury

ABSTRACT

Traumatic brain injury (TBI) is an insult to the brain that impacts neuronal and non-neuronal cells/tissues. The study aimed to understand TBI-induced early changes in the brain and systemic physiology. The male rats were subjected to mild and moderate TBI, where serum and urine metabolic fingerprints of mild TBI rats showed a hypermetabolic response with increased energy metabolites, amino acids, and gut metabolites in serum and increased TCA cycle intermediates in urine. In contrast, the moderate TBI rats showed decreased lactate, pyruvate, amino acids (glycine and leucine) and gut metabolites [trimethylamine N OXIDE (TMAO), choline and acetate] in serum. The urine showed increased pyruvate, creatinine, and allantoin levels. To understand the brain's role in altered metabolic physiology, hypothalamus structure was assessed using diffusion tensor imaging (DTI) and stress levels were observed using serum corticosterone. The injured rats exhibited changes in DTI metrics in the hypothalamus, suggesting a potential disruption in the regulation of the hypothalamus-pituitary–adrenal axis (HPA) axis. These alterations were accompanied by increased TNF-α levels after moderate TBI. The injury induced allostatic overload, accompanied by impaired hypothalamic structure, and metabolic physiology also showed gut microbiome dysbiosis. The gut microbiome showed an increased Firmicutes: Bacteroidetes ratio after injury, with variable gut composition after both injuries. Therefore, the present study provides insight into an interplay between the HPA axis, metabolism, and gut microbiome following TBI. Importantly, this crosstalk between the regulatory systems was different after mild and moderate injury, highlighting the need to assess injury phenotype based on the severity.

1 | Introduction

Traumatic brain injury-induced impact or contusion to the brain leads to irreversible brain damage that is followed by secondary injury cascades within minutes to hours to months post injury. Secondary injury progresses slowly and involves complex mechanisms including axonal degeneration, metabolic perturbation, mitochondrial dysfunction, oxidative stress, inflammation, glial activation, and neuronal death (Ng and Lee 2019). Importantly, the time window from a primary

© 2025 Wiley Periodicals LLC.

Summary

- TBI management is critical during the early time window post-traumatic event to aid in better diagnosis and treatment strategies.
- Until now, no research has comprehensively observed the changes in systemic physiology with a focus on injury severity (mild and moderate TBI).
- Therefore, the present study provides a snapshot of systemic metabolomics along with hypothalamic damage, systemic inflammation, and gut microbiome changes.

injury to the initiation of secondary injury events becomes important for the management of injury-associated disabilities. Fundamental critical care management involves various methods including mechanical ventilation, transfusion of required blood products, management of sympathetic hyperactivity, and nutritional therapy (Abdelmalik et al. 2019). Nutritional therapy is an integrated approach that optimizes systemic function and delivery of substrates to the brain. It is usually initiated as early as 24 h post injury to mitigate the effects of trauma. While nutritional supplementation suffers from issues like hyperglycemia, immunosuppression, and diminished gut integrity and function it becomes important to understand the basic metabolic and physiological mechanisms that initiate at early timepoints post injury (Kurtz and Rocha 2020).

Potential mechanisms that induce systemic alterations post TBI include the interplay of various pathways like the HPA axis and endocrine abnormalities that can further induce changes in systemic metabolism and inflammation (Tapp et al. 2019). Brain injury-induced stress can have variable effects on neurological and neurobehavioral functions. The hypothalamus plays an important role in adapting and responding to injury-associated stress and metabolic regulation under allostatic load (Zheng, Pang, et al. 2022). Our earlier study on closed head injury (CHI) has also revealed that an activated microglial population in the hypothalamic region with increased serum CORT after mild and moderate TBI is associated with a distinct metabolic and inflammatory phenotype in rats (Arora et al. 2022). While the hypothalamus is a deeply situated region, TBI-induced neuroendocrine dysfunction along with anti-hypothalamus antibodies have been observed in patients and boxers (Vijapur et al. 2020). The hypothalamus is an important region regulating energy homeostasis and metabolism in the body (Wen, Wang et al. 2019). In addition, the HPA axis is the primary responder during the states of internal and external stress to the body (Leistner and Menke 2020). While the structural changes in the hypothalamus along with its effects on HPA function are yet to be fully understood post TBI.

Since metabolites are an integral part of the biological processes and the intricate biochemical pathways that regulate the physiological system, any change in the body will reflect a unique metabolic fingerprint (Banoei et al. 2018). Metabolomics quantitatively assesses small molecules present, especially in biological fluids, including serum and urine. Since blood permeates every tissue and organ in the body, it effectively acts as a fluid network through which molecules secreted, excreted, or discarded by various tissues in response to physiological demands or stress are exchanged. Tissue damage, organ malfunctions, or pathological conditions can modify the chemical and protein composition of blood plasma or serum. Similarly, urine has been discovered to be the most convenient biological fluid for metabolic signatures related to systemic distress (Bykowski et al. 2021).

Nuclear magnetic resonance (NMR) spectroscopy is accepted as a reliable technique for metabolomic investigations due to its ability to identify novel metabolites without any chemical modification. Metabolomic fingerprinting has thus emerged as a valuable tool for identifying biomarkers and physiological indicators in various neurological conditions, including stroke, spinal cord injury (SCI), and sports-related concussions (Li and Sirko 2018). Recent studies show promise that additional research is essential to enhance our understanding of the pathophysiological processes associated with TBI severity and recovery. Extensively, TBI studies have found that the injury-associated metabolic phenotype gets repositioning or reprogramming. Injury leads to altered glucose metabolism and potentiates chronic metabolic syndrome (Li and Sirko 2018). Interestingly, a human study has also reported protective systemic lipid metabolism that can help in maintaining lipid homeostasis for optimal brain function after injury (Thomas et al. 2022). This underscores the importance of comprehending the impact of injury on systemic metabolism.

TBI-induced sympathetic activity and stress response can further progress into dysautonomia, dysbiosis, and systemic inflammation leading to gastrointestinal dysfunctions (Hanscom et al. 2021). The inflammatory shift post-injury causes gut dysmotility and increased mucosal permeability. These changes in the gut can lead to gut dysbiosis along with changes in the gut microbiome. Studies have shown that in animal models, TBI decreases smooth muscle contractility and delays gastrointestinal transit between 1 and 7 days after injury (Sun et al. 2015). Other studies have also shown microscopic changes in the intestine hours post fluid percussion injury (Feighery et al. 2008). Additionally, it has been observed that gut-derived systemic serotonin homeostasis and energy balance are disrupted post TBI (Mercado et al. 2022). Another study on TBI has also reported changes in the gut microbiome composition with altered tryptophan metabolism and microglial activation post injury day 7 and 28 (Zheng, Wang, et al. 2022). Despite this, as of now, no study has identified early changes in systemic metabolism in relation to microstructural alterations in the brain and gut microbiome after mild and moderate TBI.

Keeping in mind the above mentioned facts, the present study is aimed to assess TBI-induced changes within the early time window of 4–8 h post mild and moderate TBI. The earliest time window is considered an important phase of the injury wherein the brain is more vulnerable and susceptible to damage, with the ability of adaptive remodeling to restore homeostasis (Hoffe and Holahan 2022). The present study included metabolic phenotyping of serum and urine using ^{1}H-NMR metabolomics along with the assessment of neuronal structure that affects

10974547, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jnr.70032 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

metabolism (hypothalamus) using in vivo imaging. In addition, serum stress (serum CORT) and inflammation levels (IL-1β AND TNF-α) were quantified using sandwich ELISA kits. The gut microbiome composition was also observed using fecal 16S rRNA sequencing.

2 | Methods

2.1 | TBI Paradigm

2.1.1 | Animals

Twenty-four, 6–8 weeks male SD rats weighing between 200 and 250 g were obtained from the Experimental Animal Facility, INMAS. The rats were kept at a temperature of 23°C ± 1°C with a 12-h light/dark cycle (lights on 08:00 a.m.), and provided with a standard pellet diet (M/S Golden Feeds, Meharuli, Delhi, India) and water *ad libitum*. All experimental protocols performed were as per the approved guidelines by the institutional animal ethical committee (8/GO/RBi/S/99/CPCSEA/INM/IAEC/2021/02). The injury protocol and blood collection were done in the light phase, specifically between 09:00 a.m. and 01:00 p.m. (Qian et al. 2012).

2.1.2 | Injury Protocol

The animals were acclimatized for seven days and were randomly divided into three groups: Control, Mild TBI, and Moderate TBI, with each group having 8 rats. The rats were subjected to TBI by modifying Marmarou's weight drop model. Briefly, the rats for mild and moderate TBI were kept on a foam bed and were subjected to a free-falling brass rod (450 g inside an acrylic tube of 1 m held in place by an iron rod) from a height of 25 and 50 cm, respectively, above the sagittal midway (medial-lateral axis, relative to the intersection of bregma-lambda) of the brain (Arora et al. 2022). The control group received the same treatment except for the weight drop injury.

After 4–8 h post-injury, rats were subjected to MR imaging (anatomical and microstructural changes), systemic metabolomics analysis (serum and urine), and 16S rRNA sequencing for fecal microbiome composition (Figure 1). To study the difference between serum and urine metabolites, other serum markers, and DTI scalars in the control and injured groups, we assumed that there is no difference between these parameters (Ho = 0). Alternatively, if a difference was observed at $\alpha = 0.05$ and 80% power of the test with a priori effect size (f) of 1, the null hypothesis was rejected. Under these assumptions, the minimum number required for the study was 6 per group.

2.2 | In Vivo Imaging

To understand the microstructural alterations in the hypothalamus, MR imaging was performed at 4 h post-injury. Briefly, the anesthetized rats were placed on a bed with a radiofrequency array coil covering their heads. The MR imaging was done using Bruker Biospec 7.0 Tesla (Avance III) with a 30-cm horizontal bore magnet (Bruker Biospin, Ettlingen, Germany) and the imaging components included a receiver (array coil of four-channel for rat brain) and a transmitter (linear coil). Isoflurane in a mixture with oxygen was delivered to the rats with maintenance at 1%–2% isoflurane until the scan time with a flow of 0.5–1 L/min.

2.2.1 | Brain Localization

A fast low-angle single shot (FLASH) sequence was used to obtain a three-slice scout image (axial, mid-sagittal, and coronal).

2.2.2 | T2 Weighted Images

A rapid acquisition with relaxation enhancement (RARE) sequence with the following parameters: Field of view, 4 × 4 cm; Matrix size, 256 × 256; Repetition time/Echo time (TR)/(TE) = 2000/13 ms; Slice thickness, 1 mm; Number of slices = 15 was used for imaging.

2.2.3 | DTI Images

A multiple-shot spin echo EPI sequence with the following parameters: Field-of-view, 4 × 4 cm; Matrix size, 128 × 128; TR/TE, 5000 ms/34.46 ms; Number of gradient encoding directions, 46; with b, 672 s mm^{-2}; Slice thickness, 1 mm; Number of slices, 15; Number of segments, 4; number of averages = 1; Total imaging time = 12 min 55 s was used.

2.2.4 | Data Processing

Image cropping and correction of distortion were performed for all the scans. The tensor field data was diagonalized by using the

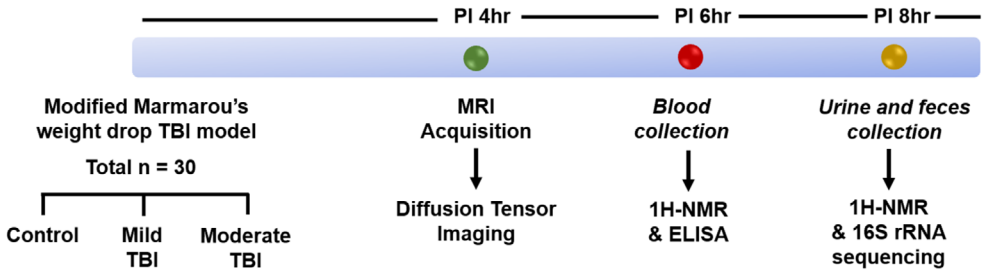


FIGURE 1 | Schematic representation of the study paradigm with the timepoint of sample collection. PI, post-injury.

10974547, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jnr.70032 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

analytic diagonalization method to obtain the eigenvalues ($\lambda 1$, $\lambda 2$, and $\lambda 3$) and the three orthonormal eigenvectors. DTI metrics including apparent diffusion coefficient (ADC), fractional anisotropy (FA), radial diffusivity (RD), and axial diffusivity (AD) for each voxel were computed using the eigenvalues. Regional FA, ADC, RD, and AD values were obtained by placing the bilateral regions of interest (ROIs; Figure 3) within the right and left sides of the hypothalamus. The size of ROIs and slices were kept constant for the analysis of each rat. For defining anatomical ROIs in the rat brain, Paxinos and Watson Rat Brain Atlas were used (Paxinos and Watson 2006). For analysis, in-house software was used, which provided the values of the DTI scalars, which were then further processed for statistical analysis (Trivedi et al. 2012).

2.3 | Sample Collection

2.3.1 | Serum

Blood (approximately 1 mL) was collected from the retro-orbital vein of each animal at 6 h post-injury in the serum vacutainers (SST Advance II BD Vacutainer) that were allowed to clot for 30 min at room temperature and then centrifuged at 2500 *g* for 15 min. The supernatant-containing serum was collected and stored at −80°C for further analysis.

2.3.2 | Urine and Feces

The rats were kept in metabolic cages as they gained consciousness for the collection of urine and fecal pellets. The samples were collected within 6–8 h post-injury and kept at −80°C for further analysis.

For NMR analysis, urine samples were centrifuged at 4722 *g* for 5 min at 4°C to remove particulate matter. The deuterated phosphate buffer solution (150 μL; 0.2 M Na_2HPO_4) (#7558-79-4, Sigma-Aldrich)/0.2 M NaH_2PO_4 (#10049-21-5, Sigma-Aldrich), pH 7.4 containing 1 mM Trimethylsilylpropanoic acid (TSP) (#7601-54-9, Eurisotop) prepared in D_2O (#7789-20-0, Sigma-Aldrich) was mixed with 450 μL of centrifuged urine to minimize variations in pH of urine samples.

2.4 | NMR Metabolomics

2.4.1 | Acquisition

Each serum (450 μL) was transferred to a 5 mm NMR tube containing 1 mM TSP in a closed capillary/insert (Wilmad, SP Industries, Buena, NJ, USA) as a reference. The urine sample (450 μL) was prepared in phosphate buffer with TSP, centrifuged, and transferred to an NMR tube. TSP in phosphate buffer acted as a chemical shift reference (0.0 ppm) and D_2O provided a signal lock. The NMR spectra were acquired at 800.20 MHz, Bruker-AV spectrometer (Bruker, Germany) at 300 K (Figure 2).

For serum NMR acquisition, water-suppressed Carr-Purcell-Meiboom-Gill (CPMG) spin echo pulse sequence [RD-90°-(τ-180°-τ) *n* acquire] with a total spin echo ($2n\tau$) of 200 ms was used to attenuate broad signals from proteins and lipoproteins. RD (relaxation delay) was 2.5 s, [RF (radio frequency) pulse], [τ (spin echo delay)], [*n* (number of loops)]. Typically, 128 scans were collected into 32 K data points over a spectral width of 12,820.51 Hz with an acquisition time of 9.29 s.

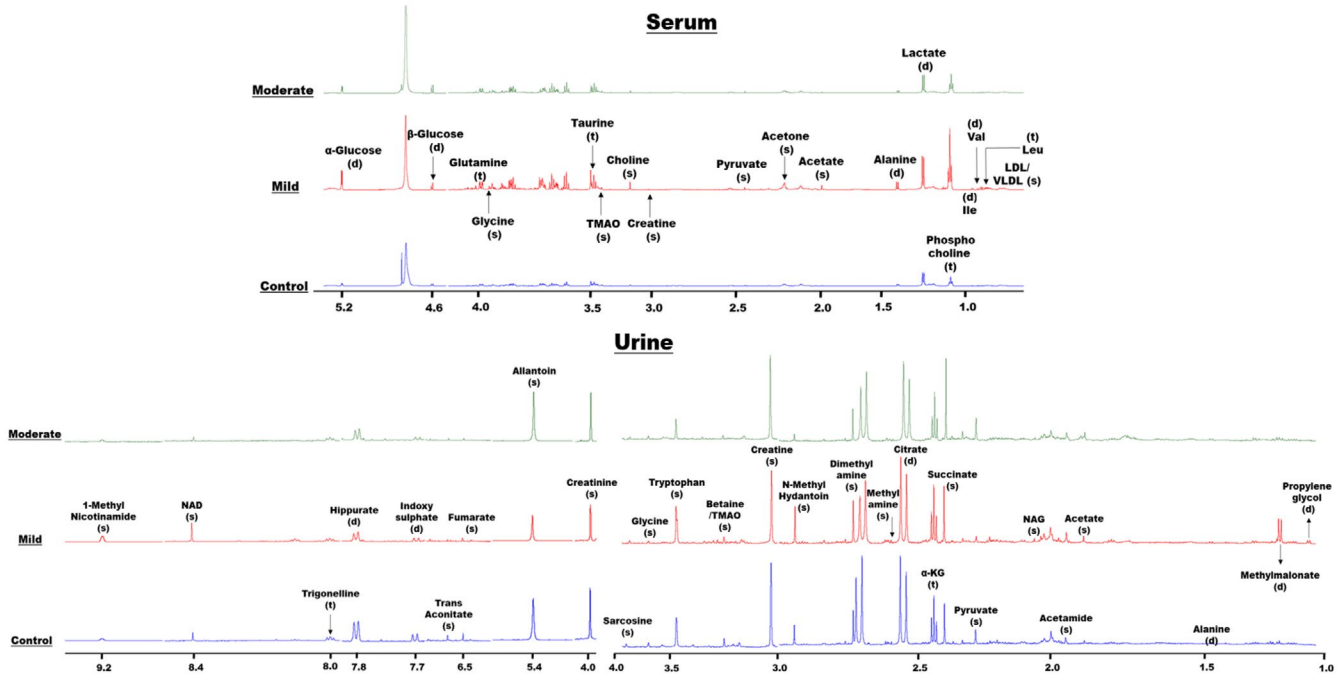


FIGURE 2 | Serum and urine spectra of ^{1}H-NMR in control, mild, and moderate groups. The spectra are labeled with the metabolites at their respective ppm. (s) Denotes singlet, (d) denotes doublet, and (t) denotes triplet. NAD, nicotinamide adenine dinucleotide; NAG, N-acetylglutamate; TMAO, trimethylamine N-oxide; α-KG, α-ketoglutarate.

10974547, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jnr.70032 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

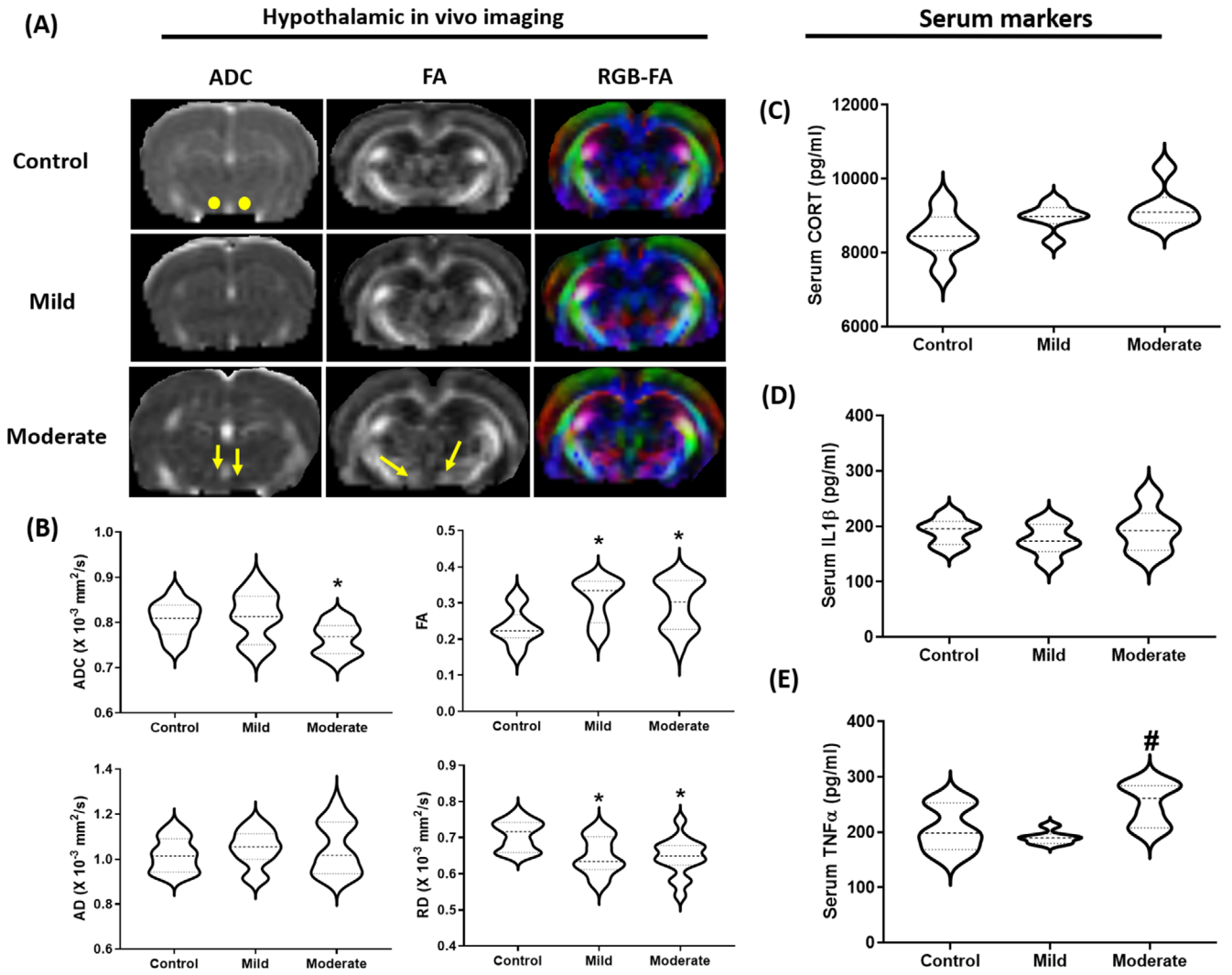


FIGURE 3 | Assessment of hypothalamic damage, systemic stress and inflammation hypothalamic apparent diffusion coefficient (ADC), fractional anisotropy (FA), and RGB-FA maps, the yellow circle shows ROI and yellow arrows depict changes in signal intensity (A) and violin plots indicating DTI scalars in all groups ($n = 7$) compared using one way ANOVA post hoc Holm's Sidak; ADC ($\times 10^{-3}\,mm^2/s$), FA, axial diffusivity (AD) ($\times 10^{-3}\,mm^2/s$) and radial diffusivity (RD) ($\times 10^{-3}\,mm^2/s$) values (B). The violin plots of serum markers of stress and inflammation in all groups ($n = 6$) were compared using one-way ANOVA post hoc Holm's Sidak; serum CORT (pg/mL) (C), serum IL-1β pg/mL (D) and serum TNF-α (pg/mL) (E) after injury. *Indicates difference versus the control group; #indicates difference versus the mild group.

For NMR acquisition of urine samples, the water signal in the samples was suppressed using a water pre-saturation pulse; Nuclear Overhauser Enhanced spectroscopy (NOESYPR1D) pulse sequence (RD-900-t-900-tm-900-acq) with relaxation delay (RD) of 5.0 s and an acquisition time of 2.5 s. 64 scans were collected into 32 K data points over a spectral width of 12,820.51 Hz.

2.4.2 | Data Processing

The serum and urine spectra were subjected to correction, including spectral phase, baseline, and chemical shift reference for each spectrum using TOPSPIN 3.7 (Bruker, Germany). Then the spectral region of 0.5–9.5 ppm was integrated, and each region was normalized with respect to the total spectral integral in TOPSPIN 3.7. The metabolites were identified using the human metabolome database (HMDB) (described in Table S1) and previous literature (Arora et al. 2022). The changes in the concentration of metabolites were compared between the control, mild, and moderate injury groups using a one-way ANOVA. The ratios were computed for serum αglucose: βglucose, lactate: pyruvate, and TMAO: choline. The relative intensity values were also analyzed in MetaboAnalyst 5.0 (https://www.metaboanalyst.ca) (Xia et al. 2009) where the data was scaled (auto-scaled and range scaled) and multivariate analysis was performed using principal component analysis (PCA) and partial least squares-discriminant analysis (PLS-DA). To identify metabolites involved in the segregation of clusters in PLS-DA, variable importance in projection (VIP) scores were used. The pathway analysis involved the selection of metabolic pathways with impact > 0.2 and $-\log 10(p) > 2$.

2.5 | Serum CORT Levels

The levels of CORT in the serum were quantified using the competitive ELISA kit (# EIACORT, Invitrogen) as described

10974547, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jnr.70032 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

previously (Arora et al. 2022). The serum samples (5 μL) were thoroughly mixed and vortexed with dissociation reagent (5 μL) and then diluted with 490 μL of 1× assay buffer from the kit. On the other hand, standards provided in the kit were serially diluted (ranging from 78.125–10,000 pg/mL) according to the manufacturers' instructions and then 50 μL were added to the well plates. The prepared serum samples and standards were then pipetted into the wells, with the addition of 1× assay buffer (75 μL), CORT conjugate (50 μL), and CORT antibody (50 μL).

The plate was kept at room temperature on a plate shaker at 400 rpm for one hour of incubation. After aspiration of the solution, the wells were washed thoroughly with 1× wash buffer from the kit for 4 times. Then, 100 μL of (3,3′, 5,5″-tetramethylbenzidine) TMB substrate was added to each well and incubated for 30 min without shaking. The reaction was terminated using a stop solution (50 μL) and end-point absorbance was read at 450 nm using a spectrophotometer (SpectraMax M2, Molecular Devices). The serum CORT levels were calculated using the standard curve and expressed in pg/mL.

2.6 | Inflammatory Markers

IL-1β (#BMS630, Invitrogen) and TNF-α (#BMS622, Invitrogen) levels were measured in serum by their respective ELISA kits. Firstly, the microwell strips were washed two times with the (1×) wash buffer. Then, diluted serum samples (50 μL) with an equal volume of sample diluents were added to the wells with (1:100) biotin-conjugate solution (50 μL) and incubated for 2 h at 25°C. For the standard curve, a serial dilution of samples ranging (TNF-α: 39.1–2500 pg/mL; IL-1β: 31.3–2000 pg/mL) was prepared and 50 μL of each was added to the wells, with (1:100) biotin-conjugate solution (50 μL) and incubated for 2 h at 25°C. After aspiration, the wells were thoroughly washed using 1× wash buffer 3–4 times. Following one hour of incubation with (1:200) streptavidin-HRP solution (100 μL), color development was observed with TMB substrate (100 μL), and the absorbance was recorded at 450 nm using a spectrophotometer (SpectraMax M2, Molecular Devices). The biotin-conjugate and HRP-streptavidin were diluted in the 1× assay buffer provided in the kit. The concentration of the cytokines was calculated using the standard curve and expressed in pg/mL.

2.7 | Fecal 16S rRNA Sequencing

2.7.1 | DNA Isolation

The genomic DNA of all the samples was isolated using DNeasy Power Soil Pro Kit (Qiagen). The purity of DNA was checked using NanoDrop 2000 (Thermofisher Scientific, Massachusetts, USA) and the integrity of DNA was evaluated on 1% (w/v) agarose (Lonza, Belgium) gel.

2.7.2 | 16S rDNA V3-V4 Metagenomics Amplification

A total volume of 12.5 ng DNA from each sample was subjected to 16S V3-V4 PCR with the primers described below. The PCR products were then loaded on a 1.5% (w/v) agarose gel along with a 100 bp ladder. Illumina adapter overhang nucleotide sequences were added to the gene-specific sequences (gene-specific sequences target the 16S V3 and V4 region).

The full-length primer sequences, using standard IUPAC nucleotide nomenclature (N = A/T/G/C, W = A/T, H = A/C/T, V = A/C/G), targeting the V3–V4 region:

16S Amplicon PCR Forward Primer: 5′TCGTCGGCAGCGTC AGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG3′ 16S Amplicon PCR Reverse Primer: 5′GTCTCGTGGGC TCG GAG ATG TGT ATA AGA GAC AGG ACT ACH VGG GTATCTAATCC3′.

After confirming the amplification on gels, 5 μL of PCR reaction were pooled for each group; control, mild, and moderate injury. The pooled PCR products were bead purified and eluted and 5 μL of eluate was subjected to another round of PCR with dual indices and adapters to generate the respective libraries as recommended by Illumina. To determine the size range of the fragments and the average library size the cleaned libraries were quantitated on a Qubit Fluorometer and appropriate dilutions were loaded on a D1000 screen tape.

2.7.3 | Data Analysis

All the analyses were performed in R version 4.2.2. The package dada2 (Callahan et al. 2016) was used to read QC, read filtering, denoising, read merging, chimera removal, and taxonomy assignment using the Silva database. A mean quality score of more than 30 was considered high quality. Multiple sequence alignment of representative sequences was done using the R package DECIPHER (Wright 2016). The bioinformatics workflow included demultiplexing of the raw sequences, which were then denoised to get the representative sequences for taxonomic classification and sequence alignment for phylogenetic studies. This data was then used for the assessment of differential abundance and diversity analysis. EdgeR, v4.2.1 (Robinson et al. 2010) an R package that uses negative binomial distribution, was used for comparative analysis. The comparisons were made based on the abundance of OTU between different sample groups.

2.8 | Statistical Analysis

All values are expressed as mean ± standard deviation (SD). The normality of the data was checked using the Shapiro–Wilk Test and then one-way analysis of variance (ANOVA) was done using pairwise comparison (Holm's Sidak) in IBM SPSS statistical software version 20 (IBM Corp., Armonk, NY, USA) and graphs were made in GraphPad version 8 (GraphPad Software Inc., California). For metabolomics data, the metabolite levels were normalized using auto and range scaling and then processed with one-way ANOVA and multivariate analysis in Metaboanalyst 5.0 (https://www.metaboanalyst.ca). (Xia et al. 2009) The fecal sequencing data analysis was performed using R software. Since the groups were pooled prior to sequencing, no statistical analysis was performed on abundance data. While EdgeR was used to find statistically significant altered genera within the groups. Values with $p < 0.05$ were considered significant. The outliers were identified

10974547, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jnr.70032 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

using adjusted box plots in R software, yielding no outliers in many cases, and if present, were not removed due to lower sample size.

3 | Results

3.1 | In Vivo Imaging Post TBI

On T2 weighted images, the hypothalamic region showed no hyperintensity and the parenchyma appeared normal after both mild and moderate TBI (Figure S1). The hypothalamic region showed changes in the DTI scalars ($n = 7$) (values including right and left hypothalamus) with a significant increase in FA [One-way ANOVA; $F_{(2, 39)} = 6.07$, $p = 0.005$] after mild ($p = 0.007$) and moderate ($p = 0.032$) injury. The RD values were significantly decreased after both injuries [One-way ANOVA; $F_{(2, 39)} = 6.54$, $p = 0.004$]; mild ($p = 0.013$) and moderate ($p = 0.007$) injury. No significant alterations were observed in AD values [One-way ANOVA; $F_{(2, 39)} = 0.54$, $p = 0.585$] after mild ($p = 0.773$) and moderate ($p = 0.744$) injury. The rats showed a significant decrease in the ADC value [One-way ANOVA; $F_{(2, 39)} = 5.02$, $p = 0.011$] in moderately injured rats ($p = 0.042$) with no change after mild injury ($p = 0.962$) (Figure 3).

3.2 | Serum CORT and Inflammatory Cytokine Levels Post TBI

Blunt impact on the brain did not significantly change the serum CORT levels ($n = 6$, each) [One-way ANOVA; $F_{(2, 15)} = 3.06$, $p = 0.077$] post mild ($p = 0.348$) and moderate ($p = 0.081$) TBI. The pro-inflammatory cytokines, IL-1β [One-way ANOVA; $F_{(2, 15)} = 0.57$, $p = 0.575$] showed no change post mild ($p = 0.789$) and moderate TBI ($p = 0.999$), while TNF-α levels were significantly upregulated [One-way ANOVA; $F_{(2, 15)} = 5.35$, $p = 0.018$] post moderate TBI ($p = 0.020$) as compared to the mild group (Figure 3).

3.3 | Metabolic Fingerprint Using ^{1}H-NMR Following TBI

Total 18 metabolites were integrated with serum ($n = 6$–7) that included metabolites from the energy metabolism (α: β glucose, lactate: pyruvate ratio, creatine), amino acids (glutamine, glycine, alanine, leucine, lysine, valine, Taurine), gut metabolism (TMAO: choline, acetate), ketone bodies (acetone), lipids (LDL/VLDL) and others (phosphocholine).

Total 27 metabolites were integrated in urine ($n = 6$–7), which included energy metabolites (fumarate, creatine, α-ketoglutarate, citrate, succinate, pyruvate, methylmalonate), amino acids (glycine, tryptophan, alanine), amino acid derivates (sarcosine, creatinine, N-acetyl glutamate (NAG)), gut (indoxyl sulfate, dimethylamine, methylamine, acetate, N-methyl hydantoin), purine metabolites (nicotinamide adenine dinucleotide (NAD), and 1-methyl nictoninamide) and others (trigonelline, hippurate, trans aconitate, allantoin, betaine, acetamide, propylene glycol).

3.3.1 | Analysis of Metabolite Alterations

The serum and urine metabolite intensities were compared using one-way ANOVA Table S2 and the heat map of the metabolites altered in both serum and urine is depicted in Figure 4.

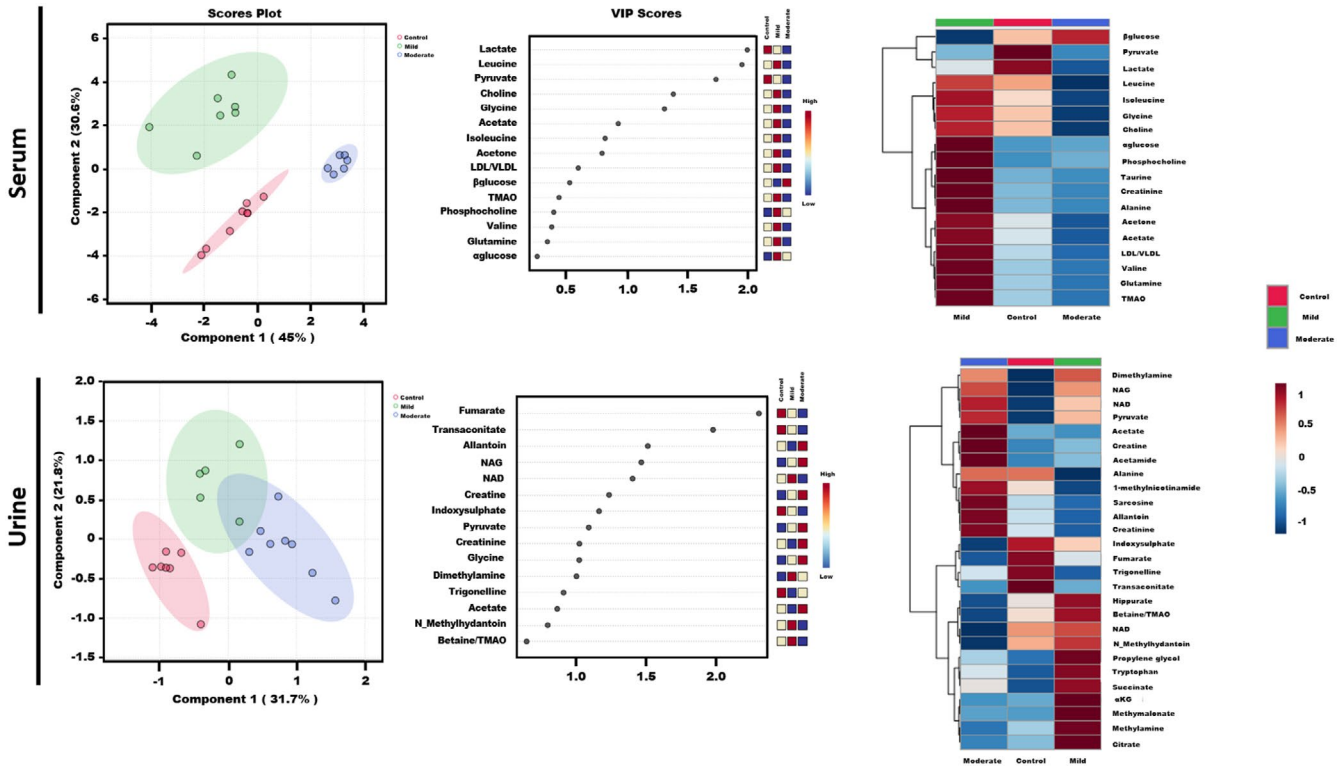


FIGURE 4 | Serum and urine metabolomics PLS-DA score plots and VIP scores of serum ($n = 6$–7) and urine ($n = 6$–7) metabolites. Metabolites with VIP > 1 were considered. The heat map shows changes in the metabolites observed in serum and urine for each group.

10974547, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jnr.70032 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The serum of mild TBI rats showed increased energy metabolites: α: β glucose ratio Table S2 and creatine intensity. The metabolic intensity of amino acids, including glutamine, alanine, valine, and taurine, was upregulated after mild TBI. The gut metabolites TMAO: choline, acetate, and ketone body (acetone) were also increased. LDL/VLDL and phosphocholine showed an increase after mild TBI. In contrast to mild injury, the serum of moderate TBI rats showed decreased lactate, pyruvate, amino acids (glycine, leucine), and gut metabolites (choline, TMAO, and acetate) Figure S2.

The urine of mild TBI rats showed an increase in energy metabolites (α-ketoglutarate, citrate, succinate, methylmalonate) accompanied by a decrease in fumarate levels. The amino acid, alanine, was decreased while the amino acid derivative, NAG, was increased. The gut metabolites, dimethylamine and methylamine, showed elevation. Trigonelline and trans aconitate decreased, while propylene glycol was increased. On the other hand, moderate TBI led to an increase in pyruvate along with a decrease in fumarate in urine. The creatinine intensity was also elevated after moderate TBI. The gut metabolites, DMA, were increased, accompanied by decreased indoxyl sulfate and N-methyl hydantoin intensities. Allantoin, an oxidative stress marker, was also increased after moderate TBI. Moderate TBI rats showed increased glycine, NAD, and NAG metabolic intensities with decreased trigonelline and trans aconitate in the urine.

3.3.2 | Multivariate Analysis

The serum PCA and PLS-DA score plots showed a complete segregation within the groups. The PLS-DA accuracy (100%) was determined by R2 (90%) and Q2 (60%). The PLS-DA score plot was calculated to identify the metabolites that had a VIP score >1 and helped in the segregation of the groups. The metabolites with VIP >1 included glycine, choline, pyruvate, leucine, and lactate.

The urine PCA score plots showed complete segregation within the groups, while the PLS-DA plot had some overlap between mild and moderate TBI groups. The accuracy (100%) was determined by R2 (90%) and Q2 (70%) values. The metabolites with VIP >1 included fumarate, trans aconitate, allantoin, NAG, NAD, creatine, Indoxyl sulfate, pyruvate, creatinine, glycine, and dimethylamine (Figure 4). The PCA plots of both serum and urine metabolites are depicted in Figure S3.

3.3.3 | Pathway Analysis

The pathway analysis was performed between control vs. mild TBI and control vs. moderate TBI groups. The selection of metabolic pathways included pathways with impact >0.2 and $-\log 10(p) >2$ (Table S3). The serum of both injuries showed changes in glycine, serine, threonine, and pyruvate metabolism. Mild TBI also resulted in changes in taurine and hypotaurine metabolism. The urine of both mild and moderate TBI rats showed an altered TCA cycle. In addition, the moderate TBI led to changes in glycine, serine, and threonine metabolism, pyruvate metabolism, and nicotinate and nicotinamide metabolism.

3.4 | Fecal Microbiome Post TBI

The gut microbiome composition analyzed using 16S rRNA sequencing showed changes following both mild and moderate TBI. The α diversity was expressed as Chao1, Ace, se.ACE, Shannon, Simpson, InvSimpson, and Fisher indexes. The β diversity was evaluated using ordination analysis. The α- diversity showed a decrease in richness within the mild TBI group, while the moderate TBI group showed no change. The β diversity has also deviated from the control group in injured rats. At the phylum level, the relative composition of Firmicutes in the control, mild, and moderate TBI groups was 38.4%, 46.8%, and 43.9%; while the Bacteroidetes composition was 58.8%, 51.7%, and 52.6%. However, the Firmicutes: Bacteroidetes ratio was increased after both mild (0.91) and moderate (0.84) injury. Other bacterial phyla including Spirochaetes, Proteobacteria, Cyanobacteria, and Epsilonbacteraeota were also altered after mild (1.4%) and moderate (3.3%) TBI, with the control group having a relative composition of 2.7%.

The percentage bacterial composition at the genus level in the pooled control group was Prevotella_9 (61.7%), Faecalibacterium (11.3%), Lactobacillus (11.0%), Ruminococcaceae_UCG-008 (6.9%) and Others (8.8%). The composition of Prevotella_9 (48.8%), Faecalibacterium (15.1%), Lactobacillus (8.7%), Ruminococcaceae_UCG-008 (19.4%), and Others (7.7%) in the mild group was altered. The bacterial composition in the moderate group included Prevotella_9 (49.8%), Faecalibacterium (14.7%), Lactobacillus (9.6%), Ruminococcaceae_UCG-008 (14.9%) and Others (11.0%) (Figure 5, Table S4). The EdgeR analysis showed significant alterations in the genus *Prevotella_9* (Bacteroidetes phylum). From the Firmicutes phylum, *Enterococcus, Blautia, Lactobacillus, Anaerovibrio, Dialister, Faecalibacterium, Subdoligranulum*, and *Treponema_2* (Spirochetes phylum) were altered. All microbes were decreased except *Lactobacillus* (Table 1).

4 | Discussion

TBI leads to bidirectional, complex, and multifaceted alteration in the systemic physiology, which has been attributed to potential mechanisms including metabolic impairment, GI tract abnormalities, and persistent inflammatory response (McDonald et al. 2020). The systemic metabolic changes are reflected in serum and urine metabolites, which can serve dual roles as signals and fuel for the gut microbiome. Conversely, metabolites originating from the gut can impact the immune, metabolic, and nervous systems (Cox et al. 2022). Therefore, to understand the TBI-induced complex interplay between the brain's metabolic regulator, hypothalamus, systemic metabolism, and gut microbiome, the present study was undertaken. The results obtained highlight an important aspect of translational studies in the diagnosis of TBI.

10974547, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jnr.70032 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

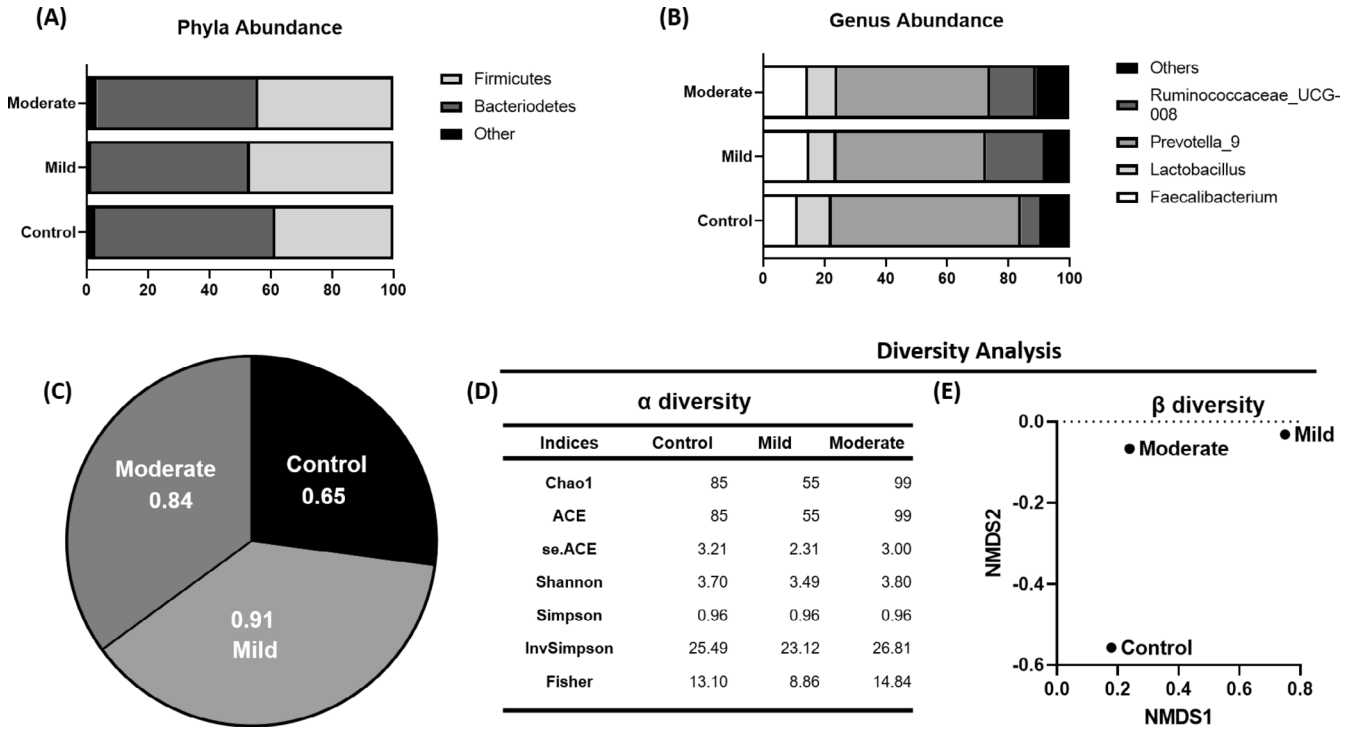


FIGURE 5 | Gut microbiome assessment the bar graph shows the relative percentage abundance of bacteria in all groups at the phylum (A) and genus level (B). The abundant phyla Firmicutes and Bacteroidetes ratio was calculated for each group shown in the pie chart (C). For diversity analysis, a table of α-diversity indices (D) and a scatter plot of β-diversity ordination values (E) in all groups. The α-diversity represents seven indices. The β-diversity was calculated using ordination analysis. All graphs shown here were not compared statistically, as the fecal samples were pooled for each group before 16S rRNA sequencing.

4.1 | Serum and Urine Metabolic Phenotypes Following TBI

Serum metabolomics yielded a variable shift in the metabolites post mild and moderate TBI. The mild TBI rats showed an increase in energy metabolites including α: β glucose ratio, (branched-chain amino acids (BCAAs), valine and leucine), glutamine, alanine and taurine, and creatine. This increase could be due to alterations in blood flow and cerebral ischemia-associated hypermetabolic conditions (Diringer et al. 2011). Increased α: β glucose ratio with glucogenic amino acids (glutamine, alanine, and valine) as a hyperacute response could be due to increased glycogen breakdown or gluconeogenesis as seen earlier in our study (Arora et al. 2022). On the other hand, increased acetone is a sign of ketosis (Manninen 2004). Therefore, both gluconeogenesis and ketosis may serve as a hyperacute injury response to maintain energy homeostasis after mild injury. Importantly, the increased taurine levels could also help in repairing tissue damage and balancing energy metabolism after TBI. As reported earlier, taurine elevation is seen after stroke and its decrease is postulated as a biomarker for stroke recovery (Wen, Li et al. 2019; Ghandforoush-Sattari et al. 2011). Glutamine has been earlier reported to promote recovery and immune modulation, resulting in an effect on the health of the intestinal lining and the balance of the gut microbiome following neuropsychiatric conditions (Kim and Kim 2017). Consequently, elevated glutamine levels in the present study suggest its immunomodulatory influence on the gut microbiome, potentially altering gut homeostasis. Corroboratively, the gut metabolites, TMAO: choline ratio, and acetate were also upregulated in mild TBI rats. Opposed to the bad effects of increased serum TMAO, the increased TMAO can also be associated with elevated glucose levels which can improve glucose tolerance (Krueger et al. 2021); (Velasquez et al. 2016). The elevation in acetate levels can be linked to the inhibition of insulin release, anti-inflammatory actions, and protein acetylation (Moffett et al. 2020). On the other hand, moderately injured TBI rats showed decreased lactate, pyruvate, leucine, glycine, TMAO, choline, and acetate intensities, indicating low levels of preferred energy substrates post injury for early response to energy requirements (Brooks and Martin 2015). The reduced BCAA is similar to the trend observed in TBI patients after severe TBI and other preclinical studies (Arora et al. 2022). While the decreased energy substrates along with glycine reduction are similar to cerebral ischemic conditions (Sidorov et al. 2019). Therefore, both injuries show unique metabolic phenotypes suggesting variable responses to the systemic effects that are dependent upon the severity of the injury.

The urine metabolic phenotype of the injured rats also suggests changes in energy metabolism, where mildly injured rats show an increase in energy metabolites: methylmalonate, α- ketoglutarate, citrate, and succinate, which suggests their increased synthesis for important functions they contribute to in the kidney, including antioxidation, immunomodulation, and gluconeogenesis, respectively. Since kidney function is directly proportional to mitochondrial activity and thereby the

10974547, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jnr.70032 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 | Genus of the fecal microbiota significantly altered within the groups: Control, mild, and moderate using EdgeR analysis.

Kingdom	Phylum	Class	Order	Family	Genus	logFC	logCPM	*p*	FDR
Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	Prevotella_9	−7.145	13.593	<0.001	<0.001
Bacteria	Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	Enterococcus	−7.223	12.044	<0.001	<0.001
Bacteria	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Blautia	−6.022	11.911	<0.001	<0.001
Bacteria	Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	Lactobacillus	5.037	10.839	<0.001	<0.001
Bacteria	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Anaerovibrio	−5.769	8.843	<0.001	<0.001
Bacteria	Firmicutes	Negativicutes	Selenomonadales	Veillonellaceae	Dialister	−5.645	11.225	<0.001	<0.001
Bacteria	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Faecalibacterium	−5.585	10.242	<0.001	<0.001
Bacteria	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Subdoligranulum	−5.043	11.221	<0.001	<0.001
Bacteria	Spirochaetes	Spirochaetia	Spirochaetales	Spirochaetaceae	Treponema_2	−4.999	10.973	<0.001	0.001

Krebs cycle, the TCA intermediates have a central role in the diagnosis of kidney dysfunctions. The elevated TCA metabolites, α- ketoglutarate and succinate, have also been found in depressive disorders, making it essential to study urine metabolomics for diagnosis (Jiménez-Uribe et al. 2021). Propylene glycol was also increased after mild injury, which is an important contributor of lactate, acetate, and pyruvate (Lim et al. 2014). The amino acid alanine was decreased after mild injury. The gut metabolites dimethylamine and methylamine were elevated in the urine, while there was a decrease in the urinary trigonelline and trans-aconitate. The oxidative stress marker allantoin was also decreased in the urine of mild TBI rats. On the other hand, the urine of moderate TBI rats showed an increase in pyruvate, creatinine, allantoin, NAD, and N-methyl hydantoin, suggesting oxidative stress and impaired energy metabolism, impaired nicotinate and nicotinamide metabolism. These findings are distinctive of TBI pathology and have been reported earlier in urine samples of TBI patients (Bykowski et al. 2021). The urine of moderate TBI rats also showed a decrease in urine indoxyl sulfate, suggesting impaired gut barrier maintenance (Gisewhite et al. 2021); (Vernocchi et al. 2020). Therefore, the urine of both injuries showed an impaired TCA cycle with increased metabolites after mild injury and decreased ones after moderate injury. The gut metabolite DMA secretion was also increased after both injuries, which could be due to proteolysis from asymmetric dimethylarginine, which is involved in the L-arginine/nitric oxide pathway and is present in several cardiovascular, renal, and metabolic diseases (Tsikas 2020). Similar to DMA, elevated MA is present during CNS disturbances with increased BBB permeability leading to urinary secretion of amines (Mitchell and Zhang 2001). In conclusion, the urinary metabolic profile indicates a compromised TCA cycle following both injuries, accompanied by heightened oxidative stress and disruptions in gut and NAD metabolism after moderate injury.

4.2 | Alterations in the Hypothalamic Structure Along with Response of Serum Markers After TBI

The microstructural changes in the hypothalamus of the injured rats using DTI showed that mild TBI leads to an increase in FA and a decrease in RD values. On the other hand, moderate TBI rats showed decreased diffusivity with decreased ADC and RD values and increased FA in the hypothalamus of the rats. The increased FA is similar to the anisotropic changes seen after acute cerebral ischemia (Alegiani et al. 2017). TBI-induced hypoxic insult, intracranial pressure, and axonal injury can lead to hypothalamic damage as earlier seen in boxers, which reported autoantibodies against the hypothalamus and pituitary post-concussion (Kelestimur 2021). Similarly, a study on pediatric TBI found a decrease in the brain volume of the corpus callosum and hypothalamus (Dennis et al. 2017). The serum CORT and inflammatory cytokine levels were also quantitated, which revealed no change, except a significant increase in the TNF-α levels after moderate injury as compared to the mild group. Importantly, TBI may alter the neural pathways responsible for transmitting stress signals to the HPA axis, slowing the overall response time for CORT secretion. Additionally, the damage to the hypothalamus (as

10974547, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jnr.70032 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

also reported in the present study) can impair the communication and coordination between the hypothalamus, pituitary gland, and adrenal glands, which could have delayed the activation of CORT release. As reported earlier, physiological and psychological stressors initiate a stress response involving the sympathetic adrenal medullary (SAM) axis and the HPA axis. The SAM axis mediates the immediate stress response, while the HPA axis drives a slower, more sustained response, where stress-induced GC secretion can be delayed (Lawrence and Scofield 2024) Additionally, our earlier study has reported increased CORT from acute till chronic timepoint post mild and moderate TBI, confirming delayed but sustained CORT secretion post TBI (Arora et al. 2022).

A similar absence of elevated inflammatory markers at the hyperacute stage post both severities were observed. However, IL-1β is involved in the early response to neuroinflammation after TBI; the kinetics of IL-1β release and its detectable levels can vary depending on the time point after injury. The increase in IL-1β expression after TBI is typically observed in the acute to chronic phases, not during the hyperacute phase, which suggests that (Morganti-Kossmann et al. 2002) production of IL-1β takes time to develop after initial injury. Cytokines like IL-1β, IL-6, TNF-α, and MCP-1 are released both locally in the brain and systemically post TBI. Delayed systemic peaks in these cytokines often occur as the inflammatory response may expand from the CNS to peripheral organs, which may be dependent upon the progressive breakdown of the BBB, allowing inflammatory mediators and immune cells to cross into the systemic circulation (Alam et al. 2020) Therefore, the present results conclude hypothalamic alterations without CORT elevation and inflammation.

4.3 | Gut Microbiome Changes After TBI

It was found that Firmicutes, including short chain fatty acids (SCFAs) producing bacteria (*Faecalibacterium*, Subdoligranulum, and *Blautia*, butyrate-producing Firmicutes) were decreased. *Faecalibacterium*, a biomarker of human gut health, is reported to decrease in diabetic patients (Ferreira-Halder et al. 2017). Along with this, *Faecalibacterium* decrease is associated with pain, sleep problems, and neurological dysfunction (Taraskina et al. 2022). The decrease in *Subdoligranulum* has been found in various metabolic disorders, where its alteration is associated with microbial richness and HDL cholesterol levels (Van Hul et al. 2020). The decrease in *Blautia* strain is also similar to gut microbiome alterations observed after stroke (Sudo 2014). The increased *Lactobacillus* is similar to that seen after acute ischemic stroke, suggesting loss of good gut health (Sudo 2014). Both mild and moderate TBI groups showed altered *Enterococcus* composition, as earlier reported after TBI and stroke (Taraskina et al. 2022). The decreased *Prevotella* composition could be an adaptive response to prevent gut immune activation and maintain intestinal permeability or could be an injury-induced effect on the *Prevotella* abundance. In conclusion, the mild and moderate TBI groups showed gut microbiota changes similar to TBI and stroke pathology, with increased Firmicutes: Bacteroidetes ratio after both injuries indicating gut dysbiosis (Taraskina et al. 2022; Rosin et al. 2021).

4.4 | Interplay Between the Hypothalamus, Systemic Metabolism and Gut Microbiome

The present study directs the potential bidirectional cross-talk between the brain and systemic physiology. The hypothalamic damage was observed to be more pronounced in the moderately injured rats, with ADC significantly decreased along with RD, indicating cellular changes in the hypothalamus. Along with this, moderate rats had impaired energy metabolism, inflammation, and oxidative stress at the systemic level. Mildly injured rats, on the other hand, showed an anisotropic increase and RD decrease (due to edema) in the hypothalamus, which also initiates a cascade of metabolic changes at the systemic level, which can consequently induce several processes involved in glucose production.

At the level of the gut, it was found that Firmicutes (involved in the synthesis of SCFAs and trimethylamine (TMA), which is converted to TMAO in the liver) were altered after both injuries. Mildly injured rats showed an increased TMAO: choline ratio with an increased Firmicutes: Bacteroidetes ratio (Liu and Dai 2020). Importantly, the TMA-TMAO axis plays an important role in the systemic metabolism and gut microbiome association, as it has been found that fecal microbiota transplantation resulted in a decrease of TMA levels in feces and TMAO levels in the brain and serum after injury (Du et al. 2021). The microbiome richness (decreased α diversity and altered β diversity) after mild injury with increased TMAO production suggests an important and compensatory role of TMAO as compared to its deleterious effects reported earlier. On the contrary, the moderate TBI rats with increased Firmicutes: Bacteroidetes ratio showed decreased gut metabolism in serum (acetate, choline, and TMAO) and urine (N-methyl hydantoin and Indoxyl sulfate), oxidative stress (allantoin in urine), and altered gut metabolites (Stojanov et al. 2020). Additionally, decreased acetate levels in the serum of moderate TBI rats could potentially affect the gut microbiome composition and brain function, as it has been earlier reported that the administration of acetate improved spatial learning in the TBI-induced rats (Opeyemi et al. 2021). Thus, highlighting that moderately injured rats with more damage and decreased metabolism can further lead to impaired gut microbiome composition. Meanwhile, the mildly injured rats also showed gut dysbiosis with increased gut metabolites, indicating an important relationship between systemic metabolism and gut health.

Therefore, this study helped us to understand an integration between systemic metabolism and gut microbiome composition, where hypothalamic damage can act as an additional arm in perpetuating the injury damage leading to more devastating effects with increased severity.

5 | Limitations

1. As females exhibit higher basal and stress-induced ACTH and CORT levels (Goel et al. 2014), and the neuroprotective effects of estrogen and progesterone influence HPA axis signaling, inflammation, cerebral blood flow, and metabolism (Liu et al. 2012), the present study used male rats to avoid confounding interactions with gonadal hormones.

10974547, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jnr.70032 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2. To better understand the interaction between CORT and the gut, a detailed morphological assessment of the intestinal lining and permeability is warranted (Leigh et al. 2023).

6 | Conclusions

The present study observed hyperacute systemic changes in mild and moderate TBI along with the underlying hypothalamic damage and gut microbiome alterations. Mild TBI showed increased glucose, amino acid, and TMAO: choline ratio. The urine metabolomics of mild-TBI rats also suggested increased levels of energy metabolites from renal gluconeogenesis. On the other hand, moderate-TBI rats showed a reduction in the energy metabolites in serum with signs of oxidative stress in urine. The rats also showed microstructural damage in the hypothalamus. The gut microbiome had an increased Firmicutes/Bacteroidetes ratio after injury, with decreased α diversity after mild TBI and altered β diversity after both injuries. Thus, the study demonstrates that both mild and moderate TBI leads to systemic metabolic alterations along with hypothalamic damage and gut dysbiosis at a hyperacute timepoint. The study holds translational significance as the animal model used closely replicates human TBI and involves different severities to understand the differential response. Additionally, the application of noninvasive techniques for hyperacute timepoint assessment could be extended to diagnostic and prognostic purposes through further investigations involving larger sample sizes. Considering the variation in the gut microbiome composition between rats and humans, additional studies are still required with the assessment of gut integrity.

Author Contributions

Palkin Arora, Megha Kumari, Kavita Singh, M Memti Devi, Richa Trivedi, Rajat Sandhir and Poonam Rana: review and editing (equal). **Richa Trivedi, Palkin Arora:** conceptualization (equal); **Palkin Arora:** writing – original draft (lead); formal analysis (lead); writing – review and editing (lead). **Palkin Arora, Megha Kumari, Kavita Singh:** data curation (equal); writing – review and editing (equal). **Palkin Arora, Megha Kumari, Kavita Singh:** methodology (lead); writing – review and editing (equal). **Kavita Singh, M Memti Devi, Richa Trivedi, Rajat Sandhir and Poonam Rana:** conceptualization (supporting); Writing – original draft (supporting); Writing – review and editing (equal).

Acknowledgments

The authors would like to acknowledge Nucleome Informatics Private Limited for the fecal 16S rRNA sequencing.

Ethics Statement

All applicable international, national, and/or institutional guidelines for the care and use of animals were followed and are approved by the INMAS research ethics committee (8/GO/RBi/S/99/CPCSEA/INM/IAEC/2021/02). This article does not contain any studies with human participants performed by any of the authors.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The datasets generated during and/or analyzed during the current study are not publicly available but are available from the corresponding author on reasonable request.

Peer Review

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/jnr.70032.

References

Abdelmalik, P. A., N. Draghic, and G. S. F. Ling. 2019. "Management of Moderate and Severe Traumatic Brain Injury." *Transfusion* 59: 1529–1538. https://doi.org/10.1111/trf.15171.

Alam, A., E. P. Thelin, T. Tajsic, et al. 2020. "Cellular Infiltration in Traumatic Brain Injury." *Journal of Neuroinflammation* 17: 328. https://doi.org/10.1186/s12974-020-02005-x.

Alegiani, A. C., S. MacLean, H. Braass, et al. 2017. "Comprehensive Analysis of Early Fractional Anisotropy Changes in Acute Ischemic Stroke." *PLoS One* 12: e0188318. https://doi.org/10.1371/journal.pone.0188318.

Arora, P., K. Singh, M. Kumari, and R. Trivedi. 2022. "Temporal Profile of Serum Metabolites and Inflammation Following Closed Head Injury in Rats Is Associated With HPA Axis Hyperactivity." *Metabolomics* 18: 28. https://doi.org/10.1007/s11306-022-01886-8.

Banoei, M. M., C. Casault, S. M. Metwaly, and B. W. Winston. 2018. "Metabolomics and Biomarker Discovery in Traumatic Brain Injury." *Journal of Neurotrauma* 35: 1831–1848. https://doi.org/10.1089/neu.2017.5326.

Brooks, G. A., and N. A. Martin. 2015. "Cerebral Metabolism Following Traumatic Brain Injury: New Discoveries With Implications for Treatment." *Frontiers in Neuroscience* 8: 408. https://doi.org/10.3389/fnins.2014.00408.

Bykowski, E. A., J. N. Petersson, S. Dukelow, et al. 2021. "Urinary Metabolomic Signatures as Indicators of Injury Severity Following Traumatic Brain Injury: A Pilot Study." *IBRO Neuroscience Reports* 11: 200–206. https://doi.org/10.1016/j.ibneur.2021.10.003.

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. "DADA2: High-Resolution Sample Inference From Illumina Amplicon Data." *Nature Methods* 13: 581–583. https://doi.org/10.1038/nmeth.3869.

Cox, T. O., P. Lundgren, K. Nath, and C. A. Thaiss. 2022. "Metabolic Control by the Microbiome." *Genome Medicine* 14: 80. https://doi.org/10.1186/s13073-022-01092-0.

Dennis, E. L., J. Faskowitz, F. Rashid, et al. 2017. "Diverging Volumetric Trajectories Following Pediatric Traumatic Brain Injury." *Neuroimage: Clinical* 15: 125–135. https://doi.org/10.1016/j.nicl.2017.03.014.

Diringer, M. N., A. R. Zazulia, and W. J. Powers. 2011. "Does Ischemia Contribute to Energy Failure in Severe TBI?" *Translational Stroke Research* 2: 517–523. https://doi.org/10.1007/s12975-011-0119-8.

Du, D., W. Tang, C. Zhou, et al. 2021. "Fecal Microbiota Transplantation Is a Promising Method to Restore Gut Microbiota Dysbiosis and Relieve Neurological Deficits After Traumatic Brain Injury." *Oxidative Medicine and Cellular Longevity* 2021: 5816837. https://doi.org/10.1155/2021/5816837.

Feighery, L., A. Smyth, S. Keely, et al. 2008. "Increased Intestinal Permeability in Rats Subjected to Traumatic Frontal Lobe Percussion Brain Injury." *Journal of Trauma* 64: 131–137. https://doi.org/10.1097/TA.0b013e3181568d9f.

Ferreira-Halder, C. V., A. V. d. S. Faria, and S. S. Andrade. 2017. "Action and Function of *Faecalibacterium prausnitzii* in Health and Disease."

10974547, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jnr.70032 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Best Practice & Research Clinical Gastroenterology 31: 643–648. https://doi.org/10.1016/j.bpg.2017.09.011.

Ghandforoush-Sattari, M., S. O. Mashayekhi, M. Nemati, and H. Ayromlou. 2011. "Changes in Plasma Concentration of Taurine in Stroke." *Neuroscience Letters* 496: 172–175. https://doi.org/10.1016/j.neulet.2011.04.010.

Gisewhite, S., I. J. Stewart, G. Beilman, and E. Lusczek. 2021. "Urinary Metabolites Predict Mortality or Need for Renal Replacement Therapy After Combat Injury." *Critical Care* 25: 119. https://doi.org/10.1186/s13054-021-03544-2.

Goel, N., J. L. Workman, T. T. Lee, L. Innala, and V. Viau. 2014. "Sex Differences in the HPA Axis." *Comprehensive Physiology* 4: 1121–1155. https://doi.org/10.1002/cphy.c130054.

Hanscom, M., D. J. Loane, and T. Shea-Donohue. 2021. "Brain-Gut Axis Dysfunction in the Pathogenesis of Traumatic Brain Injury." *Journal of Clinical Investigation* 131, no. 12: e143777. https://doi.org/10.1172/JCI143777.

Hoffe, B., and M. R. Holahan. 2022. "Hyperacute Excitotoxic Mechanisms and Synaptic Dysfunction Involved in Traumatic Brain Injury." *Frontiers in Molecular Neuroscience* 15: 831825. https://doi.org/10.3389/fnmol.2022.831825.

Jiménez-Uribe, A. P., E. Y. Hernández-Cruz, K. J. Ramírez-Magaña, and J. Pedraza-Chaverri. 2021. "Involvement of Tricarboxylic Acid Cycle Metabolites in Kidney Diseases." *Biomolecules* 11: 1259. https://doi.org/10.3390/biom11091259.

Kelestimur, F. 2021. "Antibodies Against the Pituitary and Hypothalamus in Boxers." *Handbook of Clinical Neurology* 181: 187–191. https://doi.org/10.1016/B978-0-12-820683-6.00014-2.

Kim, M.-H., and H. Kim. 2017. "The Roles of Glutamine in the Intestine and Its Implication in Intestinal Diseases." *International Journal of Molecular Sciences* 18: 1051. https://doi.org/10.3390/ijms18051051.

Krueger, E. S., T. S. Lloyd, and J. S. Tessem. 2021. "The Accumulation and Molecular Effects of Trimethylamine N-Oxide on Metabolic Tissues: It's Not all Bad." *Nutrients* 13, no. 8: 2873. https://doi.org/10.3390/nu13082873.

Kurtz, P., and E. E. M. Rocha. 2020. "Nutrition Therapy, Glucose Control, and Brain Metabolism in Traumatic Brain Injury: A Multimodal Monitoring Approach." *Frontiers in Neuroscience* 14: 190. https://doi.org/10.3389/fnins.2020.00190.

Lawrence, S., and R. H. Scofield. 2024. "Post Traumatic Stress Disorder Associated Hypothalamic-Pituitary-Adrenal Axis Dysregulation and Physical Illness." *Brain, Behavior, & Immunity-Health* 41: 100849. https://doi.org/10.1016/j.bbih.2024.100849.

Leigh, S.-J., F. Uhlig, L. Wilmes, et al. 2023. "The Impact of Acute and Chronic Stress on Gastrointestinal Physiology and Function: A Microbiota-Gut-Brain Axis Perspective." *Journal of Physiology* 601: 4491–4538. https://doi.org/10.1113/JP281951.

Leistner, C., and A. Menke. 2020. "Chapter 4: Hypothalamic–Pituitary–Adrenal Axis and Stress." In *Handbook of Clinical Neurology, Sex Differences in Neurology and Psychiatry*, edited by R. Lanzenberger, G. S. Kranz, and I. Savic, 55–64. Elsevier. https://doi.org/10.1016/B978-0-444-64123-6.00004-7.

Li, M., and S. Sirko. 2018. "Traumatic Brain Injury: At the Crossroads of Neuropathology and Common Metabolic Endocrinopathies." *Journal of Clinical Medicine* 7: 59. https://doi.org/10.3390/jcm7030059.

Lim, T. Y., R. L. Poole, and N. M. Pageler. 2014. "Propylene Glycol Toxicity in Children." *Journal of Pediatric Pharmacology and Therapeutics* 19: 277–282. https://doi.org/10.5863/1551-6776-19.4.277.

Liu, J., P. H. Bisschop, L. Eggels, et al. 2012. "Intrahypothalamic Estradiol Modulates Hypothalamus-Pituitary-Adrenal-Axis Activity in Female Rats." *Endocrinology* 153: 3337–3344. https://doi.org/10.1210/en.2011-2176.

Liu, Y., and M. Dai. 2020. "Trimethylamine N-Oxide Generated by the Gut Microbiota Is Associated With Vascular Inflammation: New Insights Into Atherosclerosis." *Mediators of Inflammation* 2020: 4634172. https://doi.org/10.1155/2020/4634172.

Manninen, A. H. 2004. "Metabolic Effects of the Very-Low-Carbohydrate Diets: Misunderstood "Villains" of Human Metabolism." *Journal of the International Society of Sports Nutrition* 1: 7–11. https://doi.org/10.1186/1550-2783-1-2-7.

McDonald, S. J., J. M. Sharkey, M. Sun, et al. 2020. "Beyond the Brain: Peripheral Interactions After Traumatic Brain Injury." *Journal of Neurotrauma* 37: 770–781. https://doi.org/10.1089/neu.2019.6885.

Mercado, N. M., G. Zhang, Z. Ying, and F. Gómez-Pinilla. 2022. "Traumatic Brain Injury Alters the Gut-Derived Serotonergic System and Associated Peripheral Organs." *Biochimica et Biophysica Acta-Molecular Basis of Disease* 1868: 166491. https://doi.org/10.1016/j.bbadis.2022.166491.

Mitchell, S. C., and A. Q. Zhang. 2001. "Methylamine in Human Urine." *Clinica Chimica Acta* 312: 107–114. https://doi.org/10.1016/s0009-8981(01)00608-8.

Moffett, J. R., N. Puthillathu, R. Vengilote, D. M. Jaworski, and A. M. Namboodiri. 2020. "Acetate Revisited: A Key Biomolecule at the Nexus of Metabolism, Epigenetics, and Oncogenesis—Part 2: Acetate and ACSS2 in Health and Disease." *Frontiers in Physiology* 11: 580171. https://doi.org/10.3389/fphys.2020.580171.

Morganti-Kossmann, M. C., M. Rancan, P. F. Stahel, and T. Kossmann. 2002. "Inflammatory Response in Acute Traumatic Brain Injury: A Double-Edged Sword." *Current Opinion in Critical Care* 8: 101–105. https://doi.org/10.1097/00075198-200204000-00002.

Ng, S. Y., and A. Y. W. Lee. 2019. "Traumatic Brain Injuries: Pathophysiology and Potential Therapeutic Targets." *Frontiers in Cellular Neuroscience* 13: 528. https://doi.org/10.3389/fncel.2019.00528.

Opeyemi, O. M., M. B. Rogers, B. A. Firek, et al. 2021. "Sustained Dysbiosis and Decreased Fecal Short-Chain Fatty Acids After Traumatic Brain Injury and Impact on Neurologic Outcome." *Journal of Neurotrauma* 38: 2610–2621. https://doi.org/10.1089/neu.2020.7506.

Paxinos, G., and C. Watson. 2006. *The Rat Brain in Stereotaxic Coordinates: Hard Cover Edition*. Elsevier.

Qian, X., S. Droste, S. Lightman, J. Reul, and A. Linthorst. 2012. "Circadian and Ultradian Rhythms of Free Glucocorticoid Hormone Are Highly Synchronized Between the Blood, the Subcutaneous Tissue, and the Brain." *Endocrinology* 153: 4346–4353. https://doi.org/10.1210/en.2012-1484.

Robinson, M. D., D. J. McCarthy, and G. K. Smyth. 2010. "edgeR: A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data." *Bioinformatics* 26: 139–140. https://doi.org/10.1093/bioinformatics/btp616.

Rosin, S., K. Xia, M. A. Azcarate-Peril, et al. 2021. "A Preliminary Study of Gut Microbiome Variation and HPA Axis Reactivity in Healthy Infants." *Psychoneuroendocrinology* 124: 105046. https://doi.org/10.1016/j.psyneuen.2020.105046.

Sidorov, E., D. K. Sanghera, and J. K. P. Vanamala. 2019. "Biomarker for Ischemic Stroke Using Metabolome: A Clinician Perspective." *Journal of Stroke* 21: 31–41. https://doi.org/10.5853/jos.2018.03454.

Stojanov, S., A. Berlec, and B. Štrukelj. 2020. "The Influence of Probiotics on the Firmicutes/Bacteroidetes Ratio in the Treatment of Obesity and Inflammatory Bowel Disease." *Microorganisms* 8: 1715. https://doi.org/10.3390/microorganisms8111715.

Sudo, N. 2014. "Microbiome, HPA Axis and Production of Endocrine Hormones in the Gut." *Advances in Experimental Medicine and Biology* 817: 177–194. https://doi.org/10.1007/978-1-4939-0897-4_8.

Sun, B., C. Hu, H. Fang, L. Zhu, N. Gao, and J. Zhu. 2015. "The Effects of *Lactobacillus acidophilus* on the Intestinal Smooth Muscle Contraction

10974547, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jnr.70032 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Through PKC/MLCK/MLC Signaling Pathway in TBI Mouse Model." *PLoS One* 10: e0128214. https://doi.org/10.1371/journal.pone.0128214.

Tapp, Z. M., J. P. Godbout, and O. N. Kokiko-Cochran. 2019. "A Tilted Axis: Maladaptive Inflammation and HPA Axis Dysfunction Contribute to Consequences of TBI." *Frontiers in Neurology* 10: 345. https://doi.org/10.3389/fneur.2019.00345.

Taraskina, A., O. Ignatyeva, D. Lisovaya, et al. 2022. "Effects of Traumatic Brain Injury on the Gut Microbiota Composition and Serum Amino Acid Profile in Rats." *Cells* 11: 1409. https://doi.org/10.3390/cells11091409.

Thomas, I., A. M. Dickens, J. P. Posti, et al. 2022. "Serum Metabolome Associated With Severity of Acute Traumatic Brain Injury." *Nature Communications* 13: 2545. https://doi.org/10.1038/s41467-022-30227-5.

Trivedi, R., A. R. Khan, P. Rana, et al. 2012. "Radiation-Induced Early Changes in the Brain and Behavior: Serial Diffusion Tensor Imaging and Behavioral Evaluation After Graded Doses of Radiation." *Journal of Neuroscience Research* 90: 2009–2019. https://doi.org/10.1002/jnr.23073.

Tsikas, D. 2020. "Urinary Dimethylamine (DMA) and Its Precursor Asymmetric Dimethylarginine (ADMA) in Clinical Medicine, in the Context of Nitric Oxide (NO) and Beyond." *Journal of Clinical Medicine* 9: 1843. https://doi.org/10.3390/jcm9061843.

Van Hul, M., T. Le Roy, E. Prifti, et al. 2020. "From Correlation to Causality: The Case of Subdoligranulum." *Gut Microbes* 12: 1–13. https://doi.org/10.1080/19490976.2020.1849998.

Velasquez, M. T., A. Ramezani, A. Manal, and D. S. Raj. 2016. "Trimethylamine N-Oxide: The Good, the Bad and the Unknown." *Toxins (Basel)* 8: 326. https://doi.org/10.3390/toxins8110326.

Vernocchi, P., F. Del Chierico, and L. Putignani. 2020. "Gut Microbiota Metabolism and Interaction With Food Components." *International Journal of Molecular Sciences* 21: 3688. https://doi.org/10.3390/ijms21103688.

Vijapur, S. M., Z. Yang, D. J. Barton, et al. 2020. "Anti-Pituitary and Anti-Hypothalamus Autoantibody Associations With Inflammation and Persistent Hypogonadotropic Hypogonadism in Men With Traumatic Brain Injury." *Journal of Neurotrauma* 37: 1609–1626. https://doi.org/10.1089/neu.2019.6780.

Wen, C., F. Li, L. Zhang, et al. 2019. "Taurine Is Involved in Energy Metabolism in Muscles, Adipose Tissue, and the Liver." *Molecular Nutrition & Food Research* 63: e1800536. https://doi.org/10.1002/mnfr.201800536.

Wen, S., C. Wang, M. Gong, and L. Zhou. 2019. "An Overview of Energy and Metabolic Regulation." *Science China Life Sciences* 62: 771–790. https://doi.org/10.1007/s11427-018-9371-4.

Wright, E. S. 2016. "Using DECIPHER v2.0 to Analyze Big Biological Sequence Data in R." *R Journal* 8: 352–359.

Xia, J., N. Psychogios, N. Young, and D. S. Wishart. 2009. "MetaboAnalyst: A Web Server for Metabolomic Data Analysis and Interpretation." *Nucleic Acids Research* 37: W652–W660. https://doi.org/10.1093/nar/gkp356.

Zheng, L., Q. Pang, H. Xu, H. Guo, R. Liu, and T. Wang. 2022. "The Neurobiological Links Between Stress and Traumatic Brain Injury: A Review of Research to Date." *International Journal of Molecular Sciences* 23: 9519. https://doi.org/10.3390/ijms23179519.

Zheng, Z., S. Wang, C. Wu, et al. 2022. "Gut Microbiota Dysbiosis After Traumatic Brain Injury Contributes to Persistent Microglial Activation Associated With Upregulated Lyz2 and Shifted Tryptophan Metabolic Phenotype." *Nutrients* 14: 3467. https://doi.org/10.3390/nu14173467.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

10974547, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/jnr.70032 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

This article has been accepted for publication in IEEE Transactions on Instrumentation and Measurement. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIM.2025.3560721

Analysis of a Digitizer Circuit with T-Network for Low Resistance Measurement

Shiraz Sohail, *Senior Member, IEEE*, Gopal Singh, *Graduate Student Member, IEEE*, Tarikul Islam, *Senior Member, IEEE*

***Abstract*—An efficient interfacing circuit designed for measuring low resistance (R_X, less than 1 Ω) and generating a digital output is presented. The digitizer (resistance to digital converter, RDC) circuit utilizes a dual-slope integrator with a T-network in feedback, three single pole single throw (SPDT) switches, and a comparator. Conventional dual-slope integrators face challenges when measuring low R_X connected to the inverting terminal, as the high input current causes the op-amp to saturate. Reducing the input voltage to even 0.01 V doesn't resolve this issue. Alternatively, connecting R_X in feedback produces a low output voltage that is susceptible to inherent circuit noise, leading to significant measurement errors. Consequently, no digitizers have been developed to accurately measure R_X below 10 Ω so far. The circuit overcomes these limitations by using a T-network which enables the incoming current (I_X, 98.09 - 71.75 μA) through R_X to amplify the feedback current (I_C), hence significantly improving sensitivity which is a novel feature. During charging, lead resistance (R_L) with fixed resistance R_3 (or R_X) is measured, while during discharging, R_X and R_L are measured together. Subtracting R_L ensures accurate R_X measurement, independent of R_L, in a two-wire configuration. The circuit is experimentally tested for R_X in the range of 0.1 - 0.9 Ω (full span) in which the power dissipation ranges from 0.96 - 4.63 nW, while the measurement time varies from 1.29 - 1.51 s. The circuit yields accurate results with an average error of ± 0.68% across the full span of R_X.**

***Index Terms*—Low resistance, T-network, dual-slope integrator, digitizer, resistance to digital converter (RDC), lead resistance compensation, two-wire type.**

I. Introduction

ACCURATE measurement of low resistance holds a prime importance in characterizing various components such as ammeter shunts, inductive coils, transformer primary and secondary windings, motor armatures and field windings, sheet resistance, and more. By definition, any resistance below 1 Ω is categorized as low resistance [1]. Traditionally, three techniques are used for measurement of low resistance: (i) Ammeter and voltmeter method (ii) Potentiometer method, and (iii) Bridge method. In ammeter and voltmeter method, a recent work, which requires a DC current calibrator, a current generator for supplying currents up to 100 Ampere (A) and 1200 A, a current-reversing switch, and two nanovoltmeters make the measurement complex [2]. In potentiometric method, Dauphinee et al. used isolating potential comparator to generate the difference between the potential drops in the two resistances without making a direct differential connection to improve the measurement [3]. Further, a modified potentiometric technique which requires a high current around 40 A to operate is also presented [4]. In bridge methods, the Kelvin double bridge (KDB) is particularly popular in industry due to its high accuracy. The errors caused by the lead resistances and contact resistances of the variable jockeys are minimized by selecting high values for the outer arms (≈ 10–100 kΩ) and inner arms (≈ 100–500 Ω). The KDB is improved by interchanging the battery and the galvanometer which results in the conjugate KDB [5]. However, both of them requires the ratio of the arms to be fixed and equal, which can deviate with temperature and aging [6]. Thereafter, the Warshawsky bridge minimizes random contact resistances occurring at the junctions of Wheatstone bridge by using multiple bridge circuit and adding an auxiliary resistor pair at each bridge corner [7]. Another improvement is the automated modified Wheatstone bridge [8] which features four 1-Ω resistors in a bridge configuration with four galvanometers. It includes six lead wires, five variable shunt resistors, and a constant current source for excitation. Both [7] and [8] proposes highly complex circuit design for measurement of low R_X.

It is to be noted that all above mentioned techniques are only suitable for four wire type R_X, and require considerable time for manual bridge balancing using precision adjustable resistors. Moreover, for proper working and accurate results, several conditions for resistance values (such as lead, contact and supply series resistances) has to be fulfilled, and it also affects the accuracy and sensitivity. The bridge balance and sensitivity, both depend upon the galvanometer's sensitivity and coil resistance [9][10] as well as bridge supply voltage. Although balancing can be avoided by calibrating the bridge voltage for R_X, the finite resistance of the galvanometer still introduces an error in the measurement. Thus, a suitable correction must be applied to the measured R_X to account for the non-ideal properties of the meters involved.

The literature also includes the following recent works. In [5], a direct current comparator ratio bridge measures low resistances by comparing current ratios that equalize voltage drops. The current ratio, determined by the turns ratio, is automatically maintained by a self-balancing direct current comparator. The current injection method [11] for enhancing the Wheatstone bridge is complex, requiring a current source, a conductance meter, and a deflection meter. The cryogenic current comparator bridge described in [12] is a research

*Corresponding Author: Shiraz Sohail and Gopal Singh (Equally contributed authors, Email: ssohail@nitt.edu, 1995.gopal.singh@gmail.com).
[1]Shiraz Sohail and Gopal Singh, are with the Department of Instrumentation and Control Engineering, National Institute of Technology, Tiruchirappalli, TN, 620015, India.
[2]Tarikul Islam is with the Department of Electrical Engineering, Jamia Millia Islamia (A Central University), New Delhi, 110025, India.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:52:49 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Instrumentation and Measurement. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIM.2025.3560721

TABLE I
LITERATURE COMPARISON TABLE ON LOW RESISTANCE MEASUREMENT

Ref.	Resistance Range	Methodology Used	Components Used	Remarks
[2]	10 $\mu\Omega$ - 10 mΩ	Volt-amperiometric	DC current calibrator, current generator, switches, two nano voltmeters (nV), two digital voltmeters and two DC shunts	A calibration lab setup for high-power (500 W) shunts, requiring 1200 A, self-heating error is compensated using three Platinum thermometers, needs a cooling system with a heat exchanger, fan, pump, and oil-contact serpentine, bulky, costly, requires two 7 1/2 - digit nanovoltmeters and proper insulation, and safety measures due to high current.
[3]	1 - 10 Ω	Isolating potential comparator	Double-pole double throw chopper, condenser, galvanometer and a couple of resistors	Bulky and complex design, need manual adjustments, errors due to contact resistances and lead connections.
[4]	In the range of $\mu\Omega$	Potentiometric method	Linear differential DC amplifier, a potentiometer, a galvanometer, standard resistors	Modified conventional potentiometric method using voltage divider and differential amplifier, need of Diesselhorst compensator to nullify thermoelectric effects, requires I_X - 40 A.
[5]	In the range of 100 Ω - mΩ	Direct current comparator	Magnetic modulator, magnetic shield, two ratio windings, galvanometer and resistors	Resolution depends on number of turns which should be uniformly distributed and tightly packed, requires 10 AT (ampere turn), bulky setup and relies on manual balancing, requires I_X - 150 A.
[7]	Random contact resistances (rcr's) < 2 Ω	Wheatstone bridge	Couple of resistors representing the main wheatstone bridge, rcr's and auxiliary resistors	Manual bridge balance technique, uses multiple bridges which is complex, power dissipation - 0.01 W.
[8]	Calibrate 4-terminal 1 Ω	Modified Wheatstone bridge	Four galvanometer detectors, current source, resistors, temperature sensor and relay drivers	Need a highly sensitive and high accuracy galvanometer detectors, highly complex circuit when its automated, requires I_X - 50 mA, power dissipation - 2.5 mW.
[11]	Calibrate 4-terminal 1 Ω	Wheatstone bridge	Switches, resistors, current source and null detector	Uncertainty in reversing current direction may lead to thermoelectric effects at junctions, highly complex circuit, power dissipation - 1 mW.
[12]	1 Ω - 100 $\mu\Omega$	Cryogenic current comparator (CCC) bridge	Buffer amplifier, current source, ADC and SQUID (superconducting quantum interference device)	Type II CCC design introduces more complexity as compared to type I CCC, requires I_X - 100 A, power dissipation - 100 kW to 1 W.
[13]	10 $\mu\Omega$ - 1 Ω	Current comparator	Commercial current extender, high current source, current meter and a nano voltmeter	Nanovoltmeter need to have a stable gain and calibrated regularly, complex setup, power dissipation - 100 mW.
This work	0.1 - 0.9 Ω	Dual-slope integrator	Integrator, comparator, switches and couple of resistors	Introduction of T-network in feedback, digital output, unaffected due to lead resistance, simple circuitry, average error of $\pm$ 0.68%, requires I_X - 98.09 to 71.75 μA, power dissipation - 0.96 to 4.63 nW.

instrument that requires a low-interference constant current source and a superconducting quantum interference device (SQUID) flux sensor. In [13] describes an automated low-ohmic resistance measurement system requiring a current source, current extender, current meter, and a stable-gain nanovoltmeter in the bridge. In [14] a complex circuit involving photo isolated synchronous rectification is used to measure the contact resistance. The circuit requires current source, an amplifier, band pass filter, gain control, phase detector and finally low pass filter. Table I provides more details about these techniques regarding range, methodology, and features. The measurement techniques presented in [2], [4], [5], [12], [13] requires a high current ($\geq$ 100 A) to be injected into the R_X, and hence not suitable for low resistance measurement of ammeter shunts, inductive coils, transformer and motor windings, where a high current may damage the coil. While the other techniques in [3], [7], [8], [11], [14] use complicated circuit design, costly and sophisticated equipment make the measurement complex, time taking, expensive, and hence not suitable for general purpose, portable, day to day measurements in industry settings. Moreover, all the previously discussed techniques provide analog output, hence an analog-to-digital converter (ADC) is required for interfacing with microcontrollers for tasks such as operational control, signal processing, lead or offset compensation, and calibration.

Recently, digitizer circuits have gained significant attention due to their ability to generate a time-sequence or square wave digital output with pulse width modulation, produced through an op-amp integrator and comparator [15]. The digitizers make the measurement fast and avoid manual balancing and calculations. The digital output allows direct connection to the microcontroller digital pin for signal processing, eliminating the need for a separate ADC. As a result, it conserves analog pins, which are often limited due to costly built-in ADCs in the microcontroller. These digital interfacing circuits can be broadly classified based on their working principles as (a) dual-slope integrator [16], [17], [18] (b) sigma-delta [19], [20] (c) relaxation oscillator [21], and (d) direct interface circuit (DIC) [22], [23].

In dual-slope integrator, typically, the R_X is connected to a dual slope integrator and its output is fed to a comparator to generate a digital output. It is noteworthy that none of the digitizers previously documented in the literature have been proposed for measuring R_X below 10 Ω. The reason is that when a low value R_X is connected to the inverting terminal of the op-amp in integrator, a large current flows, causing the op-amp to saturate. Even reducing the input voltage to 0.01 V doesn't resolve the issue because the input current will be in the range of 100 mA (when keeping R_X = 0.1 Ω). Else, if R_X is connected in the feedback of op-amp, the output voltage will be very small, and prone to circuit inbuilt-noises which result in large measurement error.

To address this significant issue, [18] presents a simulation and initial investigation of an digitizer designed for measuring

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:52:49 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Instrumentation and Measurement. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIM.2025.3560721

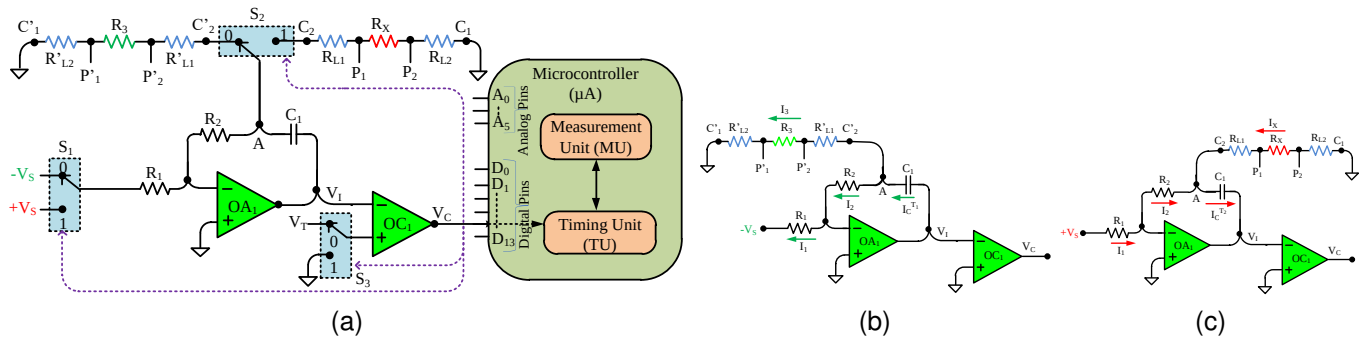


Fig. 1. (a) Proposed digitizer circuit for measurement of low resistance based on dual slope integrator with a T-network in the feedback. Simplified circuit model of the digitizer during (b) charging (T_1) interval, and (c) discharging (T_2) interval. These sub-figures show the current directions and proper labels.

R_X below 1 Ω. This study comprehensively analyzes the circuit featuring a dual-slope integrator with a T-network in its feedback loop. It also consists of three single pole double throw (SPDT) switches and a comparator. In this configuration, R_X is connected to the middle branch of the T-network, while a fixed resistor and a feedback capacitor are attached to the other two branches. Due to T-network, the incoming current (I_X) through R_X is responsible for amplifying the feedback (capacitor) current (I_C) which charges and discharges the integrating capacitor, and hence measure R_X. The I_X due to T-network enables faster R_X measurement in a two-wire configuration which is another feature of the circuit.

The proposed circuit features: 1) R_X range: 0.1 – 0.9 Ω; 2) measurement time: 1.29 – 1.51 s; 3) I_X: 98.09 – 71.75 μA; 4) power dissipation: 0.96 – 4.63 nW. In summary, the circuit offers the following novelty: 1) no manual balancing; 2) low power dissipation; 3) T-network feedback for boosted input current; 4) R_X measurement independent of R_L in two-wire configuration; 5) time-sequence digital output; 6) minimal error (± 0.68%) with good linearity (0.54% NL); and 7) no need for costly components like a waveform generator.

II. Proposed Digitizer Circuit with T-Network

TABLE II
Switching Position

	Switches		
State	S_1	S_2	S_3
Charging (T_1 interval)	0	0	0
Discharging (T_2 interval)	1	1	1

A. Circuit description

The proposed circuit is based on a dual-slope integrator which consists of an op-amp (OA_1) with a T-network in feedback as shown in Fig. 1. One branch of the T-network consists of a fixed resistance R_2, and the other one consists of either unknown low value R_X or fixed resistance R_3. The capacitor C_1 forms the third branch of the T-network and integrates the capacitor current (I_C). The R_2 is used to develop a fixed bias voltage (V_A = I_2R_2) so that incoming current (I_X) is a function of R_X and is equal to I_2R_2/R_X during discharging mode, and hence R_X can be measured. The R_X also includes lead resistance (R_L) since it is of two-wire type. To compensate it, a known R_3 with equally valued R_L is included in the circuit during charging, ensuring that the R_L is canceled out in the final expression of R_X.

The three single pole double throw (SPDT) switches S_1, S_2 and S_3 facilitate the charging and discharging modes of operation as per Table II. Green and red arrows represent the current paths during charging and discharging modes in Fig. 1b and 1c respectively. Initially, S_1, S_2 and S_3, all are at '0' position (refer Table II), and the equivalent circuit is shown in Fig. 1b, where $I_C^{T_1}$ splits into I_2 and I_3 at A, and exits C_1. The current flows from integrator output (V_I) to negative input (-V_S) and charges C_1, which results in gradual increase of V_I. When V_I reaches to threshold voltage (V_T) of comparator, the switches S_1, S_2 and S_3, changed to '1' position (refer Table II). The equivalent circuit is shown in Fig. 1c, where I_X and I_2 are summed as $I_C^{T_2}$ at A, and enters C_1. The current flows from the positive input (+V_S) to the V_I, and discharges C_1, due to which V_I gradually decreases to zero potential. Thus, I_C is equal to I_2 + I_3 and I_2 + I_X during charging and discharging modes respectively. The output of the integrator (V_I) is a triangular waveform which is fed to the comparator (OC_1) to compare with the threshold voltage (V_T) and generates a quasi-square waveform (V_C, pulse-type). Thus, R_X is converted into a digital waveform for direct and easy interfacing with microcontrollers. The integrator's charging (T_1) and discharging (T_2) times are measured using a timing unit (TU), and the measurement unit (MU) is used to calculate the R_X using analytical model equations. Both TU and MU are realized inside the microcontroller. The controlling signal, necessary to operate the switches is also derived from V_C.

For accurate measurement of R_X, cancelling R_L associated with R_X is crucial. Initially, S_2 connects the middle branch of the T-network to R_3 during the charging mode to calculate R_L. During discharging mode, S_2 connects the T-network to R_X to find $R_X + R_L$. Finally, R_L is subtracted from $R_X + R_L$ to measure R_X accurately. This converts the measurement scheme to a two-wire type, improving upon the typically used four-wire type [24], [25]. The switches S_1 and S_3 are used for switching the input voltages ($+V_S$ and $-V_S$) and comparator references (V_T and ground) respectively.

B. Derivation of unknown resistance R_X

Initially, switches S_1, S_2 and S_3 are set to position 0 as V_C is HIGH. In this state, the charging current ($I_C^{T_1}$) flows

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:52:49 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Instrumentation and Measurement. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIM.2025.3560721

from V_I back to $-V_S$ through C_1. At node A, the current $I_C^{T_1}$ splits into two paths. In first path, current I_3 passes through R_3 and R_L ($= R_{L1}^{'} + R_{L2}^{'}$) while in second path current I_1 passes through R_1, ultimately sinking to ground voltage and $-V_S$ respectively. In practice, S_1 and S_2 switches [26] have a very small on-resistance of approximately 1.25 Ω (r_1, r_2). The expressions are as follows:

$$I_1 = \frac{|V_S|}{R_1 + r_1} \tag{1}$$

$$I_C^{T_1} = I_2 + I_3 \tag{2}$$

$$I_2 = I_C^{T_1}\left(\frac{R_3 + R_L + r_2}{R_3 + R_L + R_2 + r_2}\right) = I_1 \tag{3}$$

$$I_3 = I_C^{T_1}\left(\frac{R_2}{R_3 + R_L + R_2 + r_2}\right) = \left(\frac{R_2}{R_3 + R_L + r_2}\right) I_2 \tag{4}$$

The expression for V_I is

$$V_I = \left(2I_2R_2 + \frac{I_C^{T_1}}{C_1}t\right) \tag{5}$$

The V_I is fed into OC_1 along with V_T. As long as V_I is lower than V_T, V_C stays HIGH. When V_I equals V_T, V_C goes low, signaling the end of the T_1 interval. The change in integrator's output voltage (ΔV_I) during the T_1 interval is:

$$\Delta V_I = \left(2I_2R_2 + \frac{I_C^{T_1}}{C_1}T_1\right) = V_T \tag{6}$$

The analytical expression for T_1 is:

$$T_1 = \frac{C_1\left(V_T - 2I_2R_2\right)}{I_C^{T_1}} \tag{7}$$

The TU within the microcontroller measures T_1. By using equations (1), (3), and (6), and assuming that $R_1 \gg r_1$ the expression for R_L is:

$$R_L = \frac{T_1R_2V_S}{C_1\left[R_1V_T - 2R_2V_S\right] - T_1V_S} - R_3 - r_2 \tag{8}$$

Now, with V_C LOW, all switches move to position 1, and the integrator enters the discharge mode (T_2 interval). The current I_1 starts from +V_S and gets amplified by (1+(R_X+R_L+r_2)/R_2) times at node A due to I_X. This amplified current enters C_1 and forms the discharge current ($I_C^{T_2}$) which finally sinks at V_I. Here R_L (=R_{L1}+R_{L2}) represents the lead resistance. The expressions are:

$$I_C^{T_2} = I_2 + I_X \tag{9}$$

$$I_X = I_C^{T_2}\left(\frac{R_2}{R_X + R_L + R_2 + r_2}\right) = \left(\frac{R_2}{R_X + R_L + r_2}\right) I_2 \tag{10}$$

The expression for V_I during discharge mode is

$$V_I = V_T - \left(2I_2R_2 + \frac{I_C^{T_2}}{C_1}t\right) \tag{11}$$

As long as V_I is higher than ground potential, V_C remains low. When V_I equals to ground potential, V_C goes HIGH, marking the end of the T_2 interval. At this moment, the circuit re-enters the charging mode for the next cycle, and all switches return to their previous positions; with S_1, S_2, and S_3 all moving to position 0. The change in integrator's output (ΔV_I) during the T_2 interval is:

$$\Delta V_I = \left(2I_2R_2 + \frac{I_C^{T_2}}{C_1}T_2\right) = V_T \tag{12}$$

The analytical expression for T_2 is:

$$T_2 = \frac{C_1\left(V_T - 2I_2R_2\right)}{I_C^{T_2}} \tag{13}$$

The net charge in C_1 will be zero at the end of the T_2 interval, and (6) and (12) can be equated as

$$\left(2I_2R_2 + \frac{I_C^{T_1}}{C_1}T_1\right) = \left(2I_2R_2 + \frac{I_C^{T_2}}{C_1}T_2\right) \tag{14}$$

Finally, the expression of R_X can be obtained from (1), (2), (3), (4), (9), (10) and (14) as:

$$R_X = \frac{T_2R_2\left(R_3 + R_L^{(8)} + r_2\right)}{(T_1 - T_2)\left(R_3 + R_L^{(8)} + r_2\right) + T_1R_2} - R_L^{(8)} - r_2 \tag{15}$$

The value of R_L is substituted from (8) to find R_X. Previously, digitizers [16], [17], [18], [19], [20], [21], [22], [23] did not show results for measuring low resistance (less than 10 Ω). This marks the initial effort in accurately measuring low resistance using digitizers, achieving results comparable to those of the Kelvin double bridge [1][6].

III. Analysis

A. Effect of Opamp and Comparator offset Noise

Parameters like input offset voltage (V_{OS}) and input offset current (I_{OS}) in an op-amp typically contribute to an output offset error (E_O). To analyze the errors introduced by I_{OS} and V_{OS} in the integrator (OA_1), the equivalent circuit of the proposed scheme, with V_S set to zero, is shown in Fig. 2(a-b). Additionally, the comparator (OC_1) includes an inherent input offset error denoted as V_{IO}.

To minimize the error caused by I_{OS}, a compensation resistor (R_P) is connected to the non-inverting terminal of OA_1. The selection of an appropriate R_P value is discussed in the following section. Since, S_1 on-resistance (r_1) is in series with R_1 whose value is typically 100 kΩ, so it is ignored. But, S_2 on-resistance (r_2) is in series with either R_X or R_3 whose values is less than 1 Ω, so it is considered. The $r_2^{T_1}$ and $r_2^{T_2}$ represent on-resistances of S_2 during the charging (T_1) and discharging (T_2) phases respectively, due to channel mismatches ($r_2^{T_1} \neq r_2^{T_2}$), with their absolute difference satisfying $|r_2^{T_1} - r_2^{T_2}| \leq 0.3$ Ω as per MAX4690 datasheet [26]. The error in the integrator output ($E_{01}^{'}$) due to I_{OS} during T_1 is:

$$E_{01}^{'}(t) = \frac{1}{C_1}\int_0^t I_C^{T_1}\,dt \tag{16}$$

where $I_C^{T_1} = I_B[a - (R_Pb)/R_1] - (I_{OS}/2)[a + (R_Pb)/R_1], a = 1 + [R_2/(R_4 + r_2^{T_1})], b = 1 + [R_2/(R_4 + r_2^{T_1})] + [R_1/(R_4 + r_2^{T_1})], R_4 = R_3 + R_L$ and $R_L = R_{L1}^{'} + R_{L2}^{'}$.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:52:49 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Instrumentation and Measurement. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIM.2025.3560721

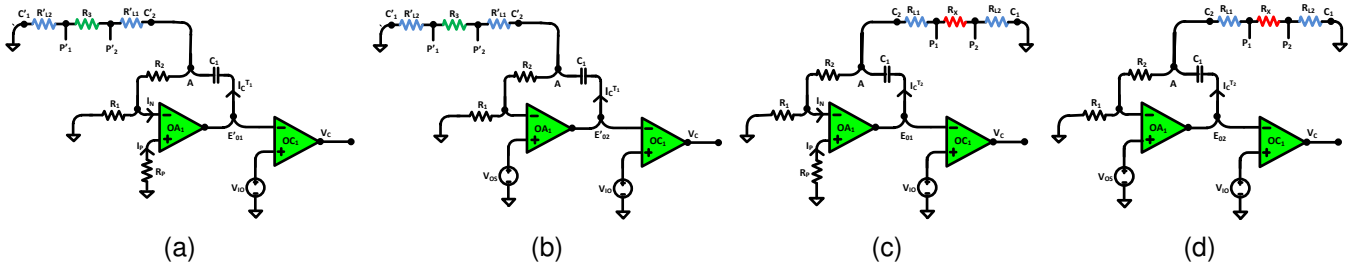


Fig. 2. The equivalent circuits for the digitizer's charging (T_1) and discharging (T_2) modes, considering offset noise. (a) Error due to I_{OS} and V_{IO} in T_1. (b) Error due to V_{OS} and V_{IO} in T_1. (c) Error due to I_{OS} and V_{IO} in T_2. (d) Error due to V_{OS} and V_{IO} in T_2.

Substituting $R_P = (aR_1/b)$ in (16), we get

$$E_{01}^{'}(t) = -\frac{aI_{OS}t}{C_1} \tag{17}$$

where I_P and I_N represent the input bias currents at the non-inverting and inverting terminals of OA_1, respectively. The input offset current (I_{OS}) is defined as $(I_P - I_N)$, while the input bias current (I_B) is given by $(I_P + I_N)/2$.

Error in integrator output ($E_{02}^{'}$) due to V_{OS} during T_1 is:

$$E_{02}^{'}(t) = \frac{1}{R_1C_1}\int_0^t (bV_{OS})dt = \frac{bV_{OS}t}{R_1C_1} \tag{18}$$

Thus, the total error in the integrator output during T_1 is

$$|E_0^{'}(t)| = |E_{01}^{'}(t)| + |E_{02}^{'}(t)| = \left[aI_{OS} + \frac{bV_{OS}}{R_1}\right]\frac{t}{C_1} \tag{19}$$

Substituting $t = T_1$ will give integrator output at the end of T_1

$$|E_0^{'}(T_1)| = \left[aI_{OS} + \frac{bV_{OS}}{R_1}\right]\frac{T_1}{C_1} \tag{20}$$

Now, the effect of V_{IO} on the output of OC_1 will be analyzed. The comparator's V_T shifts to $V_T + V_{IO}$. When the integrator output (including the offset voltage) reaches $V_T + V_{IO}$, V_C transitions from +V_{SAT} to -V_{SAT}. Thus, the voltage balance equation for finding T_1 is:

$$\begin{aligned} V_T + V_{IO} &= V_0(T_1) + |E_0^{'}(T_1)| \\ &= \frac{V_S}{R_1}\left[2R_2 + \frac{aT_1}{C_1}\right] + |E_0^{'}(T_1)| \end{aligned} \tag{21}$$

$V_O(T_1)$ expression is substituted from (5), and finally, T_1 is calculated as:

$$T_1 = \frac{[(V_T + V_{IO})\,R_1 - 2V_SR_2]\,C_1}{a\,[I_{OS}R_1 + V_S] + bV_{OS}} \tag{22}$$

Since, the typical value of I_{OS} is 0.5 pA (for TL072, [27]), it can be easily neglected. Similarly, $V_T + V_{IO}$ can also be approximated as V_T since V_{IO} (or V_{OS}) is 1 mV (for TL072, [27]), is negligible when compared to $V_T = 3V$. Additionally, $V_S + V_{OS}$ is treated as V_S. Thus, T_1 can be re-written as:

$$T_1 = \frac{[V_TR_1 - 2V_SR_2]\,C_1}{aV_S + bV_{OS}} \tag{23}$$

From (22), the expression of R_L in terms of I_{OS} and V_{OS} by considering the approximations mentioned above, can be written as:

$$\begin{aligned} R_L = &\frac{T_1\,[R_2V_S + (R_1 + R_2)\,V_{OS}]}{[V_TR_1 + V_{IO} - 2V_SR_2]\,C_1 - T_1\,[V_S + V_{OS}]} \\ &- R_3 - r_2^{T_1} \end{aligned} \tag{24}$$

After T_1 interval, all switches toggles to position 1, causing OA_1 to discharge which results in a decrease in V_I. At the same time, I_{OS} and V_{OS} slowly charge the integrator; however, their combined effect still contributes to decrease in V_I. This introduces a measurement error. Additionally, a V_{IO} appears at the non-inverting terminal of OC_1. Therefore, both error sources must be incorporated into the analytical model. The complete equivalent circuit is shown in Fig. 2(c-d) during the discharging of the integrator when V_S is set to zero. The error in integrator output (E_{01}) due to I_{OS} during T_2 is:

$$E_{01}(t) = \frac{1}{C_1}\int_0^t I_C^{T_2}dt \tag{25}$$

where $I_C^{T_2} = I_B[c-(R_Pd)/R_1]-(I_{OS}/2)[c+(R_Pd)/R_1], c = 1 + [R_2/(R_5 + r_2^{T_2})], d = 1 + [R_2/(R_5 + r_2^{T_2})] + [R_1/(R_5 + r_2^{T_2})], R_5 = R_X + R_L$ and $R_L = R_{L1} + R_{L2}$.

Substituting $R_P = (cR_1/d)$ in (25), we get

$$E_{01}(t) = -\frac{cI_{OS}t}{C_1} \tag{26}$$

Error in the integrator output (E_{02}) due to V_{OS} during T_2 is:

$$E_{02}(t) = \frac{1}{R_1C_1}\int_0^t (dV_{OS})dt = \frac{dV_{OS}t}{R_1C_1} \tag{27}$$

Thus, the total error in the integrator output during T_2 is:

$$|E_0(t)| = |E_{01}(t)| + |E_{02}(t)| = \left[cI_{OS} + \frac{dV_{OS}}{R_1}\right]\frac{t}{C_1} \tag{28}$$

Substituting $t = T_2$ will give integrator output at the end of T_2

$$|E_0(T_2)| = \left[cI_{OS} + \frac{dV_{OS}}{R_1}\right]\frac{T_2}{C_1} \tag{29}$$

During discharge mode, V_I decreases from $(V_T + V_{IO})$ to V_{IO}. At t = T_2, V_C transitions from -V_{SAT} to +V_{SAT}. Thus, the voltage balance equation for determining T_2 is:

$$\begin{aligned} &V_0(T_2) - |E_0(T_2)| \\ &= \frac{V_S}{R_1}\left[2R_2 + \frac{cT_2}{C_1}\right] - |E_0(T_2)| = V_T \end{aligned} \tag{30}$$

$V_O(T_2)$ expression is substituted from (11), and finally, T_2 is calculated as:

$$T_2 = \frac{[V_TR_1 - 2V_SR_2]\,C_1}{cV_S - dV_{OS}} \tag{31}$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:52:49 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Instrumentation and Measurement. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIM.2025.3560721

From (23) and (31), R_X is obtained as:

$$R_X = \frac{T_2\left(R_4 + r_2^{T_1}\right)\left[V_S R_2 - V_{OS}\left(R_1 + R_2\right)\right]}{p} - R_L^{(24)} - r_2^{T_2} \quad (32)$$

where $p = \left(R_4 + r_2^{T_1}\right)\left[V_S\left(T_1 - T_2\right) + V_{OS}\left(T_1 + T_2\right)\right] + T_1\left[R_2 V_S + V_{OS}\left(R_1 + R_2\right)\right]$.Thus, if we compare (24) with (8) for R_L, and (32) with (15) for R_X, it can be concluded that to select OA_1 with a minimal offset voltage (V_{OS}) to ensure accurate measurements.

B. Effect of Deviation of Fixed Resistance (R_1, R_2 and R_3) and Capacitance (C_1)

The deviation in the value of C_1 does not impact the measurement of R_X as it is independent of C_1 according to (15). The following analytical analysis demonstrates that deviations in the fixed resistances (R_1, R_2, and R_3) have a negligible effect on the measurement. Assuming the fixed resistors are changed to $R_1^{'}$, $R_2^{'}$, and $R_3^{'}$, which are defined as:

$$\begin{aligned} R_1^{'} &= R_1(1+\alpha) \\ R_2^{'} &= R_2(1+\beta) \\ R_3^{'} &= R_3(1+\gamma) \end{aligned}$$

where α, β and γ represents the percentage changes in R_1, R_2 and R_3 respectively due to ageing. Then (15) can be re-written as:

$$R_X^{'} = \frac{T_2 R_2^{'}\left(R_3^{'} + R_L + r_2\right)}{\left(T_1 - T_2\right)\left(R_3^{'} + R_L + r_2\right) + T_1 R_2^{'}} - R_L - r_2 \quad (33)$$

Substituing $R_1^{'}$, $R_2^{'}$, and $R_3^{'}$ in (33) we get:

$$R_X^{'} = \frac{T_2 R_2(1+\beta)\left[R_3(1+\gamma) + R_L + r_2\right]}{\left(T_1 - T_2\right)\left(R_3(1+\gamma) + R_L + r_2\right) + T_1 R_2(1+\beta)} - R_L - r_2 \quad (34)$$

From (34), we can conclude that the impact on the measurement due to deviations in the fixed resistors is negligible, as the terms associated with $(1+\alpha)$, $(1+\beta)$, and $(1+\gamma)$ can be easily neglected compared to the original terms.

C. Effect of Deviation of V_S

The measurement of R_X will remain unaffected by changes in the $\pm\, V_S$ value. However, if there is a discrepancy between $+V_S$ and $-V_S$, the following analysis will explore how this mismatch impacts the measurement of R_X. Let's assume that the input voltages are changed to V_{S1} for $-V_S$ and V_{S2} for $+V_S$, defined as follows:

$$\begin{aligned} V_{S1} &= -V_S(1+\delta) \\ V_{S2} &= +V_S(1+\sigma) \end{aligned}$$

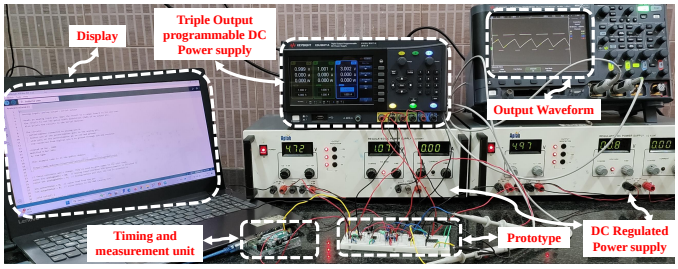


Fig. 3. Experimental setup for testing the proposed digitizer circuit with T-network for measurement of low resistance.

Here, δ and σ represent the percentage changes in $-V_S$ and $+V_S$, respectively. Thus, the charge balance in (14) can be expressed in terms of node voltage as follows:

$$\frac{V_{S1}}{\left(R_1 + r_1\right)}\left[2R_2 + \frac{T_1}{C_1}\left(1 + \frac{R_2}{R_3 + R_L + r_2}\right)\right] = \frac{V_{S2}}{\left(R_1 + r_1\right)}\left[2R_2 + \frac{T_2}{C_1}\left(1 + \frac{R_2}{R_X + R_L + r_2}\right)\right] \quad (35)$$

By substituting the values of V_{S1} and V_{S2} into (35), we obtain the following simplified expression:

$$R_X^{'} = \frac{T_2 R_2(1+\sigma)\left(R_3 + R_L + r_2\right)}{q\left(R_3 + R_L + r_2\right) + T_1 R_2\left(1+\delta\right)} - R_L - r_2 \quad (36)$$

where $q = 2R_2C_1\left(\delta - \sigma\right) - T_2\left(1+\sigma\right) + T_1\left(1+\delta\right)$. Notably, since δ and σ oppose each other, their combined impact on the measurement of R_X is minimal.

IV. Experimental Setup

A. Experimental components

The physical prototype of the dual-slope digitizer has been tested on a breadboard, and both the prototype and the experimental setup are shown in Fig. 3. The integrator (OA_1) and the comparator (OC_1) are realized using a TL072 IC [27] and a LM311 IC respectively. The SPDT switches S_1, S_2, and S_3 are realized with single pole single throw (SPST) CMOS MAX 4690 IC [26] and an inverter SN7404 IC [28]. The controlling signal for the switches is derived from the output of the comparator (V_C). The positive and negative supply (VCC and VEE) of $\pm$ 5 V are obtained from DC-regulated power supply (LQ6324T, Aplab) to OA_1, OC_1, MAX 4690 and SN7404. The values of fixed resistors R_1, R_2 and capacitor C_1 are set at 98.5 $k\Omega$, 25.5 Ω, and 27.6 μF respectively. On the other hand, R_3 is a known four-terminal low-resistance component with a value of 1.2 Ω, which is connected in circuit through its two current terminals ($C_1^{'}$ and $C_2^{'}$) only. The unknown two-terminal low value R_X ranges from 0.1 - 0.9 Ω. It is to be noted that, R_3 and R_X are connected to the circuit through R_L which is also unknown. A triple-output programmable DC power supply, precisely calibrated to an accuracy of $\pm$ 5 μV (EDU36311A, Keysight) is used to provide the input voltage (V_S) of $\pm$ 1 V and V_T of 3 V. For measuring R_X = $f(T_1, T_2)$, the timing unit is implemented using Arduino Uno ATMEGA328P. T_1 represents the time period required for V_I to gradually rise to V_T, while T_2 signifies the period for V_I to fall from V_T to ground. Initially, T_1 is used to determine R_L using (7), and subsequently R_X

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:52:49 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Instrumentation and Measurement. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIM.2025.3560721

TABLE III
SLOPE VARIATION IN SIMULATION FOR DIFFERENT VALUES OF R_X DURING T_1 AND T_2 INTERVALS

Value of R_X (Ω)	Rising Slope (V/s) (T_1 interval)	Falling Slope (V/s) (T_2 interval)
0.1	3.7781	6.0494
0.2	3.7807	5.7267
0.3	3.7763	5.4404
0.4	3.78	5.1741
0.5	3.7768	4.9507
0.6	3.7796	4.7382
0.7	3.7755	4.5394
0.8	3.7795	4.3665
0.9	3.7804	4.2015

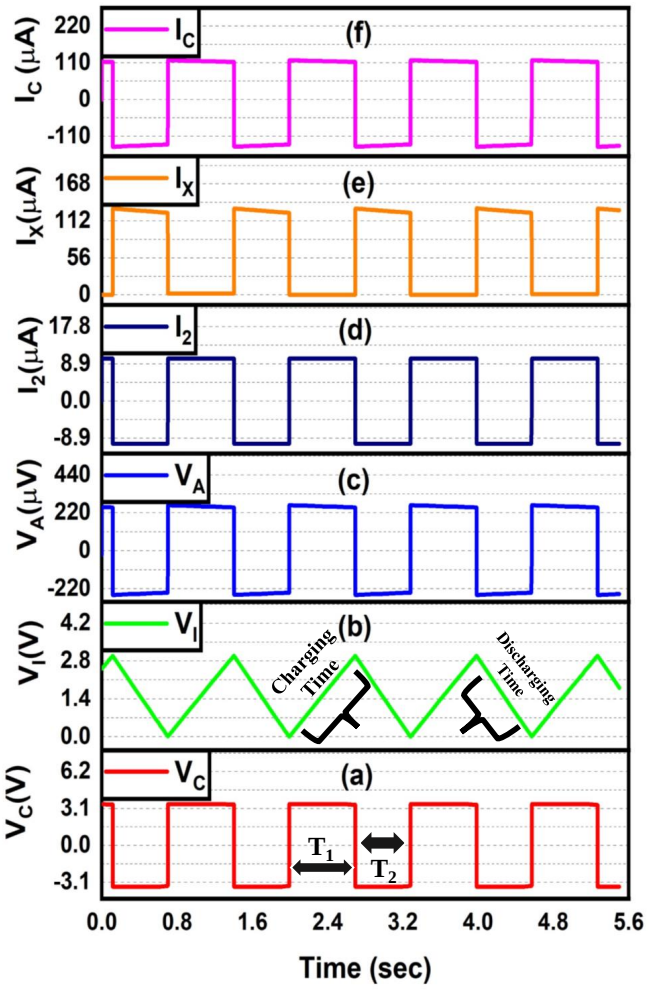


Fig. 4. Plots illustrating various branch currents (I_2, I_X, I_C) and node voltages (V_I and V_C) of the digitizer circuit for R_X = 0.5 Ω.

is calculated using (15). The circuit's performance is assessed by measuring R_X within the specified range. Additionally, a handheld LCR meter (U1733C, Keysight) is also used for verifying the measured R_X values. I_X and I_C are separately measured using a source measurement unit (SMU, B2910BL, Keysight), and a 6 1⁄2 Digit Table-top multi-meter (DM3068, Rigol) respectively.

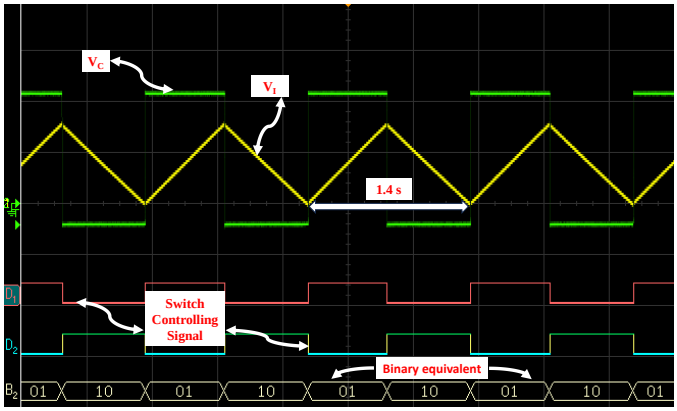


Fig. 5. The integrator (V_I) and comparator (V_C) experimental waveforms, along with control signals for switch operation are shown on mixed signal oscilloscope for R_X = 0.5 Ω.

V. RESULTS AND DISCUSSION

In the proposed digitizer, the V_S is kept at 1 V ensuring that the amplified I_C generated due to T-network remains in the operating range of OA_1, preventing it from entering into a non-linear region. The V_C is a quasi-square waveform, with its duty cycle determined by V_I and V_T. Since V_T is fixed, the time periods T_1 and T_2 primarily depend on V_I which is a triangular waveform. The V_C remains HIGH during the rising of V_I and vice-versa. As predicted in (7) and 13), T_1 remains unaffected by changes in R_X, maintaining a constant slope for V_I despite variations in R_X. In contrast, T_2 has a dependency on R_X, resulting in a different slope of V_I with variation in R_X as shown in Table III.

Fig. 4 (a-c) shows the voltage waveforms at different nodes, namely V_C, V_I, and V_A, for R_X = 0.5 Ω. The T_1 = 794.32 ms and T_2 = 605.98 ms which are measured from V_C using Arduino Uno. The V_A is a bipolar square voltage waveform with equal amplitude in both polarities. Fig. 4 (d-f) shows different current waveforms of I_2, I_X and I_C for R_X = 0.5 Ω. The nature of I_2 waveform is similar to V_A waveform as both remains constant during the charging and discharging mode irrespective of the change in R_X value. The I_X is a unipolar square nature waveform with an amplitude of 82.54 μA as switch S_2 is closed during T_2 interval. During the charging mode, $I_C^{T_1}$ is exited from C_1 and divided into two unequal halves as I_2 and I_3. Similarly, during the discharging mode, I_2 and I_X are summed at node A as $I_C^{T_2}$ and enters the C_1. Since, I_X varies with R_X, current $I_C^{T_2}$ also varies accordingly.

As, I_2 is a bipolar square waveform with equal amplitude in both polarities, its root mean square (RMS) value is equal to modulus of its peak value. For I_C, since its an bipolar unequal square waveform due to presence of I_X only during T_2, its RMS value is calculated using (37), which is given following:

$$I_C^{RMS} = \sqrt{\frac{1}{T}\left(\int_0^{T_1} {I_C^{T_1}}^2\, dt \; + \; \int_{T_1}^{T} {I_C^{T_2}}^2\, dt\right)} \qquad (37)$$

where $T = T_1 + T_2$. The values of $I_C^{T_1}$ and $I_C^{T_2}$ are obtained from (2) and (9) respectively.

The experimental waveforms of V_I, V_C, and control signals in mixed signal oscilloscope (MSOX3014G, Keysight) is

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:52:49 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Instrumentation and Measurement. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIM.2025.3560721

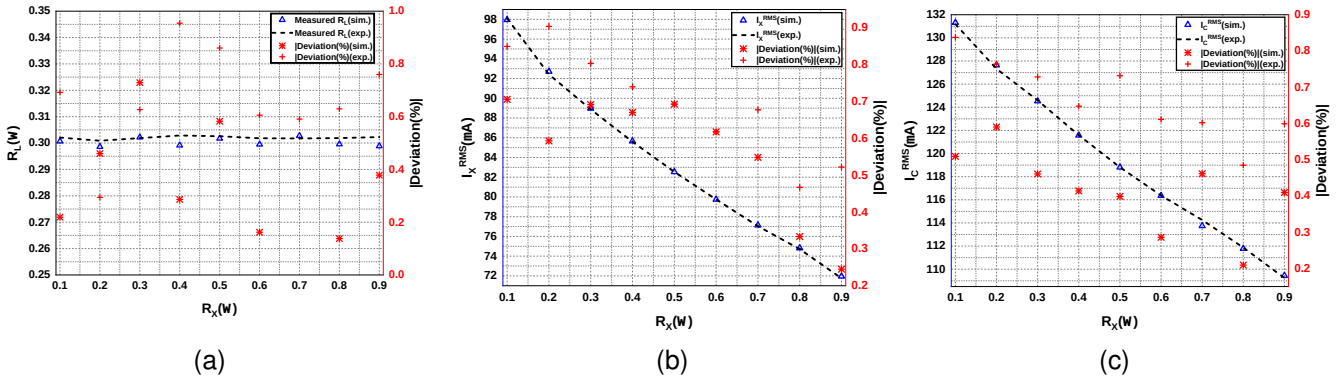


Fig. 6. Plots of parameters and branch currents for R_X varying in the range (0.1 Ω to 0.9 Ω, full span). (a) Lead resistance (R_L). (b) Incoming current (I_X^{RMS}) through R_X. (c) Capacitor current (I_C^{RMS}).

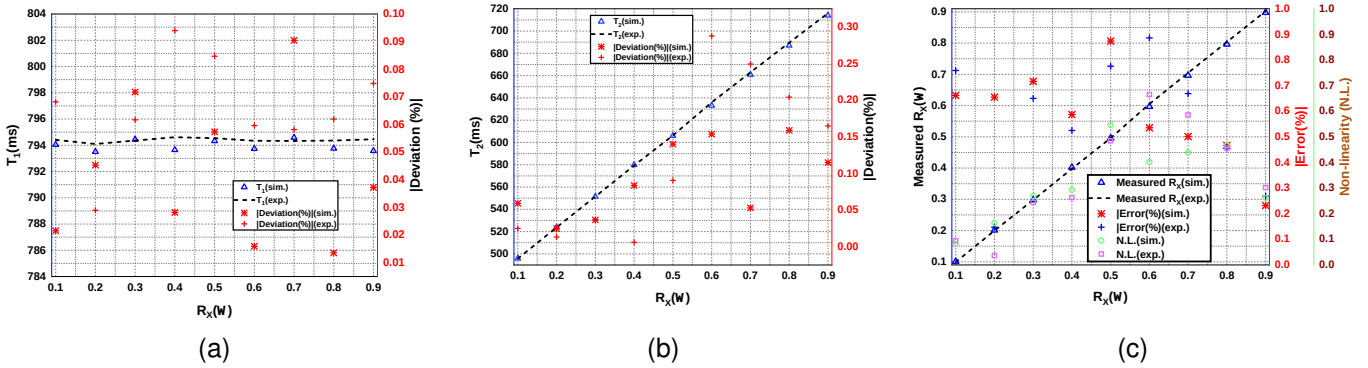


Fig. 7. Plots of different parameters of the digitizer circuit with known R_X varying for the full span. (a) Charging interval (T_1). (b) Discharging interval (T_2). (c) Measured resistance R_X. For each plot, deviation (%) is calculated from the analytical model equations (7), (13), and (15) respectively.

shown in Fig. 5 for R_X = 0.5 Ω. The T_1 and T_2 intervals measured is 794.5 ms and 606.3 ms respectively. Hence, the experimental waveform closely resembles the simulated waveforms of Fig. 4(a-b). Thereafter, when R_X varies across the entire span (0.1 Ω - 0.9 Ω), the corresponding I_X^{RMS} varies from 97.95 μA to 71.95 μA in simulation, while in experiment, it varies from 98.09 μA to 71.75 μA. Similarly, in case of I_C^{RMS}, it varies from 131.32 μA to 109.45 μA in simulation, while in experimental, it varies from 131.18 μA to 109.24 μA. The current I_2 remains constant throughout the span with a value of 10.15 μA.

Fig. 6a shows the simulation and experimental plots between R_L and R_X. The percentage deviation (%) for simulation and experimental is also depicted, and is calculated by comparing it with the analytical model (8) using formula $((\mid R_L - R_L^* \mid)/(R_L)) \times 100$ where R_L^* is the measured lead resistance. The average deviation in simulation and experimental is 0.43% and 0.67% respectively. Fig. 6b shows the simulation and experimental plots showing the relation between I_X^{RMS} and R_X as per (10). The average and maximum deviation in simulation is 0.57% and 0.71% while in experimental it is 0.7% and 0.9% respectively. Fig. 6c shows the simulation and experimental plots for the relation between I_C^{RMS} and R_X. The average and maximum deviation in simulation is 0.41% and 0.59% while in experimental it is 0.66% and 0.83% respectively. Again, (37) is used to calculate the rms value of I_C^{RMS} analytically.

The novelty of this work lies in the use of a T-network in the feedback path to measure R_X particularly when it is below 10 Ω, a task that is typically challenging with conventional dual-slope integrator based measurement techniques [15]. In conventional methods, R_X is usually placed at the input side of the op-amp integrator, and keeping it below 10 Ω, allows a large amount of current to flow, causing the op-amp to saturate. Even reducing the input voltage to 0.01 V does not resolve this, as the input current would still be in the range of 100 mA. Alternatively, connecting R_X in feedback produces a low output voltage that is susceptible to inherent circuit noise, leading to significant measurement errors. To address these challenges, a T-network is introduced. This design allows a permissible amount of current to flow through the input side of the op-amp while directing most of the current through the middle arm of the T-network. The I_X through R_X in the T-network is responsible for amplifying the I_C which charges and discharges the integrating capacitor C_1, and hence measure R_X. The I_X due to T-network enables faster R_X measurement in a two-wire configuration with accurate results.

The simulation and experimental plots of T_1 and T_2 for R_X is shown in Fig. 7a and Fig. 7b respectively. The T_1 remains almost constant while T_2 varies linearly with R_X for the full span. These plots closely align themselves with the prediction in the analytical model (7) and (13) which are also used for finding the deviation %. For T_1, the average deviation is 0.04% in simulation, and 0.06% in experimental. For T_2, the average

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:52:49 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Instrumentation and Measurement. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIM.2025.3560721

and maximum deviations are 0.09% and 0.15% respectively in simulation while in experiment, the average and maximum deviations are 0.11% and 0.28% respectively.

Fig. 7c shows the simulation and experimental plots of known R_X with measured R_X along with error %. The plot also shows the nonlinearity (NL) with a maximum value of 0.54% in simulation and 0.66% in experimental. In simulation, the average error % and maximum error % throughout the spam is 0.48% and 0.77% respectively while in experimental, the average error % is 0.68%. Thus, the experimental results presented in Fig. (5-7), along with the preceding discussion, demonstrate that the proposed digitizer circuit performs effectively and is well-suited for measuring low-value resistances.

VI. Conclusion

This study thoroughly examines the proposed digitizer circuit which produces digital output for measuring low resistance (R_X) in the range of 0.1 - 0.9 Ω (full span). The circuit features a dual-slope integrator with a T-network in the feedback path, three SPDT switches, and a comparator. The R_X is connected to the middle branch of the T-network, while a fixed resistor and a feedback capacitor are connected to the other two branches. The incoming current (I_X, 98.09 - 71.75 μA) in the T-network through R_X is responsible for amplifying the feedback (capacitor) current (I_C) which charges and discharges the integrating capacitor. The R_L connected to R_3 (or R_X) is measured during the integrator's charging interval. During the discharging interval, the combined resistance ($R_X + R_L$) is measured. Finally, R_L is subtracted to accurately determine R_X which is independent of R_L in a two-wire type configuration. The total measurement time ranges from 1.29 - 1.51 s, and the average error is $\pm$ 0.68% over the full span. Therefore, the proposed digitizer circuit is suitable for measuring low value resistances.

VII. Acknowledgement

The authors extend their gratitude to the Advanced I&M Laboratory of the Instrumentation and Control Engineering Department of the NIT Trichy, Tiruchirappalli, TN, India, for providing the facilities necessary for this research.

References

[1] E. W. Golding and F. C. Widdis, *Electrical measurements and measuring instruments*. Pitman, 1963.

[2] F. Galliana, P. Capra, and E. Gasparotto, "A traceable technique to calibrate dc current shunts and resistors in the range from 10 $\mu\Omega$ to 10 mΩ," *Measurement*, vol. 46, no. 4, pp. 1630–1640, 2013.

[3] T. Dauphinee, "An isolating potential comparator," *Canadian Journal of Physics*, vol. 31, no. 4, pp. 577–591, 1953.

[4] L. J. Verheyden and G. W. Akkermans, "An application of the potentiometer method for measurements of small electrical resistances," *IEEE Transactions on Instrumentation and Measurement*, vol. 17, no. 1, pp. 47–48, 1968.

[5] M. MacMartin and N. Kusters, "A direct-current-comparator ratio bridge for four-terminal resistance measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 15, no. 4, pp. 212–220, 1966.

[6] A. B. David, *Electronic instrumentation and measurements*. Prentice Hall, 2003.

[7] I. Warshawsky, "Multiple-bridge circuits for measurement of small changes in resistance," *Review of Scientific Instruments*, vol. 26, no. 7, pp. 711–715, 1955.

[8] H. Sasaki, H. Nishinaka, and K. Shida, "Automated measurement system for 1-omega standard resistors using a modified wheatstone bridge," in *Conference on Precision Electromagnetic Measurements*. IEEE, 1990, pp. 122–123.

[9] A. D. Helfrick and W. D. Cooper, *Modern electronic instrumentation and measurement techniques*. Prentice Hall Englewood Cliffs, NJ, 1990.

[10] A. Sawhney and P. Sawhney, *A course in Electrical and Electronic Measurements and Instrumentation*. Dhanpat Rai & Company, 2016.

[11] B. P. Kibble and D. Legg, "A generalized wheatstone bridge for comparing 1-Ω resistors," *IEEE Transactions on Instrumentation and Measurement*, no. 2, pp. 282–284, 1985.

[12] J. M. Williams and P. Kleinschmidt, "A cryogenic current comparator bridge for resistance measurements at currents of up to 100 a," *IEEE Transactions on Instrumentation and Measurement*, vol. 48, no. 2, pp. 375–378, 1999.

[13] E. Houtzager and G. Rietveld, "Automated low-ohmic resistance measurements at the $\mu\Omega/\Omega$ level," *IEEE Transactions on Instrumentation and Measurement*, vol. 56, no. 2, pp. 406–409, 2007.

[14] H. Yoshida and K. Takahashi, "Measurement of contact resistance with microampere currents," *IEEE Transactions on Instrumentation and Measurement*, vol. 39, no. 5, pp. 711–714, 1990.

[15] N. M. Mohan, B. George, and V. J. Kumar, "A novel dual-slope resistance-to-digital converter," *IEEE Transactions on instrumentation and measurement*, vol. 59, no. 5, pp. 1013–1018, 2010.

[16] G. Singh, S. Sohail, U. Mangalanathan, U. Gandhi, and T. Islam, "A novel dual-slope resistance to digital converter with lead resistance compensation," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–10, 2023.

[17] G. Singh, S. Sohail, A. D. Joshi, and T. Islam, "Performance evaluation of dual-slope digitizer for self-heating compensation in resistive sensor," *IEEE Transactions on Instrumentation and Measurement*, 2024.

[18] G. Singh, S. Sohail, and T. Islam, "A dual-slope rdc using t-network for low resistance measurement," in *2023 IEEE Applied Sensing Conference (APSCON)*. IEEE, 2023, pp. 1–3.

[19] N. M. Mohan, B. George, and V. J. Kumar, "Analysis of a sigma–delta resistance-to-digital converter for differential resistive sensors," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 5, pp. 1617–1622, 2009.

[20] J. A. Hidalgo-López, "Sigma-delta approach in direct interface circuits for readout of resistive sensors," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–8, 2022.

[21] K. Elangovan and A. C. Sreekantan, "Evaluation of new digital signal conditioning techniques for resistive sensors in some practically relevant scenarios," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–9, 2021.

[22] J. A. Hidalgo-López, J. A. Sánchez-Durán, and Oballe-Peinado, "Method to reduce quantization error in direct interface circuits for resistive sensors," *IEEE Sensors Journal*, vol. 20, no. 23, pp. 13 910–13 918, 2020.

[23] J. A. Hidalgo-López, Oballe-Peinado, J. Castellanos-Ramos, J. C. Tejero-Calado, and F. Vidal-Verdú, "Wide range calibration method for direct interface circuits and application to resistive force sensors," *IEEE Sensors Journal*, vol. 21, no. 20, pp. 22 956–22 966, 2021.

[24] S. Sen, T. Pan, and P. Ghosal, "An improved lead wire compensation technique for conventional four wire resistance temperature detectors (rtds)," *Measurement*, vol. 44, no. 5, pp. 842–846, 2011.

[25] K. Elangovan and A. C. Sreekantan, "An efficient digital readout for four-lead resistance thermometers," *IEEE Sensors Letters*, 2023.

[26] "Data sheet—max4690," 1.25 Ohm, Dual SPST, CMOS Analog Switches, Maxim Corp, Accessed on: 20 August 2023. [Online]. Available: https://www.analog.com/media/en/technical-documentation/data-sheets/MAX4680-MAX4700.pdf

[27] "Data sheet—tl072," low-Noise FET-Input Operational Amplifiers, Texas Instruments, Accessed on: 04 June 2024. [Online]. Available: https://https://www.ti.com/lit/ds/symlink/tl072.pdf

[28] "Data sheet—sn7404," hex inverters, Texas Instruments, Accessed on: 20 December 2024. [Online]. Available: https://www.ti.com/lit/ds/symlink/sn74ls04.pdf

Gopal Singh (Graduate Student Member, IEEE) received his bachelor's degree in instrumentation and control engineering from the Institute of Engineering and Rural Technology (IERT), Prayagraj, India, in 2017, and his master's degree in control and instrumentation from Delhi Technological University (DTU), New Delhi, in 2020. He is currently a Research Scholar at NIT Tiruchirappalli, specializing in sensors, measurements, and electronic instrumentation.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:52:49 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

Analyzing Digital Image Anti-Forensics Techniques and Countermeasures

Saurabh Agarwal[1], Rajeev Kumar[2,3], Abha Devi[4], Ki-Hyun Jung[3]
[1]Department of Information and Communication Engineering, Yeungnam University, Gyeongsan 38541, Republic of Korea
[2]Blockchain Technology Research Lab, Department of Computer Science and Engineering, Delhi Technological University, New Delhi, 110042, India
[3]Department of Software Convergence, Andong National University, Gyeongbuk, 36729, Republic of Korea
[4]Department of English, JS(PG) College, Sikandrabad, Uttar Pradesh, India
saurabh@yu.ac.kr, rajeevkumar@dtu.ac.in, abhachauhan17@gmail.com, kingjung@anu.ac.kr

***Abstract*—With the increasing use of digital media, digital forensics has become a crucial method for ascertaining authenticity and detecting manipulations in images. However, intentional anti-forensic manipulations to hide forensic clues have made verification more difficult. The efforts to develop countermeasures against anti-forensics have contributed to the evolution toward robust techniques in the field of anti-forensics; hence, the development of anti-forensic methods and their respective countermeasures will remain continuous. This work provides a scientometric analysis, a powerful tool for the study of the growth of research output in accordance with time and citation, author, organizational, and keyword co-occurrence analysis. It enhances state-of-the-art analysis by underlining highly influential authors, organizations, and works and stipulates knowledge gaps. This would consequently provide a platform for building a community where insights and inclusiveness among researchers, industry practitioners, and policymakers would all come together to guide their strategies and decisions for innovation in this dynamic field.**

***Index Terms*—Image Forensics, Scientometric Analysis, Image Forgery, Fake Image**

I. Introduction

Mobile phones and computers have become an integral part of life. They correspondingly improve with rapid development in information and communication technology. They bear big personal data in digital format, such as photographs, videos, and messages. They are more susceptible to tampering by resorting to available multimedia tools such as Adobe Photoshop, Corel Photo-Paint, and ACDSee. The ease of accessibility with which such images are manipulated poses a severe challenge. Manipulated images may result in defamation of character or tampered evidence in courts of law. In fact, the distribution of manipulated images can also lead to social unrest or political upheaval. Thus, the authenticity and integrity of digital images are very crucial in today's digital era, pertaining to image forgery attacks like image retouching, image splicing, copy-move, etc. Because of this, two primary methods can be used for digital image forgery detection: an active approach and a passive approach. In an active approach, specific details are embedded in the image so that it can detect even the slight form of digital tampering. Examples of such can include name, date, signature, and other meta information of the image. This often requires a peculiar system to verify the authenticity of the digital image. In passive methods, duplicated objects in manipulated images are detected without using the original image watermark. Instead, the technique looks for traces left behind due to dissimilar steps involved in the image manipulation process. It can also locate the place and region where the forgery has taken place within the image. The source camera identification and alteration trace analyses are included in passive digital image forensics that verifies image manipulation. These methods take advantage of unique fingerprints left by the acquisition devices or apply signal analysis in order to disclose forgeries. However, these verifications have been complicated by the emergence of anti-forensic techniques that mask traces of manipulation. As a result of these developments, countermeasures against anti-forensics have driven further advancement of robust forensic techniques. This dynamic evolution keeps shaping anti-forensic methods and their countermeasures in the dynamic field of digital image forensics.

Sitara and Mehtre [1] discussed a survey of the anti-forensic domain through passive techniques for image and video tampering detection, along with their respective countermeasures. The authors also pointed out different limitations that have caused anti-forensic methods to grow along with digital forensics. Taneja et al. [2] presented a study related to anti-forensic attacks on JPEG compression, contrast enhancement, and median filtering. Their work presented a robust discussion on methodologies and summarized some existing state-of-the-art methods in tabular form. Al-Fehani and Al-Kuwari [3] discussed image and video forensics, including source identification, forgery detection, and problems arising from generative media. This work has stressed the attacks' liability in using the adversarial and, in turn, motivating the need for more robust methods of forgery detection. They have mentioned the generalized detection technique for forgery and the under-representation of research in GAN detection, in particular

ISBN: 979-11-88428-13-7 Copyright 2025. Global IT Research Institute (GIRI). All right reserved.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 05:53:40 UTC from IEEE Xplore. Restrictions apply.

for image forensics. They have also outlined that video anti-forensics, compared to the image, is underrepresented in literature and countermeasures, too. They also talked about the legality of forensic evidence and were motivated to emphasize the justification of the authenticity of the evidence. Some problems arise from the fact that deep learning has started to replace traditional solutions in the understanding of decisions taken by classifiers, being an issue for forensic experts. Their overall review highlighted the changing nature of multimedia forensics and the need for adaptive methods in order to meet several new challenges.

Qureshi and El-Alfy [4] presented a bibliographic review about anti-forensics and relative countermeasures; the goal was to help researchers and practitioners find their place in the current panorama. This review underlines that countering AFMs is an essential task for making forensic techniques reliable against tampering with images. It points out the never-ending conflict between forensic and anti-forensic techniques, which imposes the necessity of constant development and adaptation. The authors firmly believe that new computational methods, such as the addition of dithering noise, have to be explored in order to find adequate countermeasures against emerging anti-forensic strategies. This investigation is based on the finding that the quality of images essentially defines forensic trace concealment. Moreover, they recommend the extension of anti-forensics and related countermeasures research to new fields of video, disk, and network forensics. The work raises questions concerning the reliability and trustworthiness of the existing forensic techniques, proposing a study of relationships among several picture artifact metrics for devising more secure methods. It advocates for thorough benchmark systems, embedding recent deep learning developments, and multi-sensor cameras that will improve forensic capabilities. Finally, the review discusses how cloud computing could support forensic solutions in a scalable way in the future.

More research is needed in the domain of anti-forensics. Only a few papers [1]–[4] present organized research on different anti-forensic schemes. Paper [2] discusses anti-forensic attacks on JPEG, contrast enhancement, and median filtering. It introduces image forgery, JPEG compression, contrast enhancement, and median filtering. After presenting the detailed anti-forensic methods, it shows the current state-of-the-art methodologies in a tabular form for clarity. Al-Fehani and Al-Kuwari, [3] in their work, review some of the recent works on image and video forensics that involve source identification, forgery detection, and challenges from generative media. Anti-forensics and countermeasures will be presented, focusing on the general problem of susceptibility to adversarial attacks. The review also highlights the need for more research on GAN detection, especially in the context of image forensics. Robustness in forgery detection methods is needed. It underlines a promising generalized forgery detection technique and calls for more research related to GAN detection, especially for image forensics. This paper points out that, compared to image forensics, the literature dealing with video anti-forensics and related countermeasures needs to be revised. It also considers in some depth legal issues with respect to forensic evidence and emphasizes the need for a description of the rationale for the authenticity of evidence. This movement into deep learning methods is at the price of well-understood classification decisions that might concern forensic practitioners. In general, it highlights the dynamic nature of multimedia forensics and the implication for adaptive approaches towards rising challenges. A review [4] on anti-forensics and countermeasures against digital image forensics gives an overview to help researchers and practitioners better understand the status quo in this arena. Therefore, countering AFMs is of vital importance to guarantee that forensic techniques against image tampering are reliable. It realizes the never-ending battle between forensic and anti-forensic techniques and further identifies the need for development and adaptation. It points out that new computational methodologies, such as adding noise with dithering, should be explored to counter emerging anti-forensic strategies successfully. Critical roles of image quality in masking the forensic traces are outlined in this work, and further investigation of the topic is presented. The paper also stresses the importance of extending research studies on anti-forensics and countermeasures to new emerging areas like video, disk, and network forensics, as this will contribute to the development of more secure forensic techniques. Furthermore, it suggests developing comprehensive benchmark systems, integrating deep learning advancements, and using multi-sensor cameras to improve forensic capabilities further. It finally discusses how cloud computing could enable scalable forensic solutions in the near future—surveys [1] anti-forensic techniques relating to passive image and video tampering detection, discussing countermeasures. Limitations within the current literature pertaining to tampering detection are noted, which have led to the development of anti-forensic methods, changing with digital forensics.

We previously examined key survey papers concerning antiforensics and its countermeasures. In the next section, we will delve into Scientometric Analysis.

II. Scientometric Analysis

This study will help explain productivity, seminal articles, productive authors, and organization and collaboration dynamics in the domain of anti-forensics. Scientometric analysis in research output related to anti-forensics encompasses several steps methodologically: the selection of appropriate bibliographic databases, formulating the search queries, retrieving the literature, and applying scientometric indicators quantifying the impact of research and collaboration patterns. It would naturally also involve a well-articulated search strategy to identify relevant publications, preprocessing for quality and consistency of data, computation of scientometric indicators such as counts of citations, impactful authors networks, and several visualization techniques to interpret.

ISBN: 979-11-88428-13-7
Copyright 2025. Global IT Research Institute (GIRI). All right reserved.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 05:53:40 UTC from IEEE Xplore. Restrictions apply.

A. Data Collection and Publication Trend

We have developed an exhaustive dataset of research publications from the WoS repository, one of the world-renowned databases containing exhaustive and quality research content for comprehensive analysis and synthesis of the anti-forensics field. The database includes all major indices, such as SCIE, SSCI, and BKCI-S. The search string (("antiforensic*" OR "counterforensic*" OR "anti-forensic*" OR "counter-forensic*" OR ((anti OR counter*) NEAR/3 (forensic*))) AND (JPEG OR image)) yielded 220 publications dealing with research. Manual screening for non-relevant entries brought us to the number of 210 relevant publications that deal with the area. Figure 1 describes the annual progress of anti-forensics research publications that had cropped up as early as 2006. From 2013 onwards, the domain continued to show a significant number of publications throughout the year, except for 2016.

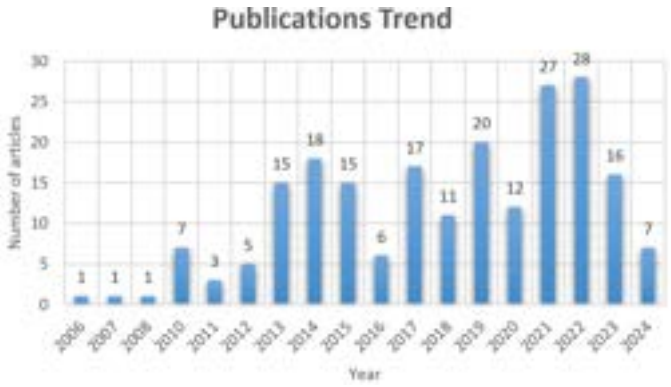


Figure 1. Yearly Publications of Anti-Forensics Research

B. Identification of Top Authors and Organizations with their country

Table I highlights the prominent researchers in the field of anti-forensics, showcasing a diverse group with significant contributions from China (four researchers), Italy (two researchers), and one each from the USA, India, Saudi Arabia, and South Korea. Barni, M. from Italy, leads the list of top contributing authors with six articles. He is followed by Stamm, MC from the USA, and Wu, JY from China, each authoring five articles.

Our analysis of the top contributing organizations, as per Table II, underscores the significant contributions of researchers from China, Italy, the USA, and India to the field of anti-forensics. Univ Maryland, Dept Elect and Comp Engn, College Pk (USA) has published the highest number of articles (8), followed by Beijing Jiaotong Univ, Inst Informat Sci, Beijing (China) with six articles, Drexel Univ, Dept Elect and Comp Engn, Philadelphia (USA) with five articles, and Sun Yat Sen Univ, Sch Informat Sci and Technol, Guangzhou (China) with four articles. Other organizations such as the Univ Siena, Dept Informat Engn and Math (Italy), TIET, Dept Elect and Commun Engn, Patiala, Punjab (India), Univ Siena, Dept Informat Engn and Math, Siena (Italy), Beihang Univ, Sch Comp Sci and Engg. (China), Shenzhen Univ, Shenzhen Key Lab Media Secur, Shenzhen (China), and Shenzhen Univ, Guangdong Key Lab Intelligent Informat Proc, Shenzhen (China) have also made significant contributions.

TABLE I
TOP AUTHORS WITH THEIR COUNTRY

Sr. No.	Author	Country	Publications
1	Barni, M	Italy	6
2	Stamm, MC	USA	5
3	Wu, JY	China	5
4	Chen, C	China	4
5	Fahmy, G	Saudi Arabia	4
6	Uddin, K	South Korea	4
7	Chen, ZP	China	3
8	Singh, G	India	3
9	Marra, F	Italy	3
10	Cao, G	China	3
11	Dou, LY	China	3
12	Valenzise, G	Italy	3
13	Ding, F	China	3
14	Kumar, A	India	3
15	Zeng, H	China	3
16	Sameer, VU	India	3
17	Fan, W	France	2
18	Bharathiraja, S	India	2
19	Berthet, A	France	2

TABLE II
TOP INSTITUTES WITH NUMBER OF PUBLICATIONS

Sr. No.	Institute	Publications
1	Univ Maryland, Dept Elect & Comp Engn, College Pk, MD 20742 USA	8
2	Beijing Jiaotong Univ, Inst Informat Sci, Beijing 100044, China	6
3	Drexel Univ, Dept Elect & Comp Engn, Philadelphia, PA 19104 USA	5
4	Sun Yat Sen Univ, Sch Informat Sci & Technol, Guangzhou, China	4
5	Univ Siena, Dept Informat Engn & Math, I-53100 Siena, Italy	3
6	TIET, Dept Elect & Commun Engn, Patiala, Punjab, India	3
7	Univ Siena, Dept Informat Engn & Math, Siena, Italy	3
8	Beihang Univ, School of Comp Sci & Engn, Beijing 100191, China	3
9	Shenzhen Univ, Shenzhen Key Lab Media Security, Shenzhen, China	3
10	Shenzhen Univ, Guangdong Key Lab Intelligent Info Proc, Shenzhen, China	3

C. Document citation linkages

A full-scale VoS analysis was performed on the WoS corpus for research related to anti-forensics, which covered 210 published documents. This paper analyzed the citations among those documents using a full counting method that had to be at least two citations each. It identified the top ten cited documents, including their total citations, publication venues, and the years of publication. The paper spearheads the list "Do GANs Leave Artificial Fingerprints?" with 127

ISBN: 979-11-88428-13-7
Copyright 2025. Global IT Research Institute (GIRI). All right reserved.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 05:53:40 UTC from IEEE Xplore. Restrictions apply.

TABLE III
MOST CITED ARTICLES WITH PUBLICATION VENUE AND NUMBER OF CITATIONS

No.	Authors	Title	Venue	Year	Citations
1	Marra et al.	Do GANs Leave Artificial Fingerprints?	IEEE MIPR, 2019	2019	[5]
2	Cao et al.	Forensic detection of median filtering	IEEE ICME, 2010	2010	[6]
3	Ansari et al.	Fluorescent materials for fingermarks detection	Coord. Chem. Rev., Elsevier	2022	[7]
4	Stamm et al.	Anti-forensics of JPEG compression	IEEE ICASSP, 2010	2010	[8]
5	Ding et al.	Anti-Forensics for Face Swapping Videos	IEEE Trans. on Multimedia	2021	[9]
6	Stamm et al.	JPEG Compression Anti-forensics	IEEE ICIP	2010	[10]
7	Cao et al.	Anti-Forensics of Contrast Enhancement	MM&Sec '10	2010	[11]
8	Fan et al.	JPEG anti-forensics via variational approach	IEEE ICASSP, 2013	2013	[12]
9	Wu et al.	Anti-forensics of median filtering	IEEE ICASSP, 2013	2013	[13]
10	Valenzise et al.	JPEG compression anti-forensics cost	IEEE ICASSP, 2011	2011	[14]

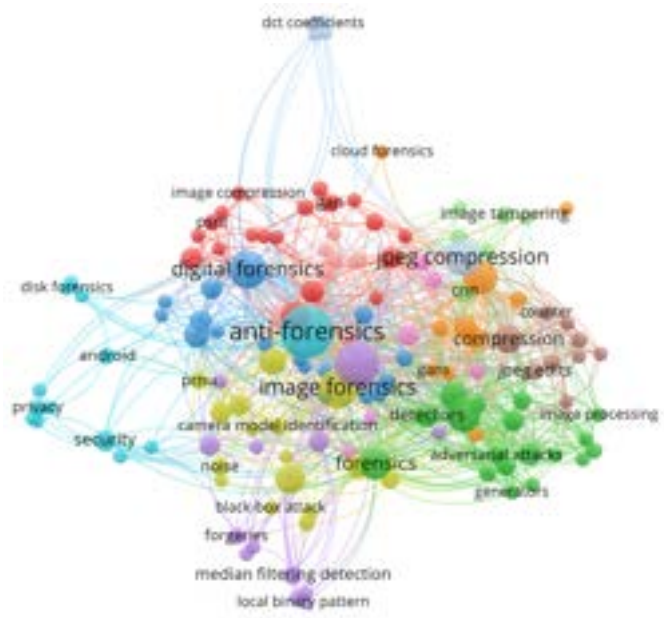


Figure 2. Keyword Co-occurrence Networks in Anti-Forensics Research

citations. Next is "Forensic Detection of Median Filtering in Digital Images" by Cao et al., which secured the second position with 107 citations. The document "Recent Progress of Fluorescent Materials for Fingermarks Detection in Forensic Science and Anti-counterfeiting" by Ansari et al. has been cited 88 times, hence placing third. The last five documents in Table III have less than 50 citations each. A key observation from this analysis is the crucial role of IEEE and conferences in promoting highly influential research work in the anti-forensics domain. The majority of high-impact papers are by IEEE, and a large fraction of them are conference papers. This underlines key works and highly influential papers that establish these publications as essential milestones in pushing the research frontier on anti-forensics.

D. Keyword Co-occurrence Networks

Figure 2 illustrates the co-occurrence networks of keywords in the anti-forensics domain based on a comprehensive analysis of the Web of Science (WoS) document corpus. The VoS search identified 649 unique keywords, with the study focusing on 140 keywords that appeared at least twice.

The co-occurrence grids are color-coded to represent different clusters:

- Cyan: The keyword "anti-forensics" forms the most significant node, frequently co-occurring with "privacy," "security," and "disk forensics."
- Blue: This network centers around "digital forensics," indicating a significant research thrust in this area.
- Yellow: Keywords related to "image forensics" dominate this cluster, highlighting its prominence in the field.
- Green: The "forensics" keyword forms the core of this cluster, encompassing a broad range of related topics.
- Smaller Clusters: These include frequent keywords such as "CNN," "cloud forensics," "forgery," "local binary pattern," and "compression," each forming distinct, smaller clusters.

The visualization provides insights into the interconnectedness and primary focus areas within the anti-forensics research landscape.

III. CONCLUSION

This scientometric analysis explains the overall growth and path of anti-forensic research, outlining the significant trends and contributions that have set a mark in the area. Growth in publication since 2006 steepens in 2013, reflecting the rising importance and complexity of digital forensics as a response against advanced challenges caused by digital manipulation. In this respect, the study identifies significant contributors from institutions worldwide, especially in the USA, China, and Italy, which are driving the research and development of new methods to respond to both established and new threats. Citation analysis has shown how keynote works, published for the most part through IEEE venues, attest to the quality of research and the pivotal role of IEEE conferences in the dissemination of critical anti-forensic studies. Keyword analysis returns a dynamic network of themes, such as "privacy," "security," and "forensics," which is necessary to understand the focus of the field and to integrate diverse research areas into effective anti-forensic strategies.

ACKNOWLEDGMENT

This research was supported by Brain Pool program funded by the Ministry of Science and ICT through the National

ISBN: 979-11-88428-13-7
Copyright 2025. Global IT Research Institute (GIRI). All right reserved.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 05:53:40 UTC from IEEE Xplore. Restrictions apply.

Research Foundation of Korea (2021H1D3A2A01099390) and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2021R1I1A3049788).

REFERENCES

[1] K. Sitara and B. M. Mehtre, "Anti-forensics for image and video tampering: A review," in *Cryptographic and Information Security*. Boca Raton, Florida: CRC Press, [2019], pp. 799–827.

[2] N. Taneja, V. S. Bramhe, D. Bhardwaj, and A. Taneja, "Understanding digital image anti-forensics: an analytical review," *Multimed Tools Appl*, vol. 83, no. 4, pp. 10 445–10 466, Jan. 2024.

[3] M. Al-Fehani and S. Al-Kuwari, "Recent advances in digital image and video forensics, anti-forensics and counter anti-forensics," Feb. 2024, [Online]. Available: http://arxiv.org/abs/2402.02089.

[4] M. A. Qureshi and E. M. El-Alfy, "Bibliography of digital image anti-forensics and anti-anti-forensics techniques," *IET Image Process*, vol. 13, no. 11, pp. 1811–1823, Sep. 2019.

[5] F. Marra, D. Gragnaniello, L. Verdoliva, and G. Poggi, "Do gans leave artificial fingerprints?" in *2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*. IEEE, 2019, pp. 506–511.

[6] G. Cao, Y. Zhao, R. Ni, L. Yu, and H. Tian, "Forensic detection of median filtering in digital images," in *2010 IEEE International Conference on Multimedia and Expo*. IEEE, 2010, pp. 89–94.

[7] A. A. Ansari, K. M. Aldajani, A. N. AlHazaa, and H. A. Albrithen, "Recent progress of fluorescent materials for fingermarks detection in forensic science and anti-counterfeiting," *Coordination Chemistry Reviews*, vol. 462, p. 214523, 2022.

[8] M. C. Stamm, S. K. Tjoa, W. S. Lin, and K. J. R. Liu, "Anti-forensics of jpeg compression," in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, Mar. 2010, pp. 1694–1697.

[9] S. Bharathiraja and B. Rajesh Kanna, "Anti-forensics contrast enhancement detection (afced) technique in images based on laplace derivative histogram," *Mobile Networks and Applications*, vol. 24, no. 4, pp. 1174–1180, Aug. 2019.

[10] M. C. Stamm, S. K. Tjoa, W. S. Lin, and K. R. Liu, "Undetectable image tampering through jpeg compression anti-forensics," in *2010 IEEE International Conference on Image Processing*. IEEE, 2010, pp. 2109–2112.

[11] G. Cao, Y. Zhao, R. Ni, and H. Tian, "Anti-forensics of contrast enhancement in digital images," in *Proceedings of the 12th ACM Workshop on Multimedia and Security*, 2010, pp. 25–34.

[12] W. Fan, K. Wang, F. Cayre, and Z. Xiong, "A variational approach to jpeg anti-forensics," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, May 2013, pp. 3058–3062.

[13] Z.-H. Wu, M. C. Stamm, and K. J. R. Liu, "Anti-forensics of median filtering," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, May 2013, pp. 3043–3047.

[14] G. Valenzise, M. Tagliasacchi, and S. Tubaro, "The cost of jpeg compression anti-forensics," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, May. 2011, pp. 1884–1887.

ISBN: 979-11-88428-13-7 Copyright 2025. Global IT Research Institute (GIRI). All right reserved.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 05:53:40 UTC from IEEE Xplore. Restrictions apply.

2025 International Conference on Pervasive Computational Technologies (ICPCT)

Anomaly Detection for Sensor Manipulation in Matter Enabled-IoT Devices with Faulty Data Injection

Jahnvi Patel
jahnvipatel1405@gmail.com
Department of Information Technology
Delhi Technological University
Delhi, India

Kapil Sharma
kapil@ieee.org
Department of Information Technology Delhi
Technological University
Delhi, India

2025 International Conference on Pervasive Computational Technologies (ICPCT) | 979-8-3315-0868-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICPCT64145.2025.10940262

***Abstract*—The security of Internet of Things (IoT) devices in Matter-enabled networks is crucial, as attacks such as sensor spoofing and data injection can lead to significant disruptions. Sensor manipulation attacks where the attackers manipulate the sensor inputs to create erroneous behavior in devices are issues of serious concern for anomaly detections. Faulty sensor data can cause devices to operate in unsafe or inefficient conditions across various applications, compromising system reliability and performance. The Isolation Forest algorithm is employed to isolate faulty data injections as outliers in sensor readings, offering a robust method for anomaly detection. The approach, therefore, focuses on the possibility of enhancing detection with greater reliability in IoT networks. The lightweight and extensible approach also assembles a process of strengthening Matter-based networks against vulnerabilities emanating from wrong control feedback.**

***Index Terms*—Matter Protocol, Internet of Things Security, Anomaly Detection, Interoperability**

I. Introduction

The expansion of the Internet of Things (IoT) has transformed industries, offering smarter solutions in areas such as home automation, healthcare, large-scale industrial systems, and smart cities. This growth, however, brings many security challenges as IoT devices generally provide open access points that can serve as gateways for various attack methods such as hacking, malware infection, phishing, and spoofing. These threats can result in unauthorized access, data manipulation, and denial of service (DoS) attacks [1], undermining the reliability of the devices and networks they operate within. In Matter-enabled IoT networks, the risk is further amplified by the mesh networking protocol [2], which facilitates communication across devices on different platforms. Matter [3], an open-source and royalty-free connectivity standard made with the intention of ensuring secure, reliable, and seamless communication between IoT devices, potentially opens new avenues for data manipulation since devices operate as relay nodes and can be exploited through vulnerabilities in the sensor systems.

Sensor spoofing attacks are a major potential threat to the security of IoT networks. Here, an unauthorized user manipulates sensor readings so as to interfere with the proper functionality of the devices. Such attacks make the systems prone to incorrect data injection, causing malfunction and posing potential safety risks to critical systems. The challenge of detecting these anomalies in real-time is crucial to maintaining the integrity of Matter-enabled networks.

To address this issue the authors have proposed the following key contributions:

- A real-time IoT network anomaly detection using an Isolation Forest algorithm-based ML security framework that secures the network.
- Provided a scalable and resource-efficient solution, making it suitable for deployment in resource-constrained IoT environments.
- Ensured the integrity of Matter-enabled networks by distinguishing legitimate sensor data from malicious alterations.

This approach not only facilitates the immediate detection and mitigation of security threats but also ensures efficient operation in resource-constrained IoT environments. It maintains data integrity, minimizes the harmful impact of data manipulation, and strengthens the resilience of Matter-enabled IoT networks. It guarantees real-time anomaly detection of sensor data that can withstand adversarial tampering, ensuring the reliability and security of the underlying network.

Section III presents a machine learning-based approach leveraging the Isolation Forest algorithm to address security challenges. It includes a pipeline diagram illustrating the data collection process and the machine learning algorithm's steps. Section IV presents the performance comparison of the proposed methodology, mentioning the highly favorable results on the limited available resources and looks into the scalability alongside the computational efficiency in diverse IoT environments. Section V provides a comprehensive discussion of the problem and the proposed solution, including limitations and potential areas for future work. Finally, a summary of the paper is offered in Section VI together with potential avenues for research in enhancing reliability and safety in IoT networks.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:04:49 UTC from IEEE Xplore. Restrictions apply.

II. Literature Review

Anomaly detection plays a pivotal role in securing and ensuring the integrity of IoT systems, especially in the Matter-enabled devices. The Matter protocol is developed by the Connectivity Standards Alliance (CSA) [4], to provide a secure and interoperable platform for IoT devices. However, like all networked systems, Matter-enabled devices are susceptible to manipulation attacks, including faulty data injection and sensor tampering, thus reducing the integrity and system faults.

A. Vulnerabilities in IoT Networks

IoT networks, based on Matter, are usually designed in the form of a mesh topology [5], where each device communicates with multiple other devices. This decentralized communication system is vulnerable to various types of attacks. A compromised device can inject faulty data, alter sensor readings, or manipulate packets. For example, a compromised sensor could provide incorrect readings to the system, leading to abnormal system behavior. Such manipulations can lead to the inefficiencies of operations or a failure-to happen in some smart homes or within industrial systems of automation.

The risks associated with IoT networks have been extensively studied. The faults data injection attacks disrupt communication between devices, and they allow false information to be distributed through the network, causing instability in device functions. As pointed out by [6], such manipulation might stunt systems reliant on sensor data, such as with the HVAC systems, where manipulated temperature data make the equipment malfunction.

B. Security Mechanisms in Matter Protocol

To maintain reliability and safety, IoT devices employ the Matter protocol with various security measures [7]. The primary security mechanism is end-to-end encryption, which secures communication and prevents eavesdropping or data alteration. Matter also includes device authentication during commissioning to prevent unauthorized access and sabotage. Granular access control policies restrict device interaction, minimizing the risk of security breaches and unauthorized data manipulation. By ensuring only trusted devices participate, the Matter protocol reduces risks from faulty data injection and sensor manipulation. Further enhancing IoT security, combining elliptic curve cryptography (ECC) with blockchain addresses privacy and unauthorized access challenges in sensitive systems like healthcare IoT. The GAO-RF model achieved 98.4% accuracy in securely handling large amounts of private data [8]. These advancements strengthen IoT protocols like Matter, providing robust mechanisms for safeguarding sensitive information in modern ecosystems.

C. Anomaly Detection in IoT Networks

Anomaly detection is fundamental for identifying attacks such as faulty data injection and sensor manipulation in IoT systems. Various machine learning techniques [9] have long been exploited to detect such anomalies in IoT environments. The application of these techniques to faulty data detection in IoT networks shows how to detect data manipulation and some form of attacks generating network congestion induced by malicious nodes [10]. Such techniques work effectively because they identify deviations in sensor data and network behavior that indicate imminent attacks. Additionally, securing IoT devices is crucial, and recent studies have explored using honeypots, like Raspberry Pi-based systems, to detect and analyze attack patterns targeting vulnerable IoT protocols [11]. A promising approach to anomaly detection in IoT networks is the use of active learning. Various observations [12] have been made for models of active learning to improve the performance and accuracy levels of anomaly detection within resource-constrained environments by training themselves on the most informative data points. The technique is especially valuable for IoT systems, where devices produce vast amounts of sensor data, making anomaly identification in (quasi-)real-time impossible without efficient methods for data selection.

D. False Data Injection Attacks and Detection Mechanisms

False data injection attacks threaten IoT systems, especially those reliant on sensor data. These attacks inject manipulated data, causing incorrect device behavior and system failures. Several detection methods identify and mitigate their impact. Machine learning models, including supervised learning, clustering, and statistical approaches, detect data patterns deviating from normal behavior, signaling potential attacks.

The importance of detecting false data injection in IoT systems have been highlighted [13], particularly in the context of smart grids, where such attacks can significantly impact system performance. Various methods have been proposed for detecting these attacks, including anomaly detection algorithms that identify irregularities in sensor data and network traffic.

III. Proposed Methodology

A. Architecture Overview

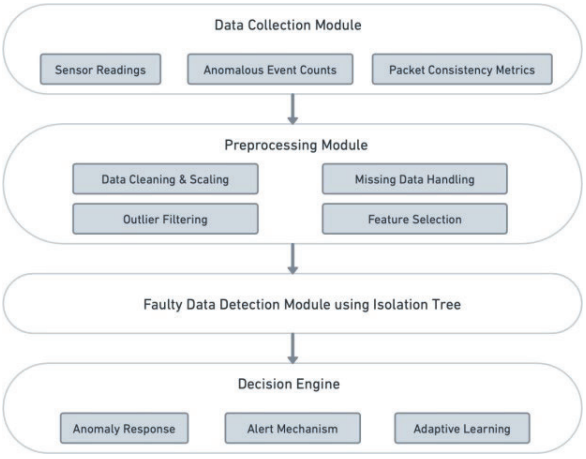


Fig. 1: Proposed Pipeline for Faulty Data Detection

Fig. 1 outlines the architecture of the IoT anomaly detection system, detailing the progression from data acquisition and

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:04:49 UTC from IEEE Xplore. Restrictions apply.

preprocessing to anomaly detection and decision-making. The architectural framework comprises:

1) Data Collection Module: The data collection module receives diagnostic clusters from the devices to collect the telemetry data. The types of data collected include:

- Sensor Readings: Temperature and humidity, sensitive data are obtained to note any variation resulting from discrepancies.
- Packet Consistency Metrics: Monitoring of packet performance in ways that can reveal the signs of alteration or insertion in the transmission medium.
- Anomalous Event Counts: Number of occurrences of decision abnormalities reflected in devices, which may indicate an attempt at manipulation or the injection of bad data.

2) Preprocessing Module: The preprocessing module ensures that data collected is clean, consistent, and prepared for analysis. Key processes include:

- Data Cleaning and Scaling: Raw data from different sources such as sensor readings, packet consistency, and event counts are normalized using StandardScaler.
- Missing Data Handling: Missing data is resolved by using mean imputation techniques so as not to distort the dataset in order to analyze it.
- Outlier Filtering: Outliers are filtered using interquartile range since they distort normal patterns and improve the extent of generalization on normal behavior of devices.
- Feature Selection: Defined the feature space using sensor readings, packet integrity, response times, and event counts, focusing on detecting faulty data injection and manipulation.

3) Faulty Data Detection Module (Anomaly Detection): This module applies the Isolation Forest algorithm for identifying the perturbation of data due to malicious attacks. The detection module identifies anomalies in sensor readings, packet integrity metrics, and event counts, signaling potential manipulation, tampering, or faulty behavior. The Isolation Forest algorithm was trained with the following parameters:

- Number of Estimators: Set to 100, specifying the number of trees in the forest. This determines how thoroughly the algorithm isolates data points to detect anomalies.
- Contamination: Set to 0.10, representing the estimated proportion of anomalies in the dataset. This value helps define the threshold for distinguishing normal from anomalous data.
- Random State: Set to 42 to ensure reproducibility. Using this value ensures that the results remain consistent across different runs of the algorithm.

4) Decision Engine: The decision engine manages responses to detected anomalies:

- Anomaly Response: Devices identified as having manipulated data are flagged, triggering immediate alerts.
- Alert Mechanism: Notifications are given to an administrator or other automated systems to take measures against suspicious activities.
- Adaptive Learning: It keeps track of predictions and results so that the model refines the levels of accuracy concerning detection based on new patterns.

B. Faulty Data Detection Algorithm Using Isolation Forest

1) Algorithm Steps:

- *Input Collection:* Generated a synthetic dataset [14] comprising 10,000 records to simulate telemetry data. This dataset includes sensor readings, packet integrity metrics and anomaly counts. The diverse input ensures comprehensive analysis for detecting inconsistencies across data sources.
- *Data Preprocessing:* Normalize the features using a StandardScaler to ensure uniformity and eliminate biases caused by varying feature scales. Remove null values and outliers, reducing the dataset to 9,470 records. Split the preprocessed dataset into 80% for training, 7,576 records and 20% for testing [15], 1,894 records to ensure reliable evaluation of the model.
- *Model Training:* Train the Isolation Forest algorithm on the preprocessed synthetic dataset. This training phase establishes a baseline for normal device behavior, making it possible to identify deviations accurately.
- *Anomaly Detection:* Apply the trained model to predict faulty data injections. Data points flagged as anomalous indicate potential manipulation, facilitating targeted response actions to protect the network's integrity.
- *Alert Generation:* Generate alerts for detected anomalies, enabling rapid intervention and restoration of normal network operations.

IV. Results and Discussion

A. Results

1) Performance of the Faulty Data Detection Methodology: The anomaly detection approach demonstrated high accuracy in identifying devices with manipulated sensor data. The classification report for the test data is as follows:

Class	Precision	Recall	F1-Score
Normal (0.0)	0.99	0.99	0.99
Anomalous (1.0)	0.95	0.94	0.95

TABLE I: Performance metrics for anomaly detection.

- Precision: 95% of the devices flagged as manipulated were correctly identified.
- Recall: The method detected 94% of all instances of faulty data injection.
- F1-Score: This metric, at 95%, highlights a strong balance between precision and recall, showcasing the model's reliability.
- Accuracy: The method correctly identified 99% of the devices in the dataset.

Table 1. demonstrates that high precision minimizes false positives, avoiding unnecessary alerts, while the strong recall supports effective detection of genuine manipulation incidents.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:04:49 UTC from IEEE Xplore. Restrictions apply.

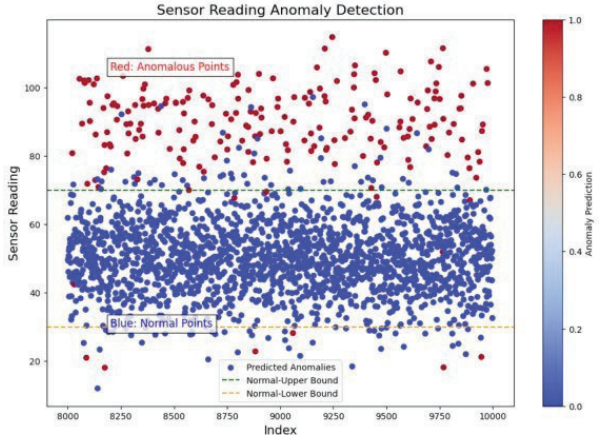


Fig. 2: Sensor Reading Anomaly Detection

Fig. 2 effectively illustrates the system's capability to differentiate between normal and anomalous sensor readings, highlighting its role in detecting deviations that could indicate potential faults, attacks, or abnormal environmental changes.

2) Feature Impact and Correlation Analysis: The correlation analysis showed that device response times had the strongest correlation (0.936) with detected manipulation events, emphasizing their critical role as an indicator of tampering. Sensor readings (0.767) also exhibited a strong positive correlation, highlighting that inconsistencies in sensor data are significant indicators of tampering. Packet integrity checks displayed a strong negative correlation (-0.838), indicating that reduced packet integrity is associated with faulty data injection scenarios.

These results imply that compromised devices tend to exhibit abnormal delays in device responses, inconsistencies in networked sensor readings, and deteriorated packet integrity metrics. These characteristics collectively point to sensor manipulation attacks or other forms of device tampering.

B. Discussion

Sensor manipulation and faulty data injection pose significant risks to Matter-enabled mesh networks, where devices may respond to false sensor readings, leading to network disruptions. Matter IoT devices, often relying on multi-hop communication, can face extensive issues when manipulated sensor data propagates throughout the network. While Matter's end-to-end encryption ensures packet contents remain secure, our findings show that faulty data injection does not require decryption of the packets. Instead, attackers manipulate sensor readings, influencing device behavior. This highlights the importance of anomaly detection that focuses on the integrity of the data rather than just the security of the packets themselves.

The authors used the Isolation Forest algorithm to detect the faulty data injections. The approach successfully identified deviations in sensor readings, learning the normal behavior of the network and flagging any anomalies. The unsupervised way is suitable for IoT, as no signature needs to be defined ahead of time, which allows adaptation. Our findings complement earlier research [16], whereby tampered sensor data were shown to compromise IoT network performance and safety of the devices. Similarly, it was discussed [17] how manipulation in one device could have network-wide repercussions in mesh networks. The studies provide endorsement for the need for robust anomaly detection systems to ensure the reliability and security of Matter-enabled IoT networks.

While Matter includes device authentication and access control to safeguard networks, these protections alone are insufficient against sensor manipulation and faulty data injection. Replay protection mechanisms help, but the intruder would still be able to inject new manipulated data in any point of the network. This aspect ultimately calls for an effective detection system.

This research underscores the effectiveness of the Isolation Forest algorithm for detecting sensor manipulation attacks in Matter-enabled IoT networks. By focusing on critical telemetry data such as sensor readings and packet integrity, the approach successfully identifies compromised devices and mitigates network disruptions. As the techniques of sensor manipulation are continually evolving, scalability, latency, and a multi-layered security approach will need to be considered for fulfilling a practical deployment of Matter IoT networks to stay ahead in these real-world scenarios.

V. Conclusion with Future Work

This paper addresses the issue of anomaly detection in sensor readings within Matter-enabled IoT networks, focusing on identifying faulty data injection and sensor manipulation. With the increasing criticality of IoT networks in various applications, real-time detection of manipulated data is essential to maintain network reliability and functionality. The proposed methodology demonstrates strong performance metrics, minimizing false positives while ensuring robust anomaly detection. These findings emphasize the importance of comprehensive sensor data monitoring to safeguard IoT networks from destabilization caused by manipulated data.

Future work will prioritize collecting high-quality telemetry data from realistic sources to enhance approach generalizability. Adaptive detection algorithms, hybrid approaches combining rule-based and machine learning techniques, and scalable, low-latency solutions will be explored to improve intrusion detection in Matter IoT environments. Cross-sensor validation and robust rate-limiting mechanisms will be integrated to mitigate the impact of faulty data injection, ensuring reliable detection across diverse application.

References

[1] Statista, "Number of Internet of Things (IoT) connected devices worldwide from 2020 to 2030," [Online]. Available: https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/. [Accessed: Jan. 11, 2025].

[2] IEEE, "Implementation of Wireless Mesh Network for IoT based Smart Homes," IEEE, 2020, doi: 10.1109/RAEE.2019.8887035.

[3] Google, "Matter Overview," Google Developers. [Online]. Available: https://developers.google.com/matter [Accessed: Jan. 11, 2025].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:04:49 UTC from IEEE Xplore. Restrictions apply.

[4] D. Belli, P. Barsocchi, and F. Palumbo, ”Connectivity Standards Alliance Matter: State of the art and opportunities,” *Internet of Things*, vol. 25, p. 101005, 2024. doi: 10.1016/j.iot.2023.101005.
[5] R. Ohlan and A. Ohlan, “Wireless Mesh Networks in IoT Networks,” presented at the 2017 International Workshop on Electromagnetics: Applications and Student Innovation Competition (iWEM), 2017, doi: 10.1109/iWEM.2017.7968828.
[6] N. Mishra and S. Pandya, ”Internet of Things Applications, Security Challenges, Attacks, Intrusion Detection, and Future Visions: A Systematic Review,” *IEEE Access*, vol. 9, pp. 59353-59377, 2021, doi: 10.1109/ACCESS.2021.3073408.
[7] L. Basyal, “IoT Security with Matter Protocol: Exploring a Secure and Reliable Home Automation,” Stevens Institute of Technology.
[8] L. Basyal, “IoT Security with Matter V. Sharma and A. Kumar, ”Digital twin: securing IoT networks using integrated ECC with blockchain for healthcare ecosystem,” *Knowledge and Information Systems*, vol. 2024, no. 1, pp. 1–32, Nov. 2024, doi: 10.1007/s10115-024-02273-6.
[9] M. Al-Garadi, A. Mohamed, A. K. Al-Ali, and X. Du, “A Survey of Machine and Deep Learning Methods for Internet of Things (IoT) Security,” July 2018, doi: 10.48550/arXiv.1807.11023.
[10] C. E. Gomez and N. M. Salim, “Security Analysis of Trust on the Controller in the Matter Protocol Specification,” presented at the 2023 IEEE Conference on Communications and Network Security (CNS), doi: 10.1109/CNS59707.2023.10288747.
[11] U. Goyal, A. Krishna, A. Kumar, and K. Sharma, ”Detection and Prevention of Cyber Attacks on Multi-purpose IoT Devices Using Honeypot,” SSRN Electronic Journal, Jan. 2022, doi: 10.2139/ssrn.4157345.
[12] A. Kumar, S. Malhotra, I. Dias, and K. Lee, “Active Blended Learning Model for Teaching IoT Application Development Course: A Case Study,” Journal of Engineering Education Transformations, vol. 37, IS2, pp. 425–431, Jan. 2024, doi: 10.16920/jeet/2024/v37is2/24070.
[13] C. Liu, W. He, R. Deng, Y. C. Tian, and W. Du, ”False data injection enabled network parameter modifications in power systems: Attack and detection,” IEEE Trans. Ind. Inform., 2022.
[14] “Dataset “Matter-IoT-Anomaly-Detection,” GitHub. [Online]. Available: https://github.com/Jahnvii/Matter-IoT-Anomaly-Detection. [Accessed: Jan. 5, 2025].
[15] A. Gholamy, V. Kreinovich, and O. Kosheleva, ”Why 70/30 or 80/20 Relation Between Training and Testing Sets: A Pedagogical Explanation,” in *Proceedings of the 2018 Conference*, 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:7467506.
[16] I. Khan, “Internet of Things (IoT) Its Applications and Enabling Technologies in Healthcare: A Review,” Educational Administration Theory and Practice Journal, vol. 30, no. 4, Apr. 2024, doi: 10.53555/kuey.v30i4.5031.
[17] J. Zhang, I. Vukotic, and R. Gardner, “Anomaly detection in wide area network mesh using two machine learning anomaly detection algorithms,” Jan. 2018, doi: 10.48550/arXiv.1801.10094.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:04:49 UTC from IEEE Xplore. Restrictions apply.

ECS Journal of Solid State Science and Technology

A Novel n+ Pocket Doped L-Shaped Channel TFET for Biosensing Applications

To cite this article: Danishk Barwa *et al* 2025 *ECS J. Solid State Sci. Technol.* **14** 047010

View the article online for updates and enhancements.

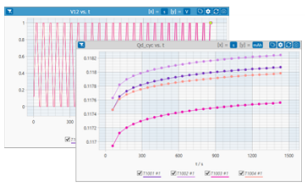
You may also like

- Dielectric Modulated InAs Pocket Heterojunction Tunnel FET for Biosensor Applications
Shashi Ranjan Upadhyay, Sumit Kale and Anukul Pandey
- Photo Sensing Properties of Heterojunction WO_3/p-Si and WS_2/p-Si Thin Films Synthesized by CBD
Avneesh Kumar, Ajeet Gupta, Arun Kumar et al.
- Effective Improvement of Chemical Mechanical Polishing for Silicon Wafers Using Nitrogen-Containing Heterocyclic Compounds Slurry
Haitao Wu, Zefang Zhang, Jiuwei Bai et al.

PAT-Tester-x-8 Potentiostat:
Modular Solution for Electrochemical Testing!

EL-CELL®
electrochemical test equipment





- **Flexible Setup with up to 8 Independent Test Channels!**
Each with a fully equipped Potentiostat, Galvanostat and EIS!
- **Perfect Choice for Small-Scale and Special Purpose Testing!**
Suited for all 3-electrode, optical, dilatometry or force test cells from EL-CELL.
- **Complete Solution with Extensive Software!**
Plan, conduct and analyze experiments with EL-Software.
- **Small Footprint, Easy to Setup and Operate!**
Usable inside a glove box. Full multi-user, multi-device control via LAN.



Scan me!

Contact us:
+49 40 79012-734
sales@el-cell.com
www.el-cell.com

This content was downloaded from IP address 14.139.251.98 on 01/05/2025 at 07:51

ECS Journal of Solid State Science and Technology, 2025 **14** 047010
2162-8777/2025/14(4)/047010/11/$40.00 © 2025 The Electrochemical Society ("ECS"). Published on behalf of ECS by IOP Publishing Limited.
All rights, including for text and data mining, AI training, and similar technologies, are reserved.

A Novel n+ Pocket Doped L-Shaped Channel TFET for Biosensing Applications

Danishk Barwa, Dhruv Khanna, Harshit Parihar, and Sumit Kale[z]

Department of ECE, Delhi Technological University, Delhi, 110042, India

This paper presents a novel channel-engineered Tunnel Field-Effect Transistor (TFET) featuring n+ pocket doping, specifically tailored for biosensing applications. The proposed TFET incorporates an L-shaped silicon (Si) channel with a distinctive configuration—an upright drain at the top and a lateral source at the bottom. Additionally, pocket doping is introduced at the source-channel interface, and a nano gate cavity under the dual gate for biomolecule immobilisation. The combination of pocket doping and the L-shaped channel significantly improves the performance of the TFET for biosensing. The device's performance in presence of different neutral as well as charged biomolecules has been analysed using energy band variation, surface potential, internal electric field, and transfer characteristics. The device's sensitivity and selectivity have been investigated in terms of I_{ON}, I_{ON}/I_{OFF}, threshold voltage (V_{th}), and subthreshold swing (SS) and the proposed device's performance has been compared with state-of-the-art biosensors.
© 2025 The Electrochemical Society ("ECS"). Published on behalf of ECS by IOP Publishing Limited. All rights, including for text and data mining, AI training, and similar technologies, are reserved. [DOI: 10.1149/2162-8777/adccf8]

Manuscript submitted February 18, 2025; revised manuscript received April 10, 2025. Published April 23, 2025.

Biosensors are miniature analytical tools that translate biological interactions into detectable electrical signals. These instruments must demonstrate high sensitivity, selectivity, rapid response, affordability, and ease of use. FET-based biosensors have earned widespread attention due to their numerous advantages, including high sensitivity, specificity, compact size, portability, fast operation, cost-effectiveness, ease of use, and suitability for large-scale analysis. The first FET-based biosensor, the Ion-Sensitive FET, was widely used for detecting charged biomolecules, but due to its inability to detect neutral biomolecules, its effectiveness is limited.[1–12]

Despite their advantages, FET-based biosensors face several challenges, such as the requirement of a minimum subthreshold slope (SS) of 60 mV decade^{-1}, vulnerability to SCEs, and a reduction in drive current due to random dopant deviations. To overcome these issues, various advanced device architectures, including IMOS,[13,14] JFET,[15,16] and Schottky Barrier MOSFET (SBMOS),[17–21] have been suggested for biosensor applications. However, these designs continue to face challenges such as low ON current (I_{ON}), issues related to doping, significant subthreshold leakage current, high SS, and increased complexity in fabrication at the nanoscale. Consequently, they demonstrate lower biomolecule detection sensitivity, higher power consumption, and slower detection speeds in label-free biosensors. In recent years, TFETs have gained attention as a promising alternative for biosensing applications.[22–31]

As a result of band-to-band tunneling (BTBT) current conduction mechanism, TFETs achieve a subthreshold slope below 60 mV decade^{-1}, making them both cost-effective and reliable. However, they also present challenges, like low ON-current and substantial ambipolar leakage, stemming from the constrained BTBT rate, which affects both analog and digital performance. Various modifications to TFET architecture and materials have been explored to mitigate these issues like Double-Gate TFETs (DG-TFETs),[32] Drain Engineered-Quasi Gate TFET(DE-QG-TFET),[33] Heterojunction TFETs (HTFETs),[34] Vertical TFETs (VTFETs),[35] and Dielectric Modulated TFETs (DM-TFETs).[9] Each of these designs has specific advantages and limitations: DG-TFETs improve electrostatic control but still suffer from low ON-current (I_{ON}), DE-QG-TFETs overcome the problem of ambipolar current but offers low ON-current (I_{ON}), HTFETs achieve higher I_{ON} using III-V materials but face fabrication challenges due to material integration issues, VTFETs increase the tunneling area, improving current drive, but require complex vertical fabrication (such as high-aspect-ratio etching, precise gate alignment, and advanced epitaxial growth techniques), DM-TFETs are highly sensitive to biomolecules but suffer from moderate I_{ON} and SS limitations. Therefore, a new device design is needed to effectively overcome these shortcomings while being a promising candidate for biosensing applications.

In this study, we propose a **novel L-shaped Channel TFET with n+ Pocket Doping** designed specifically for biosensing applications. The L-shaped channel introduces a localized electric field enhancement at the source-channel junction that facilitates band-to-band tunnelling and reduces unwanted leakage currents due to SCEs (Short Channel Effects), resulting in higher drive currents. Additionally, the n+ pocket doping at the source-channel interface plays a crucial role in improving the tunneling probability by creating a sharper potential gradient, reducing subthreshold swing (SS), and enhancing threshold voltage (V_{th}) stability. This configuration increases device sensitivity and improves charge carrier transport efficiency, making it an ideal candidate for biosensing applications. The central purpose of this research is to assess the device's performance by varying K and N_f values. Key electrical parameters, including energy band variation, surface potential, electric field distribution, and transfer characteristics, are analysed under different K and N_f conditions. Additionally, sensitivity and selectivity analysis are conducted using varied values of K and N_f.

Device Design, Calibration and Simulation

The 3D structure of the device is shown in Fig. 1. The device includes a L-shaped intrinsic channel with a drain on the top of the channel. Secondly, it has an in-built pocket at the source/intrinsic channel interface. Table I shows the parameters and specifications used for the simulation of the device. The simulations for this study were conducted using **Silvaco TCAD** for accurately modelling the tunneling mechanisms and carrier transport in the proposed TFET structure. Several models were incorporated to enhance simulation accuracy, including **bbt.nonlocal** (for nonlocal band-to-band tunneling), **bbt.forward** (forward tunneling enhancement), **qtunn.dir = 1** (enables quantum tunneling in a specific direction), **bgn** (bandgap narrowing in heavily doped regions), **consrh** (Shockley-Read-Hall recombination with constant lifetimes), and **conmob** (modelling carrier mobility under different doping and field conditions). The material used in the simulations include Silicon as the semiconductor material with a tunneling effective mass (me.tunnel) of 0.14, SiO_2 as oxide layer and oxide material with different K values to model different biomolecule interactions. The non-local BBT model is calibrated using conventional DG-TFET of K. Boucart and A. M. Ionescu.[32] The calibration has been performed which shows close agreement between reported results and our simulation, which confirms the validity of the simulations performed in this work.

[z]E-mail: sumitkale785@gmail.com

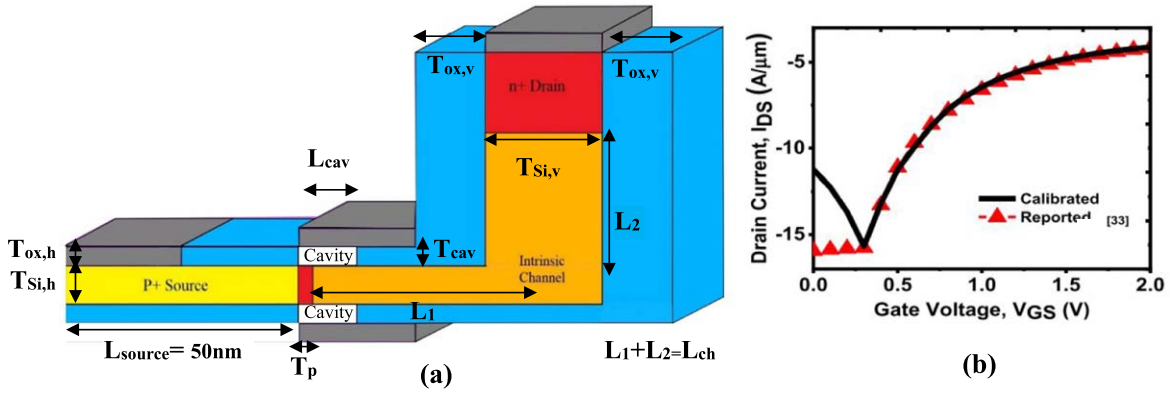


Figure 1. (a) Schematic view of the device (b) The calibrated simulation model with prefabricated TFET.

Table I. Device parameters.

Parameters	Values
Source Doping	1×10^{20} cm^{-3}
Drain Doping	5×10^{18} cm^{-3}
Channel Doping	1×10^{17} cm^{-3}
Pocket Doping	3×10^{19} cm^{-3}
Silicon Thickness, Horizontal ($T_{si,h}$)	10 nm
Silicon Thickness, Vertical ($T_{si,v}$)	25 nm
SiO_2 Thickness, Horizontal ($T_{ox,h}$)	5 nm
SiO_2 Thickness, Vertical ($T_{ox,v}$)	15 nm
Pocket Thickness (T_p)	3 nm
Channel Length (L_{ch})	90 nm
Nanogap Cavity Length (L_{cavity})	12.5 nm
Nanogap Cavity Thickness (T_{cavity})	5 nm
Gate Metal Work Function	4.5 eV

Potential fabrication flow for the proposed device.—As illustrated in Figs. 2a–2h, the proposed device can be fabricated using a CMOS-compatible process. In Fig 2a, a thin silicon film with upper and lower oxide layers can be achieved through methods detailed in Ref. 36. Figure 2b shows the formation of the p+ source via highly precise ion implantation. A SiO_2 layer can then be developed on this structure, as illustrated in Fig. 2c. Additionally, selective etching is performed to create a vertical cavity on the right side of the device, as depicted in Fig. 2d. This process defines a seed window for the selective epitaxial growth of intrinsic silicon,[37] forming the vertical drain. Any excess silicon that extends beyond the oxide can be removed using Chemical Mechanical Polishing, as shown in Fig. 2e. The n+ drain is implemented using energized ion implantation, and pocket doping is carried out, as depicted in Fig. 2f. Furthermore, Fig. 2g illustrates the selective etching of oxides to form vertical and horizontal gate dielectrics. Finally, metallization is performed to establish the source, drain, and gate contacts, as seen in Fig. 2h,[36] followed by another selective etching step to create the cavity shown in Fig. 2i.

The biomolecules are modeled by varying the dielectric constant (K), associated to different biomolecules. When biomolecules are present in the nanogap cavity, K is greater than 1, whereas K = 1 indicates their absence. In this study, biomolecules are modeled within a nanogap cavity measuring 12.5 nm × 5 nm. The proposed device considers biomolecules with varying K values, as outlined in Table II.

Results and Discussion

This section discusses the response of the proposed biosensor device when exposed to both neutral and charged biomolecules. The

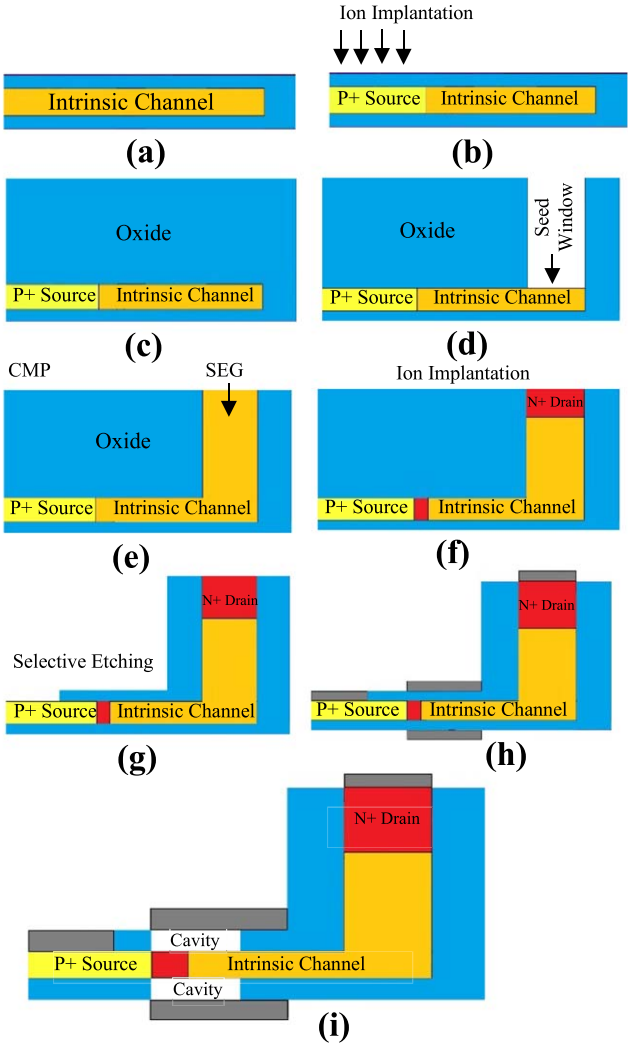


Figure 2. Key steps in the fabrication of proposed TFET.

device's performance has been evaluated by varying dielectric constant values and charge densities. For consistency, a reference value of K = 1 has been maintained throughout the analysis. Figure 3 illustrates the energy band diagram corresponding to neutral biomolecules, for different values of K. An increase in the K value from 2.63 to 12 results in a narrower tunneling width, facilitating the movement of charge carriers from the source to the drain. Consequently, this enhances carrier transport efficiency, leading to

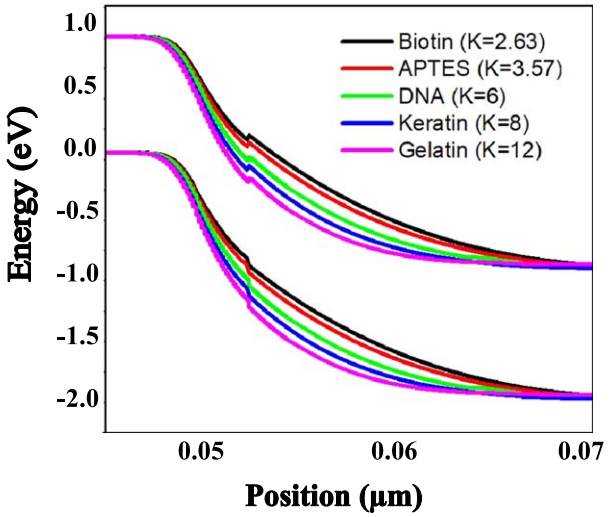


Figure 3. Energy bands of (this device) for $V_{gs} = 1.5$ V, $V_{ds} = 1$ V, $N_f = 0$ at different dielectric constant.

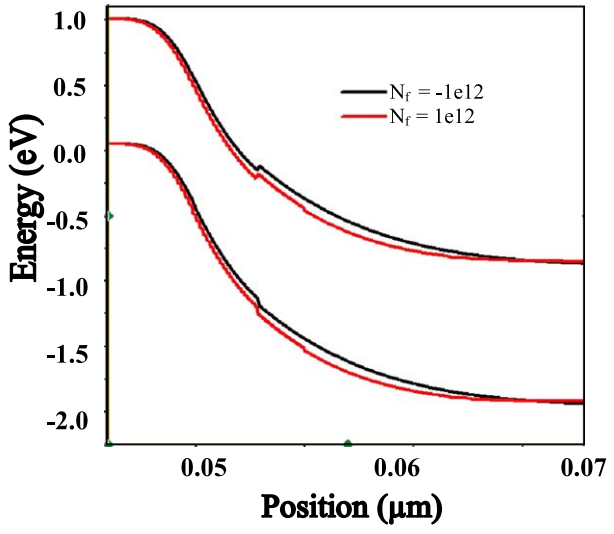


Figure 4. Energy band graphs of the (this device) for $K = 2.63$, $V_{gs} = 1$ V, $V_{ds} = 1$ V at different charge densities (N_f).

Table II. Biomolecules Dielectric Constant.[38]

Biomolecule	Dielectric constant (K)
Biotin	2.63
APTES	3.57
DNA	6
Keratin	8
Gelatin	12

a higher drain current, as depicted in Fig. 9. Additionally, Fig. 4 presents the energy band diagram of the proposed device under different charge density variations.

From Fig. 4, it can be observed that as the charge density becomes more negative, the tunneling width increases. Positively charged biomolecules lower the energy bands, reducing the tunneling barrier and enhancing carrier injection, leading to a lower V_{th} and increased tunneling probability. In contrast, negatively charged biomolecules raise the energy bands, increasing the tunneling barrier

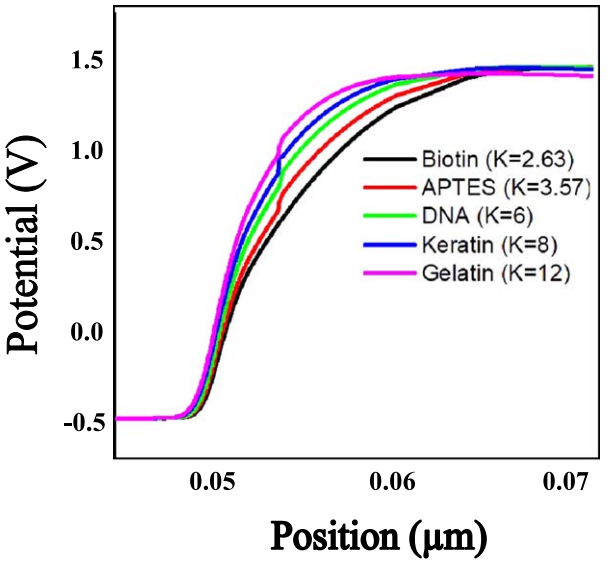


Figure 5. Surface potential of the (this device) for $V_{gs} = 1.5$ V, $V_{ds} = 1$ V, $N_f = 0$ at different dielectric constant.

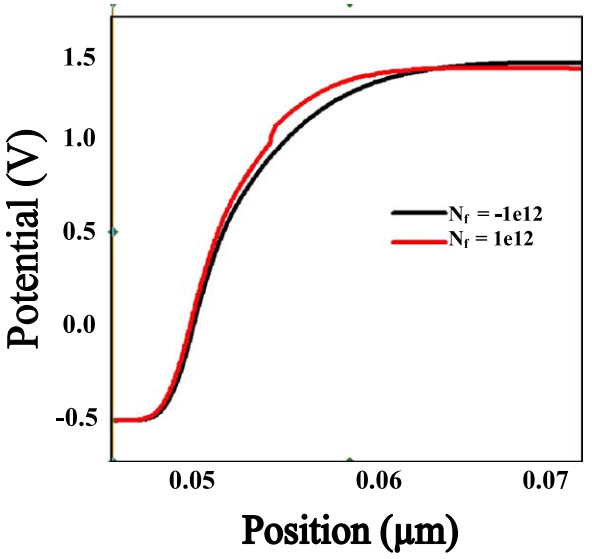


Figure 6. Surface potential variation of the (this device) for $V_{gs} = 1$ V, $V_{ds} = 1$ V, $K = 2.63$ at different charge densities (N_f).

and shifting Vth to a higher value, thereby reducing carrier injection. These charge-induced modulations alter the surface potential, redistributing the electric field in the gate cavity region and directly influencing tunneling efficiency and switching behavior. Figure 5 illustrates how the surface potential varies in response to neutral biomolecules. Notably, the surface potential exhibits a direct correlation with the dielectric constant (K). As illustrated in Fig. 5, the surface potential increases as the dielectric constant(K) rises.

In Fig. 6, the presence of negatively charged biomolecules lead to a decrease in surface potential as the charge density magnitude increases. Figures 7 and 8 depict variations in the electric field caused by changes in K and N_f. These variations in surface potential directly impact the electric field distribution. The electric field strengthens with an increase in the K of neutral biomolecules, whereas it weakens as the magnitude of negative charge density rises. Figures 9 and 10 present the transfer characteristics of

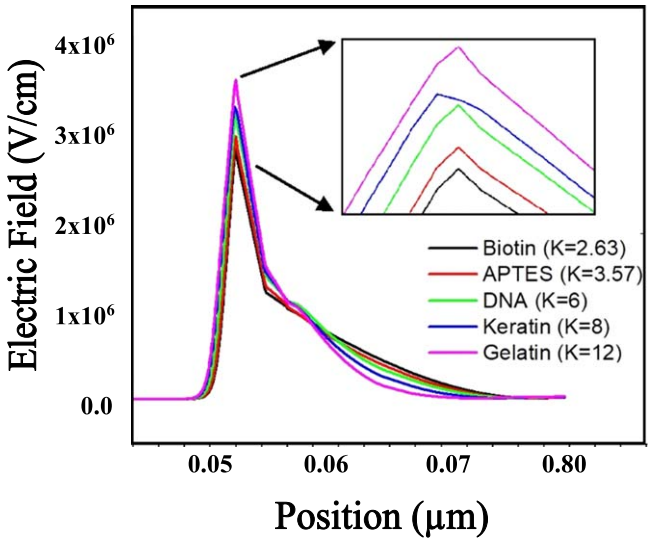


Figure 7. Electric field variation for $V_{gs} = 1.5$ V, $V_{ds} = 1$ V, $N_f = 0$ at different dielectric constant.

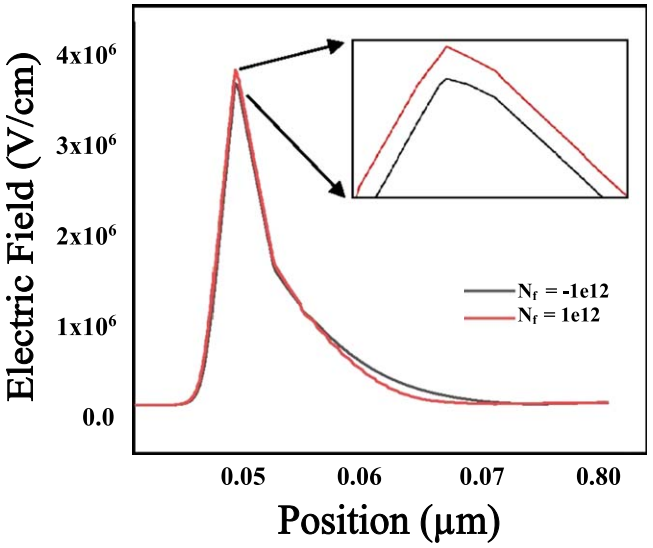


Figure 8. Electric field variation of the (this device) for $K = 2.63$, $V_{gs} = 1$ V, $V_{ds} = 1$ V at different charge densities (N_f).

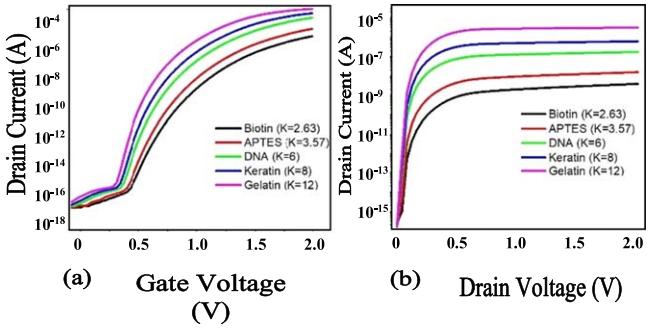


Figure 9. Transfer characteristics of (this device) at $N_f = 0$ for different K values (a) I_{ds} vs V_{gs} curve (b) I_{ds} vs V_{ds} curve.

the proposed device, analysed for different dielectric constants ($K = 2.63$, 5, 8, and 12) and varying charge densities ($N_f = -1 \times 10^{12}$ C cm^{-2} and 1×10^{12} C cm^{-2}), respectively.

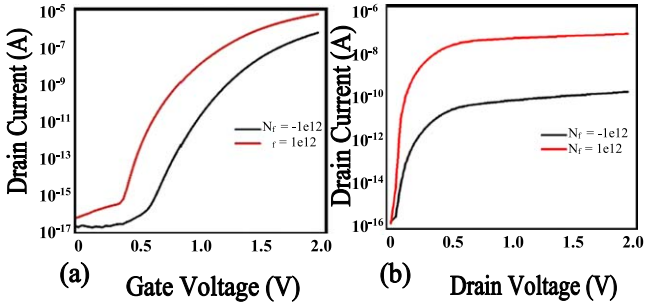


Figure 10. Transfer characteristics of (this device) at $K = 2.63$ for different charge densities (a) I_{ds} vs V_{gs} curve (b) I_{ds} vs V_{ds} curve.

Sensitivity Parameters

Sensitivity is a crucial parameter for evaluating a sensor's response. A highly sensitive and fast device is essential for providing rapid analysis results. To determine the sensitivity parameter, a reference material is required. In this study, air ($K = 1$) is used as the benchmark for sensitivity calculations. The impact of various biomolecules with $K = 2.63$, 3.57, 6, 8, and 12 on the sensitivity of the proposed device structure has been analysed. This analysis involves evaluating different sensitivity parameters, including V_{th} sensitivity, I_{ON} sensitivity, SS sensitivity, and I_{ON}/I_{OFF} ratio sensitivity. These sensitivity parameters are determined using Eq. 1.

$$X_{sensitivity} = \frac{X_{air} - X_{bio}}{X_{air}} \quad [1]$$

X = Sensitivity parameter
X_{air} = Value in the absence of biomolecule ($K = 1$)
X_{bio} = Value of Sensitivity parameter in the presence of biomolecules ($K = 2.63$, 3.57, 6, 8, & 12)

Similarly, sensitivity parameters for various charge densities (N_f) can be determined using $N_f = 0$ as the reference point. Increasing the value of K increases the effective capacitance, which increases the

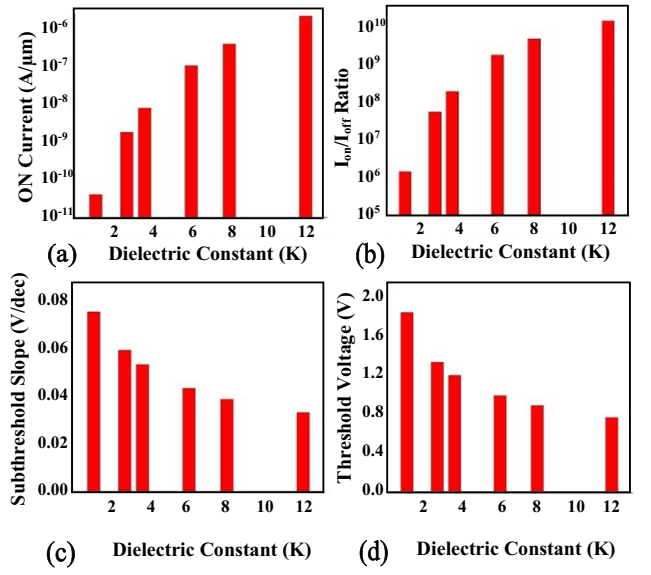


Figure 11. (a) I_{ON} current, (b) I_{ON}/I_{OFF} Ratio (c) subthreshold slope (SS) (d) threshold voltage (V_{th}) at $N_f = 0$, $V_{gs} = 1.5$ V, $V_{ds} = 1$ V for different K values.

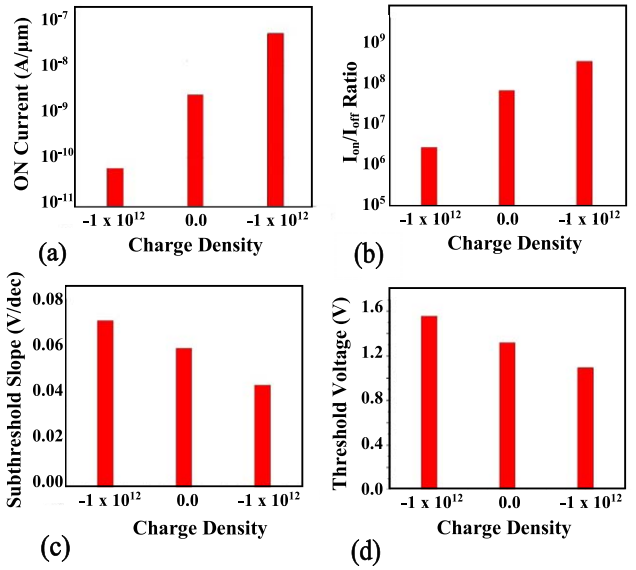


Figure 12. (a) I_{ON} current, (b) I_{ON}/I_{OFF} Ratio (c) subthreshold slope (SS) (d) threshold voltage (V_{th}) at K = 2.63, V_{gs} = 1 V, V_{ds} = 1 V for different charge densities.

Table III. Sensitivity parameters at different K, with reference K = 1 (V_{ds} = 1 V, N_f = 0, L_{ch} = 90 nm, L_{cavity} = 12.5 nm, T_{cavity} = 5 nm).

Parameters	Dielectric constant	Value
I_{on} Sensitivity	2.63	45.898
	3.57	215.96
	6	2.98×10^3
	8	1.15×10^4
	12	6.47×10^4
I_{on}/I_{off} Sensitivity	2.63	38.74
	3.57	139.5
	6	1.35×10^3
	8	3.6×10^3
	12	1.09×10^4
SS Sensitivity	2.63	0.214
	3.57	0.293
	6	0.424
	8	0.486
	12	0.559
V_{th} Sensitivity	2.63	0.276
	3.57	0.349
	6	0.461
	8	0.515
	12	0.579

Electric field in the cavity region. A high electric field leads to steeper band bending and a reduction in tunnel width, which increases the tunneling rate and enhances the device's drive current. This explains why the drive current rises with an increasing K. As shown in Fig. 11, the SS decreases as K increases. This occurs because a thinner barrier enhances tunneling probability, improving switching characteristics and lowering SS. Additionally, a higher K-value strengthens the gate's control over the channel region, facilitating carrier tunneling at a lower gate voltage, which in turn reduces the threshold voltage (V_{th}). The sensitivity parameters corresponding to different K can be seen in Table III.

Similarly, Fig. 12 demonstrates that a positive charge density intensifies the electric field and reduces the tunneling width, leading to a significant increase in ON current. Conversely, a negative

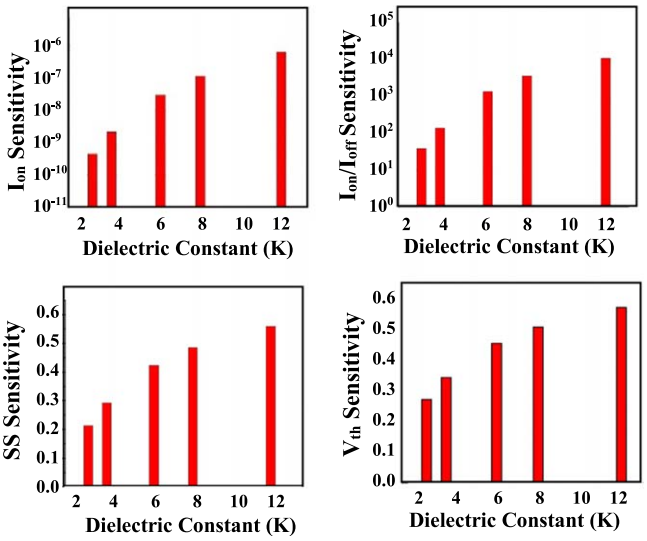


Figure 13. Sensitivity of (this device) for different K values at N_f = 0, V_{gs} = 1.5 V, V_{ds} = 1 V.

Table IV. Sensitivity parameters at different charge density, with reference N_f = 0 (V_{ds} = 1 V, K = 2.63, L_{ch} = 90 nm, L_{cavity} = 12.5 nm, T_{cavity} = 5 nm).

Parameters	Charge density($C \cdot cm^{-2}$)	Value
I_{on} Sensitivity	-1×10^{12}	0.969
	1×10^{12}	17.16
I_{on}/I_{off} Sensitivity	-1×10^{12}	0.957
	1×10^{12}	3.968
SS Sensitivity	-1×10^{12}	0.198
	1×10^{12}	0.260
V_{th} Sensitivity	-1×10^{12}	0.179
	1×10^{12}	0.172

Table V. Sensitivity Parameters at Different Channel Length, with reference L_{ch} = 90 nm (V_{ds} = 1 V, K = 5, N_f = 0, L_{cavity} = 12.5 nm, T_{cavity} = 5 nm).

Parameters	Channel length(nm)	Value
I_{on} Sensitivity	75	0.264
	105	0.219
	120	0.199
I_{on}/I_{off} Sensitivity	75	0.185
	105	0.207
	120	0.284
SS Sensitivity	75	0.151
	105	0.0675
	120	0.107
V_{th} Sensitivity	75	0.0159
	105	0.0172
	120	0.0144

charge density expands the tunneling width, thereby reducing the ON current. The SS associated with a positive charge density is lower than that with a negative charge density due to the reduced tunneling width as charge density increases. Figures 13 and 14 depict the variations in sensitivity parameters for different values of K and charge densities, respectively. In Fig. 13, the parameters are observed to increase with increasing K values. Likewise, Figure 14

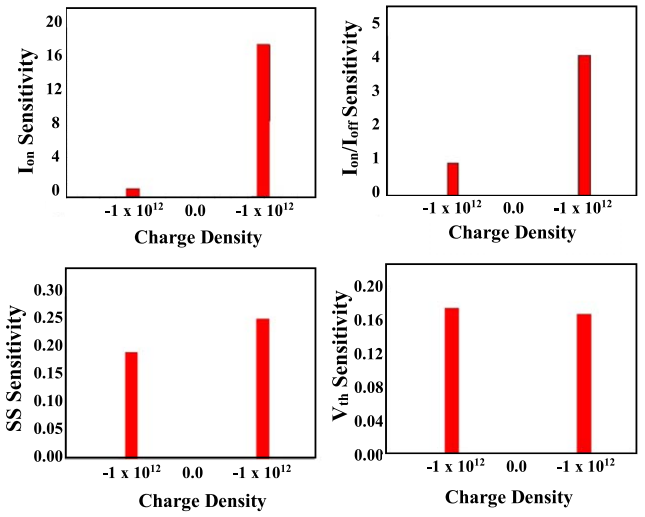


Figure 14. Sensitivity of (this device) for different N_f values K = 2.63, V_{gs} = 1.5 V, V_{ds} = 1 V.

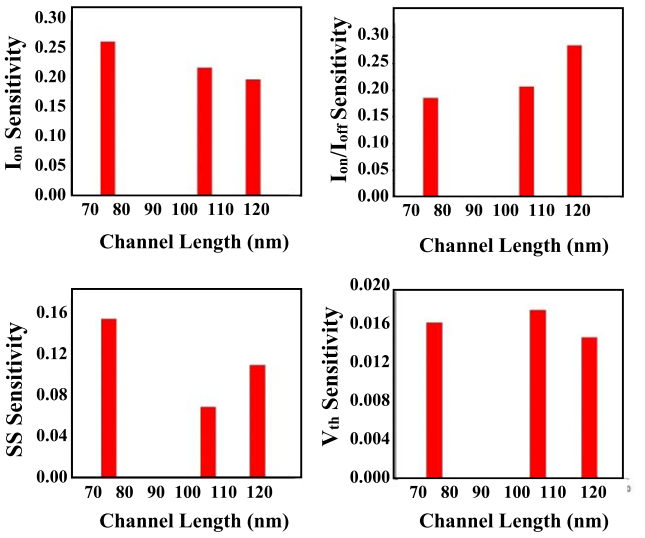


Figure 16. Sensitivity of (this device) for different channel length at K = 5, V_{gs} = 1 V, V_{ds} = 1 V.

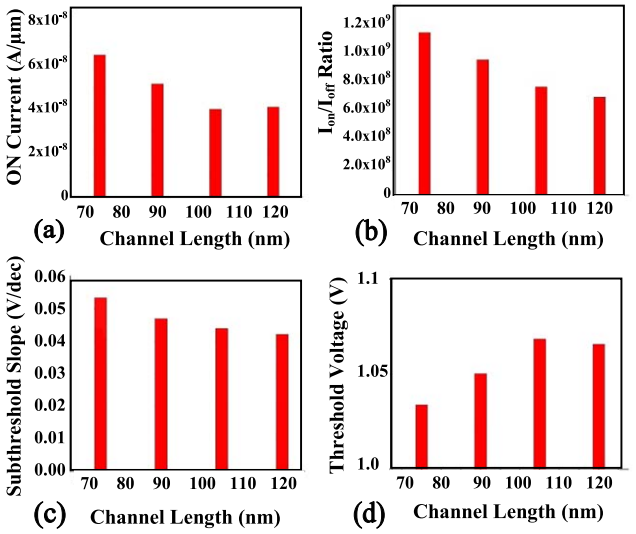


Figure 15. (a) I_{ON} current, (b) I_{ON}/I_{OFF} ratio(c) subthreshold Slope (SS) (d) threshold voltage (V_{th}) at K = 5, V_{gs} = 1 V, V_{ds} = 1 V for different channel length.

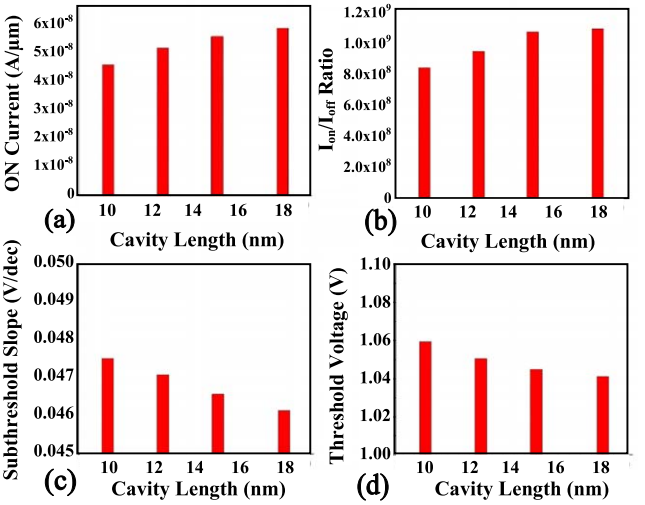


Figure 17. (a) I_{ON} current, (b) I_{ON}/I_{OFF} ratio (c) subthreshold Slope (SS) (d) threshold voltage (V_{th}) for different cavity length for neutral biomolecules at V_{gs} = 1 V, V_{ds} = 1 V.

shows a general rise with increasing charge density. Table IV summarizes the sensitivity parameters for various charge densities. Figures 15 and 16 illustrate the impact of channel length (L_{ch}) variation on DC parameters and sensitivity parameters, respectively. In this study, neutral biomolecules have been analysed at different channel lengths (L_{ch}). Table V presents the sensitivity parameters for the proposed device biosensors across varying channel lengths.

The influence of cavity length (L_{cavity}) variation on DC and sensitivity parameters has been analysed in Figs. 17–22, with the cavity thickness (T_{cavity}) maintained at 5 nm throughout the simulation. Figures 17 and 18 examine the impact of different cavity lengths on neutral biomolecules, revealing that sensitivity parameters improve as L_{cavity} increases.

Figures 19 and 20 illustrate the relationship between DC and sensitivity parameters and change in cavity length for biomolecules with negative charge density. Compared to neutral biomolecules, sensitivity parameters exhibit a decline. This reduction occurs because negatively charged biomolecules lower the surface potential, leading to a decrease in drain current.

Figures 21 and 22 depict the DC and sensitivity parameter trends for biomolecules with positive charge density as a function of cavity length. When positively charged biomolecules are placed within the cavity, an increase in cavity length enhances the surface potential in the cavity region, which further raises the drain current. Consequently, sensitivity parameters also improve. The sensitivity parameters for the proposed biosensor at various cavity lengths are summarized in Table VI.

The impact of varying cavity thickness, on sensitivity parameters for both charged and neutral biomolecules has been analysed in Figs. 23–28. In this study, the cavity length of 12.5 nm is taken as reference, with a K = 5. Figures 23 and 24 illustrate the relationship between cavity thickness and DC/sensitivity parameters for neutral biomolecules. Similarly, Figs. 25 and 26 present the variation in DC and sensitivity parameters for negatively charged biomolecules at different cavity

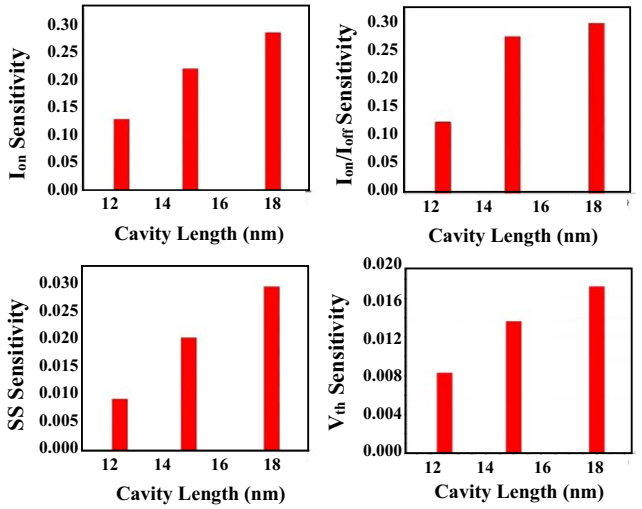


Figure 18. Sensitivity of (this device) for different cavity length for neutral biomolecules at $V_{gs} = 1$ V, $V_{ds} = 1$ V.

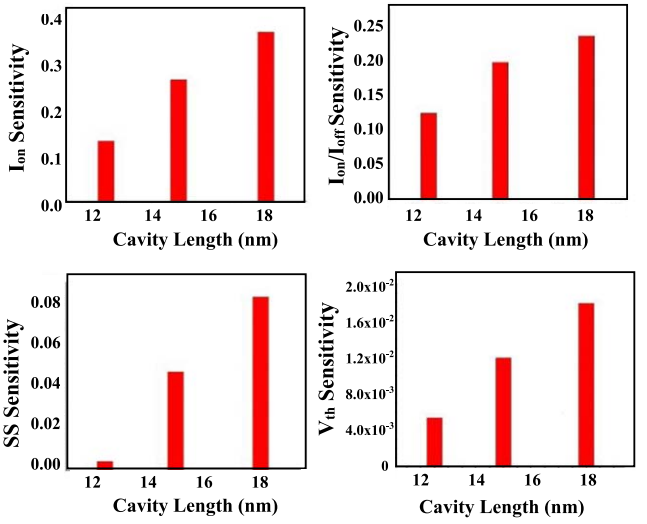


Figure 20. Sensitivity of (this device) for different cavity length for negative biomolecules at $V_{gs} = 1$ V, $V_{ds} = 1$ V.

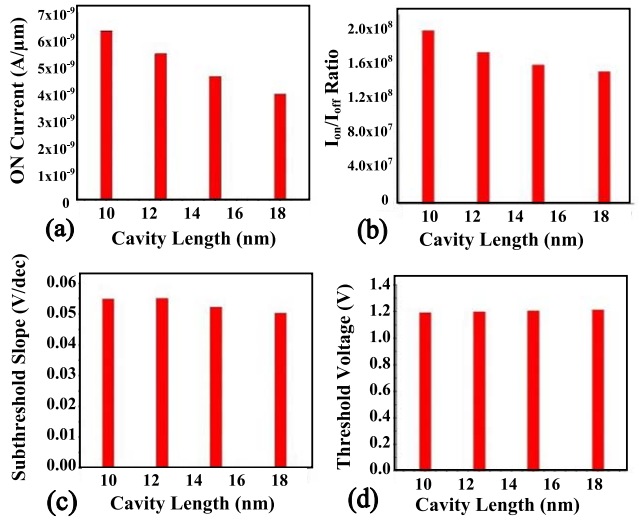


Figure 19. (a) I_{ON} Current, (b) I_{ON}/I_{OFF} Ratio (c) subthreshold Slope (SS) (d) threshold voltage (V_{th}) for different cavity length for negative biomolecules at $V_{gs} = 1$ V, $V_{ds} = 1$ V.

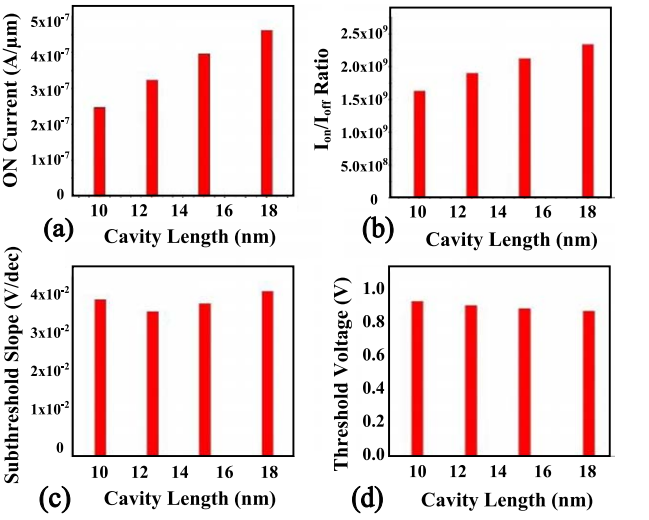


Figure 21. (a) I_{ON} current, (b) I_{ON}/I_{OFF} Ratio (c) subthreshold slope (SS) (d) threshold voltage (V_{th}) for different cavity length for positive biomolecules at $V_{gs} = 1$ V, $V_{ds} = 1$ V.

thickness values, showing that sensitivity parameters increase as cavity thickness (T_{ch}) increases. Figures 27 and 28 display the trends of DC and sensitivity parameters for positively charged biomolecules at varying cavity thicknesses. Table VII provides the sensitivity parameters of the proposed TFET biosensor at different cavity thicknesses.

The results point out that the device has High SS and V_{th} Sensitivity compared to the other existing TFET designs. The high SS and V_{th} sensitivity in our device is due to the localized electric field enhancement caused by the L-shaped channel and n+ pocket doping, which makes the device highly responsive to charge variations from biomolecules, leading to sharp changes in SS and Vth. However, the low I_{ON} sensitivity occurs because the device maintains a stable tunneling current, meaning biomolecular adsorption primarily affects the tunneling barrier without significantly altering the overall current flow. This enhances device stability, ensuring consistent performance and reducing unwanted fluctuations. This trade-off makes the sensor highly reliable for biosensing applications. Table VIII compares its sensitivity performance with earlier reported TFET-based biosensors.

Selectivity Parameters

Selectivity analysis is a technique for evaluating a biosensor's ability to accurately detect a target biomolecule while distinguishing it from other biomolecules present in the environment. The selectivity for I_{ON} and V_{th} is shown in Eqs. 2 and 3.

$$\text{Selectivity, } \Delta S, I_{on} = \frac{I_{on}(\text{target}) - I_{on}(\text{undesirable})}{I_{on}(\text{undesirable})} \quad [2]$$

$$\text{Selectivity, } \Delta S, V_{th} = \frac{V_{th}(\text{target}) - V_{th}(\text{undesirable})}{V_{th}(\text{undesirable})} \quad [3]$$

The target biomolecule in this study has a K value of 12, while other biomolecules have K values of 2.63, 3.57, 6, and 8. As the difference in the K value between the target and unwanted

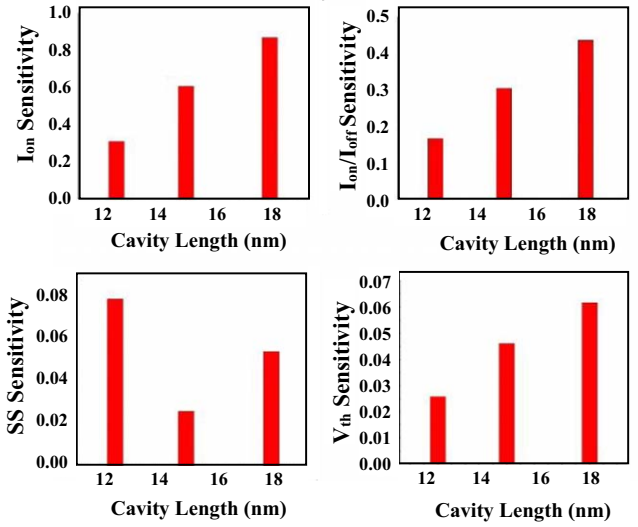


Figure 22. Sensitivity of (this device) for different cavity length for positive biomolecules at $V_{gs} = 1$ V, $V_{ds} = 1$ V.

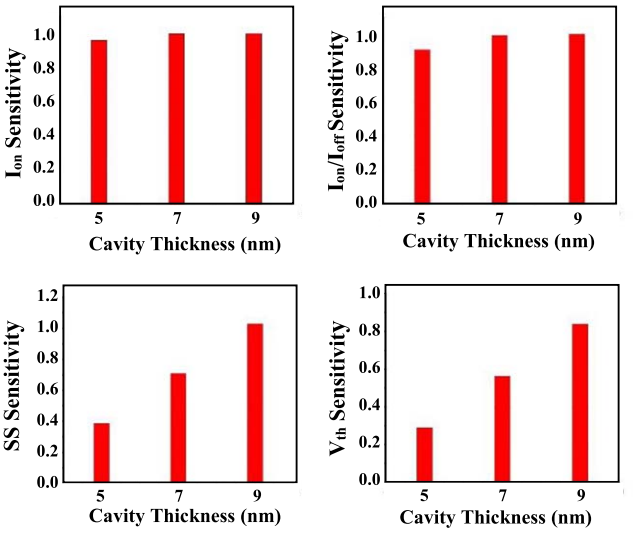


Figure 24. Sensitivity of (this device) for different cavity thickness for neutral biomolecules at $V_{gs} = 1$ V, $V_{ds} = 1$ V.

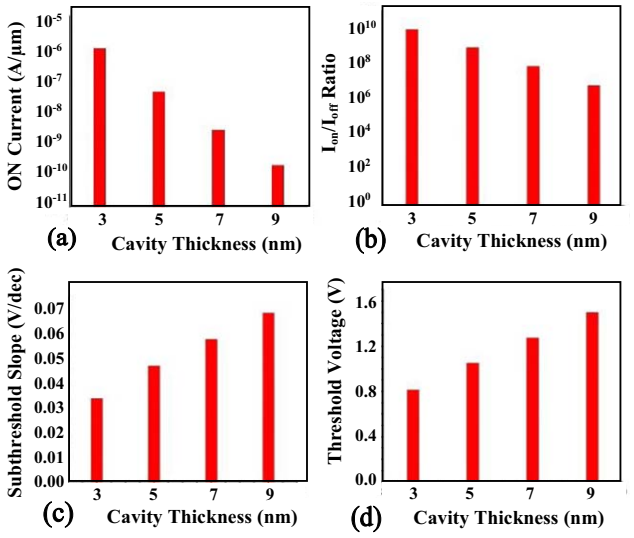


Figure 23. (a) I_{ON} current, (b) I_{ON}/I_{OFF} ratio (c) subthreshold slope (SS) (d) threshold voltage (V_{th}) for different cavity thickness for neutral biomolecules at $V_{gs} = 1$ V, $V_{ds} = 1$ V.

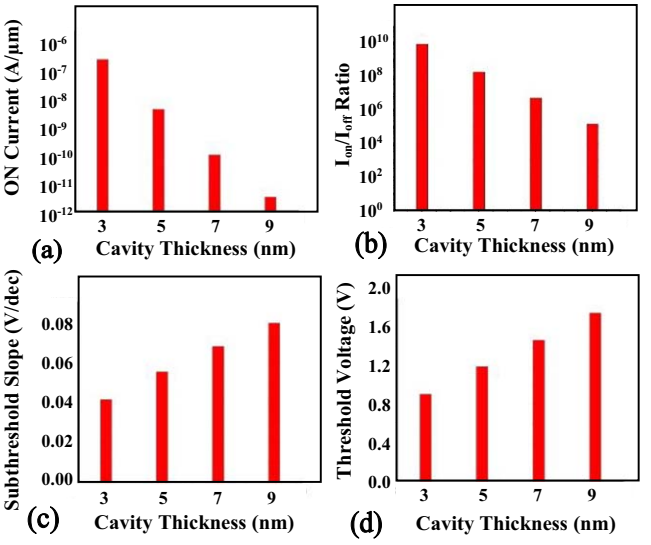


Figure 25. (a) I_{ON} current, (b) I_{ON}/I_{OFF} ratio (c) subthreshold slope (SS), (d) threshold voltage (V_{th}) for different cavity thickness for negative biomolecules at $V_{gs} = 1$ V, $V_{ds} = 1$ V.

Table VI. Sensitivity Parameters at Different Cavity Length, with reference L_{cavity} = 10 nm (V_{ds} = 1 V, K = 5, L_{ch} = 90 nm, T_{cavity} = 5 nm).

Parameters	Cavity length(nm)	Value for ($N_f = 0$ C·cm^{-2})	Value for ($N_f = -1 \times 10^{12}$ C·cm^{-2})	Value for ($N_f = 1 \times 10^{12}$ C·cm^{-2})
I_{on} Sensitivity	12.5	0.129	0.134	0.309
	15	0.218	0.269	0.606
	18	0.282	0.374	0.869
I_{on}/I_{off} Sensitivity	12.5	0.125	0.124	0.167
	15	0.275	0.198	0.305
	18	0.299	0.236	0.438
SS Sensitivity	12.5	9.17×10^{-3}	4.28×10^{-3}	0.0756
	15	0.0200	0.0472	0.0251
	18	0.029	0.0832	0.0519
V_{th} Sensitivity	12.5	8.34×10^{-3}	5.38×10^{-3}	0.0258
	15	0.0136	0.0119	0.0462
	18	0.0172	0.0179	0.0617

Table VII. Sensitivity parameters at different cavity thickness, with reference T_{cavity} = 3 nm (V_{ds} = 1 V, K = 5, L_{ch} = 90 nm, L_{cavity} = 12.5 nm).

Parameters	Cavity thickness(nm)	Value for ($N_f = 0$ C·cm^{-2})	Value for ($N_f = -1 \times 10^{12}$ C·cm^{-2})	Value for ($N_f = 1 \times 10^{12}$ C·cm^{-2})
I_{on} Sensitivity	5	0.959	0.985	0.911
	7	0.998	0.9996	0.988
	9	0.999	0.9999	0.998
I_{on}/I_{off} Sensitivity	5	0.909	0.978	0.854
	7	0.992	0.9993	0.974
	9	0.999	0.9999	0.995
SS Sensitivity	5	0.390	0.337	0.428
	7	0.709	0.644	0.845
	9	1.029	0.926	1.17
V_{th} Sensitivity	5	0.289	0.314	0.256
	7	0.563	0.616	0.499
	9	0.839	0.929	0.741

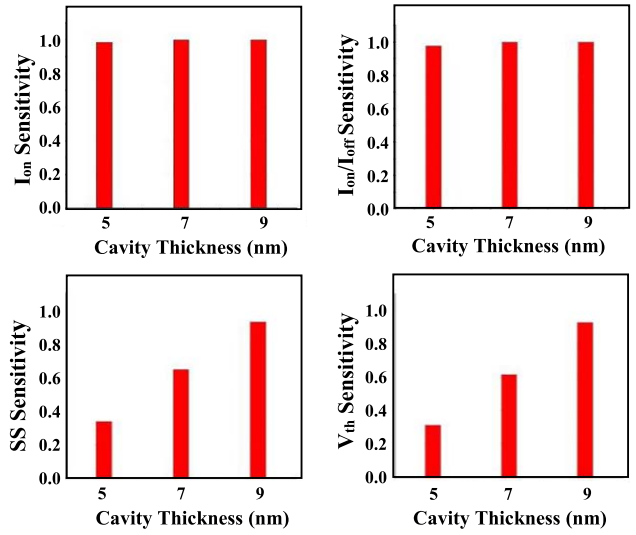


Figure 26. Sensitivity of (this device) for different cavity thickness for negative biomolecules at V_{gs} = 1 V, V_{ds} = 1 V.

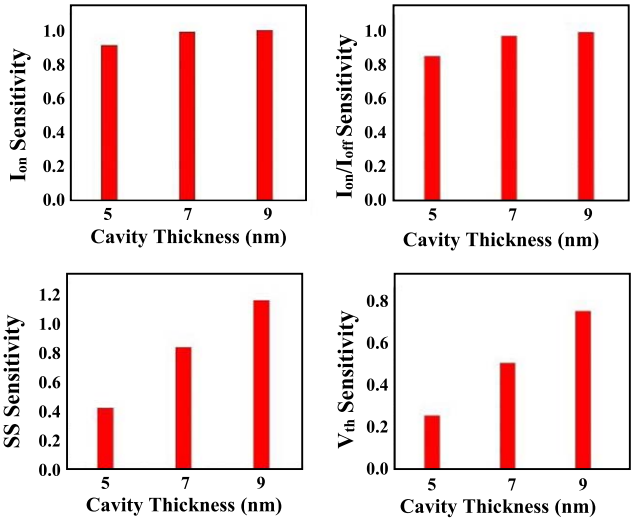


Figure 28. Sensitivity of (this device) for different cavity thickness for positive biomolecules at V_{gs} = 1 V, V_{ds} = 1 V.

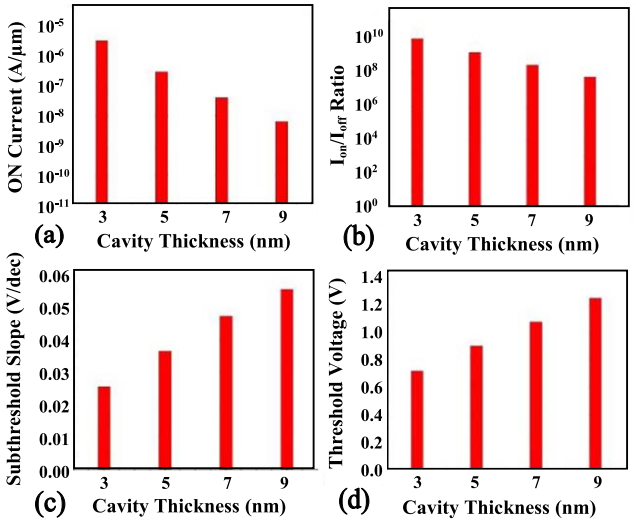


Figure 27. (a) I_{ON} current, (b) I_{ON}/I_{OFF} ratio (c) subthreshold slope (SS), (d) threshold voltage (V_{th}) for different cavity thickness for positive biomolecules at V_{gs} = 1 V, V_{ds} = 1 V.

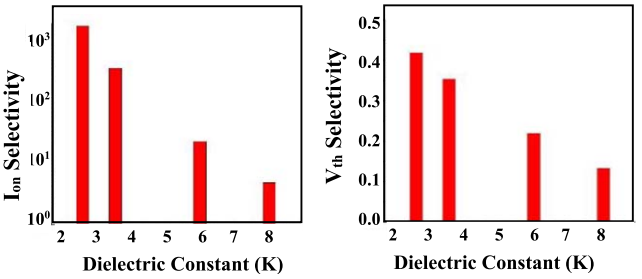


Figure 29. Selectivity of (this device) for target biomolecule K = 12, with neutral biomolecules at V_{gs} = 1 V, V_{ds} = 1 V.

biomolecules decreases, the biosensor's selectivity also declines. Figure 29 presents the relationship between selectivity parameters and the K for neutral biomolecules. Figure 30 illustrates the change in selectivity parameters for negatively charged biomolecules at different K values, while Fig. 31 depicts the selectivity trends for positively charged biomolecules under varying K values.

The selectivity parameters of the proposed TFET-based biosensor, considering a target biomolecule with K = 12 and different undesired biomolecules, are summarized in Table IX.

Table VIII. Sensitivity Comparison of Proposed Device with State-of-the-Art TFET Biosensors. (K = 12).

TFET based biosensors	V_{th} Sensitivity	SS sensitivity
DM-CP-TFET with gate cavity[39]	0.073	—
Full Gate DM-SB-TFET[40]	0.25	0.47
Short Gate DMDS-SB-TFET[40]	0.27	0.50
Staggered heterojunctions-based tunnel-FET[41]	0.45	0.25
DM-NT-TFET(K = 10)[42]	0.44	—
Electrically doped SiGe-heterojunction TFET[43]	0.196	—
TMG-JLTFET[44]	—	0.33
A Charge-Plasma-Based DM-JLTFET (K = 10)[45]	0.51	—
This Work	0.58	0.56

Table IX. Selectivity parameters at different dielectric constant, with target K = 12 (V_{ds} = 1 V, L_{ch} = 90 nm, L_{cavity} = 12.5 nm, T_{cavity} = 5 nm).

Parameters	Dielectric constant (K)	Value for ($N_f = 0$ C·cm^{-2})	Value for ($N_f = -1 \times 10^{12}$ C·cm^{-2})	Value for ($N_f = 1 \times 10^{12}$ C·cm^{-2})
I_{on} Selectivity	2.63	1378.755	18066.023	167.196
	3.57	297.251	2225.38	59.388
	6	20.686	57.325	8.964
	8	4.617	8.309	2.739
V_{th} Selectivity	2.63	0.419	0.460	0.366
	3.57	0.354	0.389	0.307
	6	0.219	0.244	0.187
	8	0.133	0.149	0.113

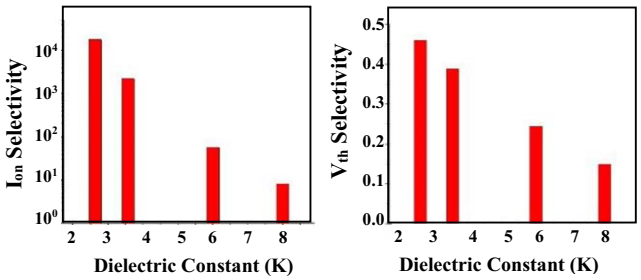


Figure 30. Selectivity of (this device) for target biomolecule K = 12, with negative biomolecules at V_{gs} = 1 V, V_{ds} = 1 V.

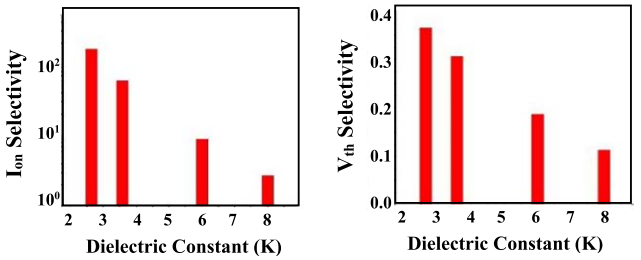


Figure 31. Selectivity of (this device) for target biomolecule K = 12, with positive biomolecules at V_{gs} = 1 V, V_{ds} = 1 V.

Conclusions

We have successfully designed and investigated the n+ pocket doped L-shaped channel TFET for biosensing applications. A comprehensive investigation and evaluation of the device's sensitivity for charged as well as neutral biomolecules modeled by different K values and charge density has been performed in terms of I_{ON}, I_{ON}/I_{OFF}, V_{th}, and SS. In addition, the device performance has been investigated for different cavity lengths and thickness on sensitivity parameters. Results confirm that the proposed device exhibits 29% better V_{th} sensitivity and 124% better SS sensitivity in comparison with state-of-the-art devices. In addition, the proposed device exhibits high I_{ON} and V_{th} selectivity. The proposed device is a promising contender for biosensing applications.

ORCID

Sumit Kale https://orcid.org/0000-0001-9761-9632

References

1. M. Barbaro, A. Bonfiglio, and L. Raffo, "A charge-modulated FET for biomolecular process detection: Concept, modeling, and simulation." *IEEE Trans. Electron Devices*, **53**, 158 (2006).
2. H. Im, X.-J. Huang, B. Gu, and Y.-K. Choi, "Dielectric modulated field-effect transistor for biosensing." *Nat. Nanotechnol.*, **2**, 430 (2007).
3. C.-H. Kim, C. Jung, H. G. Park, and Y.-K. Choi, "A novel dielectric modulated field-effect transistor for label-free DNA detection." *Biochip Journal*, **2**, 127 (2009).
4. X. Chen, Z. Guo, G. M. Yang, J. Li, M. Q. Li, J. H. Liu, and X. J. Huang, "Electrical nanogap devices for biosensing." *Mater. Today*, **13**, 28 (2010).
5. X. P. A. Gao, G. Zheng, and C. M. Lieber, "Optimal sensitivity in the subthreshold regime for nanowire FET biosensors." *Nano Lett.*, **10**, 547 (2010).
6. J. M. Choi, J.-W. Han, S.-J. Choi, and Y.-K. Choi, "Analytical modeling of a nanogap-embedded FET for biosensor applications." *IEEE Trans. Electron Devices*, **57**, 3477 (2010).
7. J. Y. Kim and J. H. Ahn, "An underlap channel-embedded field-effect transistor for biosensor applications in both wet and dry environments." *IEEE Trans. Nanotechnol.*, **11**, 390 (2012).
8. C.-T. Lee, Y.-S. Chiu, L.-R. Lou, S.-C. Ho, and C. T. Chuang, "Integrated pH sensors and performance enhancement mechanism of ZnO-based ion-sensitive field-effect transistors." *IEEE Sens. J.*, **14**, 490 (2014).
9. R. Narang, K. V. S. Reddy, M. Saxena, R. S. Gupta, and M. Gupta, "Dielectric-modulated tunnel-FET biosensor: Analytical modeling and sensitivity analysis." *IEEE Trans. Electron Devices*, **59**, 2809 (2012).
10. R. Gautam, M. Saxena, R. S. Gupta, and M. Gupta, "Numerical modeling of gate-all-around MOSFET with vacuum gate dielectric for biomolecule detection." *IEEE Electron Device Lett.*, **33**, 1756 (2012).
11. D. Sarkar and K. Banerjee, "Tunnel field-effect transistor as ultra-sensitive and label-free biosensors." *Appl. Phys. Lett.*, **100**, 143108 (2012).
12. M. J. Kumar and S. Janardhanan, "Doping-less tunnel field effect transistor: design and investigation." *IEEE Trans. Electron Devices*, **60**, 3285 (2013).
13. D. Sarkar, H. Gossner, W. Hansch, and K. Banerjee, "Impact-ionization field-effect-transistor-based biosensors for ultra-sensitive detection of biomolecules." *Appl. Phys. Lett.*, **102**, 203110 (2013).
14. N. Kannan and M. J. Kumar, "Dielectric-modulated impact-ionization MOS transistor as a label-free biosensor." *IEEE Electron Device Lett.*, **34**, 1575 (2013).

15. C.-W. Lee, I. Ferain, A. Afzalian, R. Yan, N. D. Akhavan, P. Razavi, and J.-P. Colinge, "Performance estimation of junctionless multigate transistors." *Solid-State Electron.*, **54**, 97 (2010).
16. C. Sahu and J. Singh, "Charge-plasma-based process variation immune junctionless transistor." *IEEE Electron Device Lett.*, **35**, 411 (2014).
17. S. Kale and P. N. Kondekar, "Suppression of ambipolar leakage current in Schottky barrier MOSFET using gate engineering." *Electron. Lett.*, **51**, 1536 (2015).
18. S. Kale and P. N. Kondekar, "Ambipolar leakage suppression in Ge n-channel Schottky barrier MOSFET." *IETE Journal of Research*, **61**, 323 (2015).
19. S. Kale and P. N. Kondekar, "Design and investigation of dielectric engineered dopant-segregated Schottky barrier MOSFET with NiSi source/drain." *IEEE Trans. Electron Devices*, **64**, 4400 (2017).
20. N. K. H. Latha and S. Kale, "Dielectric modulated Schottky barrier TFET for label-free biosensor applications." *Silicon*, **12**, 2673 (2020).
21. R. Singh, S. Kaim, R. MedhaShree, A. Kumar, and S. Kale, "Dielectric engineered Schottky barrier MOSFET for biosensor applications: proposal and investigation." *Silicon*, **14**, 4053 (2022).
22. S. Kanungo, S. Chattopadhyay, P. S. Gupta, and H. Rahaman, "Comparative performance analysis of the dielectrically modulated full-gate and short-gate tunnel FET-based biosensors." *IEEE Trans. Electron Devices*, **62**, 994 (2015).
23. S. Kanungo, S. Chattopadhyay, P. S. Gupta, K. Sinha, and H. Rahaman, "Study and analysis of the effects of SiGe source and pocket-doped channel on sensing performance of dielectrically modulated tunnel FET-based biosensors." *IEEE Trans. Electron Devices*, **63**, 2589 (2016).
24. S. Anand, A. Singh, S. I. Amin, and A. S. Thool, "Design and performance analysis of dielectrically modulated doping-less tunnel FET-based label-free biosensor." *IEEE Sens. J.*, **19**, 4369 (2019).
25. A. D. Singh and S. Sinha, "Design and simulation of silicon-on-insulator based dielectric-modulated field-effect transistor for biosensing applications." *Proceedings of the 3rd International Conference on IEMENTech*29 (2019).
26. R. Narang, M. Saxena, and M. Gupta, "Comparative analysis of dielectric-modulated FET and TFET-based biosensor." *IEEE Trans. Nanotechnol.*, **14**, 427 (2015).
27. D. Singh and G. C. Patil, "Dielectric-modulated bulk-planar junctionless field-effect transistor for biosensing applications." *IEEE Trans. Electron Devices*, **68**, 3545 (2021).
28. P. Raut and D. K. Panda, "Threshold voltage model development of N+ pocket vertical junctionless TFET (V-JL-TFET) as a label free biosensor." *Microelectron. J.*, **151**, 106331 (2024).
29. N. N. Reddy, P. Raut, and D. K. Panda, "Low-bandgap material engineering based TFET device for next-generation biosensor application-A comprehensive review on device structure and sensitivity." *Micro and Nanostructures*, **194**, 207935 (2024).
30. P. Raut, D. K. Panda, U. Nanda, and C. C. Hsu, "Simulation and modeling of high-sensitive JL-TFET based biosensor for label free detection of biomolecules." *Microsyst. Technol.*, **31**, 1103 (2024).
31. N. Nelaturi and D. K. Panda, "Design and investigation of dielectric modulated triple metal gate-oxide-stack Z-shaped gate horizontal pocket TFET device as a label-free biosensor." *J. Micromech. Microeng.*, **32**, 085001 (2022).
32. K. Boucart and A. M. Ionescu, "Double-Gate tunnel FET with high-κ gate dielectric." *IEEE Trans. Electron Devices*, **54**, 1725 (2007).
33. S. J. M. R. Uddin and S. Loan, "Drain-engineered TFET with fully suppressed ambipolarity for high-frequency application." *IEEE Trans. Electron Devices*, **66**, 1628 (2019).
34. G. Dewey, B. Chu-Kung, J. Boardman, J. M. Fastenau, J. Kavalieros, and R. Kotlyar, "Fabrication, characterization, and physics of III–V heterojunction tunneling field effect transistors (H-TFET) for steep sub-threshold swing." *Technical Digest - International Electron Devices Meeting* (IEDM, Washington, DC, USA) (2011).
35. E. Ko, H. Lee, J.-D. Park, and C. Shin, "Vertical tunnel FET: design optimization with triple metal-gate layers." *IEEE Trans. Electron Devices*, **63**, 5030 (2016).
36. H.-S. P. Wong, K. K. Chan, and Y. Taur, *Self-Aligned (Top and Bottom) Double-Gate MOSFET with a 25 nm Thick Silicon Channel* (IEDM Tech. Dig., Washington, DC, USA)427 (1997).
37. J. P. Denton and G. W. Neudeck, "Fully depleted dual-gated thin-film SOI P-MOSFETs fabricated in SOI islands with an isolated buried polysilicon backgate." *IEEE Electron Device Lett.*, **17**, 509 (1996).
38. S. Jain, A. Joshi, and L. R. Cenkeramaddi, "Dielectric modulated bilayer electrode top contact OTFT for label-free biosensing." *IEEE Access*, **11**, 23714 (2023).
39. A. Kumar and S. Kale, "A comparative analysis of cavity positions in charge plasma-based tunnel FET for biosensor application." *IETE Journal of Research*, **70**, 2819 (2023).
40. N. N. Poojari et al., "A comparative study of short gate and full gate dielectric modulated Schottky barrier TFET for biosensing application." *Silicon*, **14**, 12223 (2022).
41. A. Chakraborty, D. Singha, and A. Sarkar, "Staggered heterojunctions-based tunnel-FET for application as a label-free biosensor." *International Journal of Nanoparticles*, **10**, 107 (2018).
42. D. Sen, S. D. Patel, and S. Sahay, "Dielectric modulated nanotube tunnel field-effect transistor as a label-free biosensor: proposal and investigation." *IEEE Trans. Nanobiosci.*, **22**, 163 (2022).
43. B. Dewan, S. Chaudhary, and M. Yadav, "Electrically doped SiGe-heterojunction TFET-based biosensor considering non-ideal hybridization issues: a simulation study." *Appl. Phys. A*, **127**, 791 (2021).
44. S. Choudhury, K. L. Baishnab, K. Guha, Z. Jakšić, O. Jakšić, and J. Iannacci, "Modeling and Simulation of a TFET-Based Label-Free Biosensor with Enhanced Sensitivity." *Chemosensors*, **11**, 312 (2023).
45. D. Singh, S. Pandey, K. Nigam, and P. N. Kondekar, "A charge-plasma-based dielectric-modulated junctionless TFET for biosensor label-free detection." *IEEE Trans. Electron Devices*, **64**, 271 (2017).

Materials Research Proceedings 49 (2025) 386-395 https://doi.org/10.21741/9781644903438-39

Assessment of shear bond strength of triplets using different joint thicknesses and mortar types

Shilpa PAL[1,a*], Anshoo YADAV[1,b] and Khyati SAGGU[1,c]

[1]Department of Civil Engineering, Delhi Technological University, New Delhi, India

[a]shilpapal@dtu.ac.in, [b]banshoo.yadav509@gmail.com, [c]khyatisaggu_2k21phdce501@dtu.ac.in

Keywords: Masonry, Mortar Thickness, Shear Bond Strength, Triplets

Abstract. Building structures with brick masonry is the ancient form implementation and is still commonly employed worldwide, despite advances in expertise. Recent investigations are still being conducted to improve the technical and economic performance of masonry. As the mechanical characteristics of ancient masonry might vary significantly from the hypothetical one, it must be properly evaluated to quantify the well-being and usability of actual structures in the context of recuperation and/or seismic dependability valuation. Initially, small-scale assessments were performed to assess the distinct features of binder and brick separately. Investigational research of triplets made using cement and hydraulic lime-based mortars is presented in this paper. This study aimed to explore the performance variation of brick triplets utilizing two dissimilar mortar mixes cement and lime-based mortar and three mortar of 13 mm, 20 mm, and 27 mm thickness respectively. When testing was performed, the results specified that masonry units with a mortar thickness of 27 mm were more prone to shear failure and 13 mm thick specimens, regardless of the type of mortar used, were able to withstand the maximum load before failure. Cement-based mortar triplets were found to withstand 254% more load compared to lime-based mortar in the case of a 13 mm thickness.

1. Introduction

Masonry buildings continue to be a considerable part of the global constructed infrastructure. Masonry was commonly used to construct historical and aesthetic heritage including standard residential constructions in historical urban and rural areas. Masonry material cannot withstand tensile stress and is so referred to as a material that lacks tensile strength. Furthermore, because it is a diverse material, its performance is determined by the shape characteristics and the constituent materials' properties. Seismic activity can exert strong lateral forces on structures that can lead to shear deformation and failure. Therefore, the shear resistance of a structure is a critical factor in combating the effects of seismic activity. Seismic events that have hit and caused extensive damage to heavily-built areas of masonry structures in recent decades have sparked the scientific community's interest in suitable techniques to assess the seismic susceptibility of these constructions. Masonry mortar rules recommend that the mortar used in construction be frailer than the brick units to guarantee that any cracks that occur in the binder joints are readily repaired. The depth of the bed joint is another critical component influencing the overall performance of the brickwork. Several factors influence the compressive strength of mortar, including binder content, binder type, and water-cement ratio. Because of their greater binder ratio, cement-based mortars are typically regarded as stronger than lime binders. Nevertheless, compression capacity should not be the only consideration when choosing a mortar. The bond strength and flexural strength of the mortar are also important since they affect its resistance to cracking. The bond strength between the binder and brick is usually greater as compared to the tensile and compressive capacity of the binder, which means that bond failures occur at lower stress levels in mortar joints.

Despite the fact that structural engineers are creating increasingly efficient numerical techniques, "accuracy" of demonstrating findings is always reliant on the accurate documentation

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 license. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. Published under license by Materials Research Forum LLC.

of limited mechanical characteristics necessary to illustrate brickwork material. As brick, stone, and earth masonry make up a huge portion of present structures and architectural legacy in Europe and around the world, understanding their mechanical performance is critical [1]. However, assessing the mechanical properties and, as a result, the well-being and serviceableness of actual masonry is a difficult job [2], and numerous damaging and non-damaging methods are currently accessible for evaluating the integrity and excellence of resources in on-site brickwork, with the goal of providing input data for structural amplification [3]. Mortars have a considerable influence on the comprehensive performance of brickwork. The fundamental goal of binder in brickwork is to glue brickwork units together to produce an assembly that serves as an integrated component with anticipated purposeful act features [4]. An absence of connection at the boundary between mortar and brick, for example, owing to shrinkage, might allow moisture to penetrate certain regions. As a result, several methods for reinforcing unreinforced masonry buildings have been presented [5]. Liable on their form, aspect ratio, and edge conditions, masonry can suffer shear- or flexure-controlled failures. Shear failures are defined by either inclined cracks or sliding cracks, which occur as soon as a piece of the masonry sideways a bed joint moves horizontally under lateral stresses. The way brick-and-mortar interacts considerably influences the strength of URM. In general, the shear response of the interface binder-brick interface is highly reliant on moistness content, porosity, binder strength, and conditioning type [6], with saturation producing a loss of strength and stiffness [7].

When compared to other building materials, masonry exhibits anisotropic behaviour. Prior to the invention of cement, it was built with lime mortar [8]. Because of the rapid strength development of cement, the use of lime in masonry buildings quickly declined after the introduction of cement [9]. The consequence of diverse binder configurations on the structural response of brickwork is investigated in this study. To characterize masonry material, various mechanical characteristics must be correctly identified. One of these factors is the shear strength of brick triplets [10]. As a result, interfacial cohesiveness and the precise characterization of mortar qualities are critical in predicting masonry behavior under any type of activity. Cement- and lime-based mortars are explored in this study. Cement binder adds strength and longevity to brickwork, but it lacks flexibility, making it susceptible to breaking. The binder thickness of the brickwork also was treated as a changeable parameter, with considered thicknesses with the strength of the binder. In a study conducted in Korea, the behavior of factory based and handmade bricks with lime mortar is addressed and this has been determined that factory made bricks shows 15% greater compressive strength but lowers specific area by 22%. Also, handmade bricks improved shear strength. Shear stresses are mostly carried by mortar joints in masonry [11]. Similar to this work, research was conducted to analyze the physical and analytical behavior of 19th-century bricks with old aerial lime mortars after 730 days of curing. The impact of curing was checked. The results are significant in refurbishment of historic brickwork [12]. China has also been working in this domain to mechanically characterize the behavior of grey clay bricks, suggesting that though cement mortar enhances the compressive strength of the prism but it has tendency to damage grey clay bricks suggesting problem judgment as per requirement of the repair and rehabilitation [13].

This has been inferred from the literature that the standards and previous studies have been silent in understanding the impact of joint thickness of the behaviour of triplets; therefore, in present study the mortar thickness is administered. Moreover, very limited work is done considering impact of lime mortar, this study can add to the knowledge about behavior of brick masonry with lime mortar facilitating the rehabilitation and renovation work for historical structures.

2. Problem statement

The utilization of brick triplets is a widespread building tendency owing to its aesthetic appeal and structural aids. Though, the choice of mortar and thickness can pointedly govern the strength and durability of the brick triplets. In spite of the importance of these factors, there is inadequate research on the effects of diverse categories of binder and thicknesses on the features of the triplet samples. This study's aims to inspect the impact of different mortar types and thicknesses on the shear bond strength of triplets. The deductions of this work will deliver a valuable understanding of the optimal mortar and thickness combinations for brick triplets, which can enhance their performance and extend their service life.

3. Materials and Methods

This section describes the experimental evaluation, which was conducted in two phases. The first phase focused on describing the constituent ingredients, while the second phase concentrated on performing triplet tests. The constituent material samples of different types, including lime and cement were tested.Binders are constituents used in construction to keep other materials, such as aggregates, intact to make a cohesive and enduring structure. Binders show a vital role in the strength and durability of building materials since they provide the substance with its adhesive and cohesive qualities. Cement is composed of limestone, clay, and other elements that have been burned at high temperatures to make a powder that hardens when combined with water. In this experiment, 53-grade OPC as described in IS: 269-2015 was used. The properties of the binder ingredients as supplied by the supplier are highlighted in Table 1.

Lime is another frequent binder used in buildings, notably in masonry and plastering. It is created from limestone that has been burnt and slaked with water to generate a powder that hardens when exposed to air. Hydrated lime as described in IS: 1540(Part I)-1990 [14] was utilized in study. Table 1 discusses the features of the binder ingredients as given by the supplier.

Table 1. Properties of Binders

Properties	**Hydrated lime**	**OPC-53**
Chemical Composition (%)		
SiO_2	13.1	24.42
$Al2O_3$	0.7	4.85
$Fe2O_3$	2.81	3.80
CaO	66.3	66.16
MgO	0.5	1.85
K_2O	0.15	0.4
LoI	16.44	1.17
Physical Properties		
Specific Gravity	2.7	3.12
Fineness(cm^2/gm)	3800	3000
Color	White	Grey

In this study, 220 mm in length, 110 mm in width, and 75 mm deep standard-fired clay bricks were used. Before evaluating the shear bond capacity of triplets, it must be emphasized that to ensure a strong shear connection at the brick-and-mortar interface, the typology of constituent materials such as brick and mortar should be carefully examined.

Testing was done to characterize bricks by performing the following tests- Dimensions test, water absorption test, efflorescence test, and Compressive strength test in the laboratory before starting the preparation of triplets for the shear strength.

4. Tests on Specimens

4.1 Test on Bricks

(a) Dimension Tests

The experiment utilized locally available bricks for the study. However, it should be noted that not all of the bricks possess exact and precise measurements. The various brick samples are referred to as B1, B2, B3, B4, and B5. Table 2 shows the measurements of these samples.

Specimen no.	Dimensions(mm)
B_1	220x110x76
B_2	221x110x75
B_3	223x110x75
B_4	221x110x73
B_5	220x110x75

Table 2. Dimensions of Bricks

Specimens	Weight after oven-dry(M1) (Kg)	Weight after 24 hours in water(M2) (Kg)	Water absorption (%)	Avg. Water absorption (%)
B_1	2.99	3.29	10.03	11.08
B_2	3.02	3.37	11.59	
B_3	3.11	3.44	10.61	
B_4	3.13	3.46	10.54	
B_5	3.01	3.39	12.62	

Table 3. Values of water absorption

(b) Bricks Porosity Test

The experiment utilized locally available bricks for the study. However, it should be noted that not all of the bricks possess exact and precise measurements. The various brick samples are referred to as B_1, B_2, B_3, B_4, and B_5. Table 2 shows the measurements of these samples.

The standard approach defined in IS 3495:1992 Part 2 [15] was followed to measure the water absorption capacity of bricks. The samples were first dehydrated in a hot air oven at 105 to 115°C, and their dry weights were recorded (M1). Next, the bricks were deep in clean water for 24 hours, then taken out and wiped down before their wet weights were measured (M2). By subtracting the dry weight from the wet weight, the water absorption of individual bricks was found, and the average water absorption was determined. The resulting data, which provides the water absorption values for 5 samples, is listed in Table 3. To calculate the water absorption of the brick specimens, Equation (1) was utilized.

$$\text{Water Absorption (\%)} = \frac{M_2 - M_1}{M_1} \times 100 \quad (1)$$

(c) Bricks Porosity Test

The average water absorption on conducting test on the specimen was 11.08%. In accordance to the Indian standards, the average value should not surpass 20% by weight for Class B by 12.5% and 15% for higher classes. This is can be further inferred that the excessive water absorption causes structural damage like crack, warp, and corrosion of the brick. Accordingly, it is vital to monitor the properties of bricks and ensure the appropriate standards for the definite class are met.

(d) Efflorescence Test of the Bricks

The bricks were subjected to an efflorescence test, and the findings showed that the efflorescence was less than 4%. As per IS 3495:1992[16], the permissible efflorescence limit is under or equal to 10% for Class 12.5 and under or equal to 5% for higher classes.

5. Evaluating the Compression Resistance of Bricks

The compression resistance test is performed on bricks to assess its capability to endure compressive loads. The bricks are kept in between two compression plates and load is applied until the failure occurs. The results are ascertained as the maximum load that the unit can resist divided by cross-section area. The lowest strength for Class 12.5 bricks is 3.5 N/mm^2, though the minimum compressive strength for higher classes ranges from 4.5 to 7.5 N/mm^2 [17]. The importance of the strength test is to ensure that the bricks meet the essential standards and are suitable to be used in construction projects. The strength of samples was determined applying equation (2).

$$\text{Compressive Strength of Bricks (MPa)} = \frac{\mathbf{P}}{\mathbf{A}} \quad (2)$$

Where P = Peak Load at collapse (N)
A = Cross-section area (mm^2)

Figure 1 displays the strength values of individual samples. The average strength was calculated to be 22.252 MPa, which indicates that they belong to Class A as per the IS code. The test was performed on CTM after wiping the water on the surface of the bricks as the specimens were inside the water for 3 days after filling the frog on the bricks to make the surface plane so that load can be applied equally on the whole surface. Figure 2 displays the brick specimen before compressive strength test in CTM.

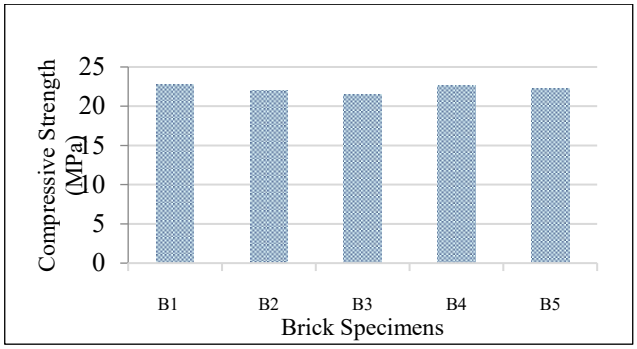


Figure 1. Crushing Strength of Brick Samples



Figure 2. Brick specimen before test

6. Test on Mortar

Mortar is a smooth paste made by mixing cement or lime, with sand. Mortars are crucial in creating a cohesive mass out of individual masonry components. For the study, lime mortars with a lime-to-sand ratio of 1:3 were chosen. This mortar was frequently used for historic masonry conservation projects because the walls' ability to naturally breathe ensures improved long-term performance. Similarly, cement is now the most efficient binder on the market, thus OPC-53 mortar was chosen because of this. The ratio of binder to aggregate in this mortar was also 1:3. River sand was used in the investigation available locally.

6.1 Mortar Compressive Strength

The mortar strength influences the maximum load limit of the brickwork, and the strength values of the mortar were computed using mortar cubes with a face area of 50 cm^2. Figure 3 depicts the test outcomes of the 28-days strength of the binder utilized in the current investigation. In comparison to lime mortar, which has a strength of 5.51MPa on average, cement mortar has a strength of 20.38MPa which is about 270% more as compared to lime mortar strength. L1, L2, L3, and C1, C2, and C3 are lime and cement mortar specimens, respectively.

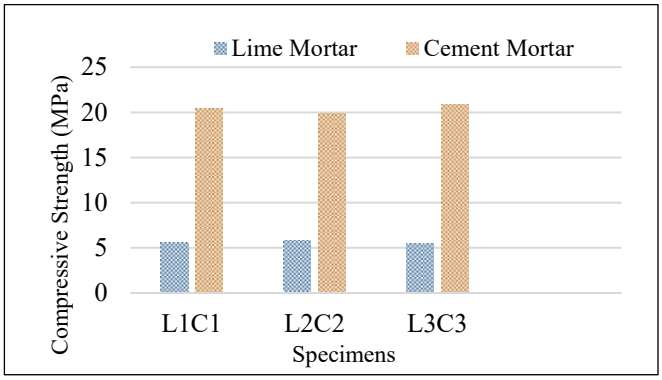


Figure 3. The individual strength of mortar cubes



Figure 4. Cement and Lime mortar triplet specimens

6.2 Triplet Specimens

This section explains the sample arrangement and testing techniques used to find the shear strength of samples. Shear capacity is one of the regulating factors in determining the capacity of brick masonry components. The shear capacity of brickwork had been calculated by experimentation by testing triplet specimens with the help of a hydraulic jack of capacity 100kN. In triplet tests, three brick assemblies are utilized to evaluate the shear bond strength. With this setup, the extreme bricks of the triplet connect against the loading edge to control movement, while the center brick is unrestrained. All of the units were submerged in water for 20 minutes before triplet casting to achieve 75% saturation. All of the samples were allowed to cure for 28 days. The failure load 'P'(N) is recorded. Equation (3) was used to compute the binding strength of a triplet.

$$\text{Shear Bond Strength} = \frac{\mathbf{P}}{\mathbf{2LB}} \quad (3)$$

Where, L= 225 mm

B = 110 mm

EN 1052-3 [18] describes the shear test, which is used to assess shear strength by evaluating specimens made up of bricks and mortar joints. Figure 4 depicts the prepared specimens with lime mortar and cement mortar respectively.

7. Results and Discussions

Two kinds of mortar were used in the current investigation, each with three distinct thicknesses, and three brick specimens were made for each thickness. Lime mortar samples were maintained in the laboratory after construction for air curing for 28 days, whereas cement mortar samples were cured in damp jute bags for 28 days.

7.1 Testing of Triplets

The masonry triplet specimens were tested using a hydraulic jack that had a maximum load capacity of 100 kilonewtons (kN).

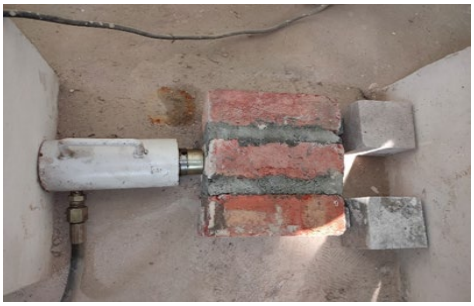


Figure 5(a). Triplet testing demonstration before testing



Figure 5(b). Triplet testing demonstration after testing

A visual representation of the testing process is presented here. A similar mode of failure was obtained in both mortars at the thicknesses of 20 mm and 27 mm as illustrated in Figure 5(b). In this kind of failure, mortar and bricks got separated from one face of bricks and at the place of the frog also mortar came out easily after applying the load. In the case of 13 mm thickness the mortar was not completely separated from the bricks in that case some mortar was on one face of the brick and some mortar was on another face of the brick and the surface of the mortar was running fracture not smooth as in previous cases. The study revealed that decreasing the strength of the mortar concluded in a decrease in the shear strength of bond values. Moreover, when comparing the different mortar thicknesses, it was observed that the 13 mm thickness showed a higher order of shear bond strength compared to the 20 mm and 27 mm thicknesses. Interestingly, the frequency of decrease of bond shear strength was found to be lower when increasing the depth of the mortar, as associated with the poor-quality binder used in brickwork structures. The results of the study indicate that the shear bond strength of brickwork is influenced to a greater extent by the strength of the mortar. While the thickness of the mortar has a relatively insignificant impact on the shear strength. Altering the thickness of the mortar layer amid the specimens can affect the stress distribution and crack propagation during testing, which can in turn influence the measured mechanical properties.

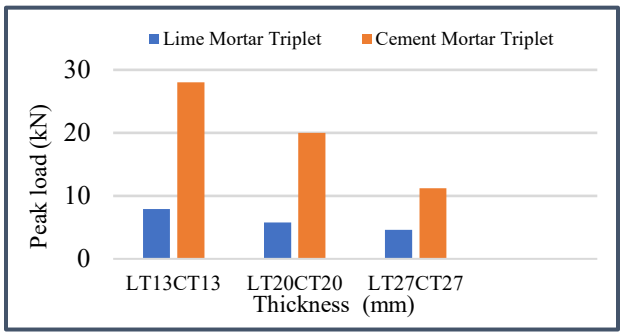


Figure 6 Relation joint thickness v/s load

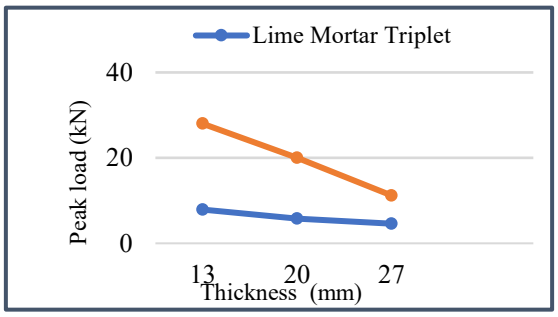


Figure 7 Relation mortar type v/s load

7.2 Mode of failure

This test can be performed using one of two distinct protocols, with or without lateral pre-compression. Figure 8 (a) and Figure 8 (b) depict the expected failure possibilities after performing this test. The first two modes (A/1 and A/2) fail owing to mortar separation from the brick, which is ascribed to poor contact between the two components (Bricks and mortar). Failure occurs in the mortar layer (B mode) when a less compressive strength mortar is used to build joints between

mortar and brick. A high-strength mortar with strong adhesive characteristics, on the other hand, would lead the specimen to fail due to the occurrence of fractures running through one or both bricks (modes C and D) [19]. In this study, mainly A/1 and A/2 were the failure mode obtained after testing in both the mortar as the compressive strength of mortar in both cases was less as compared to the strength of bricks.

7.3 Shear Bond Strength

Equation 3 demonstrates the calculation of the investigational shear bond strength, which involves dividing the maximal force that a triplet can withstand by twice the area of its cross-section [20]. Triplets with joint thicknesses of 13 mm, constructed with cement mortar or lime mortar, were found to have the highest load capacity. The 13 mm cement mortar triplet had the maximum load capacity, measuring 28.0 kN, while the 27 mm lime mortar prism had the lowest capacity, measuring 4.60 kN. Here LT13CT13, LT20CT20, and LT27CT27 show the different Lime mortar thickness and Cement mortar thickness. Table 4 display the mean peak load and shear bond strength values for cement mortar and lime mortar, respectively. For all the cases the size of specimens was the same as triplets were constructed with the help of 3 bricks.

Table 4. Shear strength of cement brick masonry and lime brick masonry

Thickness (mm)	**Specimen Size (mm)**	**Avg. Peak load (P) (kN)**	**Shear Bond Strength (MPa)**	**Avg. Peak load (P) (kN)**	**Shear Bond Strength (MPa)**
		Cement Brick Masonry		Lime Brick Masonry	
CT13±3	225*110*225	28.00	0.565	7.90	0.160
CT20±3	225*110*225	20.75	0.404	5.80	0.120
CT27±3	225*110*225	11.20	0.226	4.60	0.093

7.4 Comparative study between the difference in mortar strength and bond shear strength for different thicknesses

As the results show after testing the shear bond strength of brick triplets, the shear capacity of triplets depends on both the parameters i.e., the strength of mortar as well as the thickness of mortars. Table 6 depicts the percentage differences in strength for mortar and triplets. There exists a variation in the strength of mortar between cement mortar and lime mortar which is about 270% after 28 days of strength. This has been illustrated in many studies [21, 22]. When these mortars were used for making the brick triplets with different thicknesses and testing was performed with the help of a hydraulic jack of capacity of 100kN after 28 days of curing period.

Table 5. Comparative Study of mortar strength and triplet shear strength

Strength	**Lime Mortar (MPa)**	**Cement Mortar (MPa)**	**% Difference**
Strength of Mortar	5.51	20.38	269.87%
Shear Bond Strength (13 mm)	7.90	28.00	254.43%
Shear Bond Strength (20 mm)	5.80	20.75	257.76%
Shear Bond Strength (27 mm)	4.60	11.20	143.48%

As per the results, the strength differences in triplets are 254.43%, 257.76%, and 143.48% for 13mm, 20mm, and 27 mm thicknesses respectively. A similar comparison can also be done within

the same mortar for different thicknesses. In both the mortar the difference in shear strength in the case of 13mm and 20 mm is about 35% and in the case of 20mm and 27mm, the difference is more in the case of cement mortar as compared to lime mortar. In the case of 27mm thickness the shear strength difference is less as compared to 13mm and 20mm and in this case, the failure mode was A/1 as shown in Figure 8 (a) in both the mortars and the possible reason was the thickness itself, as the load applied mortar separated from bricks without any cracks or something like that in all three bricks.

Conclusions

This paper reports on the findings of an empirical study conducted to investigate the shear bond strength of masonry when subjected to zero normal stress conditions. The focus of this research was on the lateral load capacity of masonry triplets with different joint thicknesses, which were built utilizing cement and lime mortar as construction materials.

1. The triplets of joint thicknesses 13 mm built using cement or lime mortar had the highest lateral load capacity. The triplets constructed with cement mortar exhibited a maximum lateral capacity of 28.0 kN while the triplets of joint thickness of 27 mm with lime mortar had the lowest lateral load capacity of around 4.60 kN.
2. Cement mortar exhibited the strong shear bond between the brick unit and the mortar component. It was observed that increasing the mortar thickness to 27 mm led to a decrease in the shear bond strength of masonry triplets. Moreover, the deterioration in the shear bond strength was greater in mortars with lower compressive strength, for both cement and lime mortar, than the rate caused by augmented mortar thickness.
3. The hypothesis that the compressive strength of mortar impact load-bearing performance of masonry construction is acknowledged by this study. The results also showed that variations in the strength of the mortar have better effect on the shear bond strength of brickwork than the thickness of the mortar.

References

[1] Theodossopoulos, D., & Sinha, B., A review of analytical methods in the current design processes and assessment of performance of masonry structures. Construction and Building Materials. 41 (2013) 990–1001. https://doi.org/10.1016/j.conbuildmat.2012.07.095

[2] Lourenço, P. B., Oliveira, D. V., Leite, J. C., Ingham, J. M., Modena, C., & da Porto, F., Simplified indexes for the seismic assessment of masonry buildings: International database and validation, Engineering Failure Analysis. 34 (2013) 585–605. https://doi.org/10.1016/j.engfailanal.2013.02.014

[3] Asteris, P. G., Chronopoulos, M. P., Chrysostomou, C. Z., Varum, H., Plevris, V., Kyriakides, N., & Silva, V., Seismic vulnerability assessment of historical masonry structural systems, Engineering Structures. 62–63 (2014) 118–134. https://doi.org/10.1016/j.engstruct.2014.01.031F14

[4] Singh, S. B., & Munjal, P., Bond strength and compressive stress-strain characteristics of brick masonry, Journal of Building Engineering. 9 (2017) 10–16. https://doi.org/10.1016/j.jobe.2016.11.006

[5] Papanicolaou, C., Triantafillou, T., & Lekka, M., Externally bonded grids as strengthening and seismic retrofitting materials of masonry panels, Construction and Building Materials, 25(2011) 504–514. https://doi.org/10.1016/j.conbuildmat.2010.07.018

[6] Franzoni, E., Gentilini, C., Graziani, G., & Bandini, S., Compressive behaviour of brick masonry triplets in wet and dry conditions. Construction and Building Materials. 82 (2015) 45–52. https://doi.org/10.1016/j.conbuildmat.2015.02.052

[7] Bompa, D. V., & Elghazouli, A. Y., Compressive behaviour of fired-clay brick and lime mortar masonry components in dry and wet conditions, Materials and Structures. 53(2020) 60. https://doi.org/10.1617/s11527-020-01493-w
[8] Costigan, A., Pavía, S., & Kinnane, O., An experimental evaluation of prediction models for the mechanical behaviour of unreinforced, lime-mortar masonry under compression, Journal of Building Engineering. 4(2015) 283–294. https://doi.org/10.1016/j.jobe.2015.10.001
[9] Kumarasamy, K., & Nagar, K., Studies on the Strength of Stabilized Mud Block Masonry Using Different Mortar Proportions. (2014)
[10] Alecci, V., Fagone, M., Rotunno, T., & De Stefano, M., Shear strength of brick masonry walls assembled with different types of mortar. Construction and Building Materials. 40 (2013) 1038–1045. https://doi.org/10.1016/j.conbuildmat.2012.11.107
[11] Lee, G., Park, J. H., Lee, C. H., Lee, S. M., & Lee, K., Mechanical and Structural Investigation of Traditional Masonry Systems with Diverse Types of Bricks and Hydrated Lime Mortars, International Journal of Architectural Heritage. (2023) 1–23. https://doi.org/10.1080/15583058.2023.2197405
[12] Boffill, Y., H. Blanco, I. Lombillo, and L. Villegas., Assessment of historic brickwork under compression and comparison with available equations, Construction and Building Materials. 207(2019) 258–72. Retrieved from. https://doi.org/10.1016/j.conbuildmat.2019.02.083
[13] Xuan Wang, Chi Chiu Lam, Vai Pan Iu, Characterization of mechanical behaviour of grey clay brick masonry in China, Construction and Building Materials. Volume 262 (2020) https://doi.org/10.1016/j.conbuildmat.2020.119964.
[14] Bureau of Indian Standards, IS 1514 (1990): Methods of sampling and test for quick lime and hydrated lime. (1990)
[15] Standard, I., Burnt Clay Building Bricks-Methods of Tests Part 2 Determination of Water Absorption (Fourth Revision). (2019) www.standardsbis.in
[16] Bureau of Indian Standards, Burnt Clay Building Bricks-Methods of Tests Part 3 Determination of Efflorescence (Fourth Revision). (2019) www.standardsbis.in
[17] British Standards Institution, Methods of test for masonry : Part 3 : determination of initial shear strength. British Standards Institution. (2002)
[18] Alecci, V., Fagone, M., Rotunno, T., & De Stefano, M., Shear strength of brick masonry walls assembled with different types of mortar, Construction and Building Materials. 40 (2013) 1038–1045. https://doi.org/10.1016/j.conbuildmat.2012.11.107
[19] Bosiljkov, V., Uranjek, M., Žarnić, R., & Bokan-Bosiljkov, V., An integrated diagnostic approach for the assessment of historic masonry structures, Journal of Cultural Heritage, 11(2010) 239–249. https://doi.org/10.1016/j.culher.2009.11.007
[20] Mazzotti, C., Sassoni, E., & Pagliai, G., Determination of shear strength of historic masonries by moderately destructive testing of masonry cores. Construction and Building Materials. 54 (2014) 421–431. https://doi.org/10.1016/j.conbuildmat.2013.12.039
[21] McCann, D. M., & Forde, M. C., Review of NDT methods in the assessment of concrete and masonry structures. NDT & E International. 34(2001) 71–84. https://doi.org/10.1016/S0963-8695(00)00032-3
[22] Mojsilović, N., Petrović, M., & Anglada, X. R., Masonry elements with multi-layer bed joints: Behaviour under monotonic and static-cyclic shear, Construction and Building Materials. 100 (2015) 149–162. https://doi.org/10.1016/j.conbuildmat.2015.09.065

International Journal of Microwave and Wireless Technologies

cambridge.org/mrf

A wideband dual polarized in-band full duplex antenna with higher order modes suppression

Sandeep Rana[1,2], Amit Kumar[1] and Akhilesh Mohan[1]

[1]Department of Electronics and Communication Engineering, Indian Institute of Technology Roorkee, Roorkee, India and [2]Department of Electronics and Communication Engineering, G.B. Pant Institute of Engineering and Technology, Pauri Garhwal, India

Research Paper

Cite this article: Rana S, Kumar A, Mohan A (2025) A wideband dual polarized in-band full duplex antenna with higher order modes suppression. *International Journal of Microwave and Wireless Technologies*, 1–11. https://doi.org/10.1017/S1759078725000315

Received: 30 April 2024
Revised: 12 February 2025
Accepted: 27 February 2025

Keywords:
differential patch antenna; dual polarized; higher order modes suppression; in-band full duplex (IBFD)

Corresponding author: Akhilesh Mohan; Email: am@ece.iitr.ac.in

Abstract

This paper presents a wideband dual-polarized in-band full duplex antenna which can also suppress the higher order modes. A dual differential capacitive feeding scheme is proposed to increase the isolation between the co-located transmitter and receiver ports. The proposed design uses $\lambda/4$ resonators in close vicinity of the radiating patch to provide spurious-free characteristics. Further, a parasitic patch placed above the radiating patch at height "h" is used to enhance the input matching, impedance bandwidth, and gain of the proposed design. The key features of the proposed structure are its wide impedance bandwidth (∼20%), high inter-port isolation (>58 dB) throughout the entire operating frequency range (2.26–2.76 GHz), and simultaneous suppression of higher order modes which makes it a suitable candidate for modern wireless application. Finally, a prototype is fabricated and measurement results are in good agreement with simulation results.

Introduction

The rapid advancements in modern wireless technology create a scarcity in the available frequency spectrum. In-band full duplex (IBFD) system is a potential candidate to address this issue where the transmission and reception of message signals are done simultaneously using a single frequency. Theoretically, a full-duplex system can increase the spectrum efficiency two times that of a half-duplex system [1]. However, while transmitting and receiving at the same frequency, the interference caused by a strong transmitting signal on a weaker receiving signal is the main concern that needs to be minimized to make the IBFD system more efficient. To address this issue, several self-interference cancellation (SIC) techniques are reported in the literature [2]. Generally, SIC techniques are collectively implemented in RF/analog [3], digital [4], and antenna domains but an antenna system having high inter-port isolation can further decrease the need for SIC in succeeding stages. Various methods such as the use of a circulator [5], increased antenna separation [6], near field cancellation [7], and beam-forming network [8] are used for achieving high inter-port isolation. Typically, these techniques involve the use of multiple antennas or complex antenna structures, resulting in a bulky form factor or larger physical dimensions.

Considering simplicity in the design process, dual-polarized antennas are preferred for the IBFD system [9]. Several dual-polarized antenna structures are reported in the literature [10–19] which suffer from one or many of the following disadvantages:

(i) Separate radiators are used for the Transmitter (Tx) and Receiver (Rx) which results in a large footprint [13, 14].
(ii) Isolation less than 40 dB is achieved between Tx and Rx which might not be sufficient for an IBFD system [10, 15].
(iii) Impedance bandwidth of the design is limited (<6%) [11–15, 17, 19].
(iv) Peak gain of the design is limited (∼5dBi) [12, 14, 19].

Also, the higher-order modes of an antenna can cause electromagnetic interferences to the different devices of any wireless communication system. Hence, to suppress undesired higher-order modes of the antenna, several research activities have been proposed in recent times. These include exciting a circular patch with common mode signals [20], loading of shorting pins [21], deploying a pair of $\lambda/4$ resonators with capacitive feed [22], and use of multi-point differential feed [23], etc.

The above discussion clearly indicates that all the previous works are designed either for the IBFD applications or for the suppression of higher-order modes. To the best of the authors' knowledge, it is the first instant a shared radiator-based wideband IBFD antenna with enhanced

© The Author(s), 2025. Published by Cambridge University Press in association with The European Microwave Association.

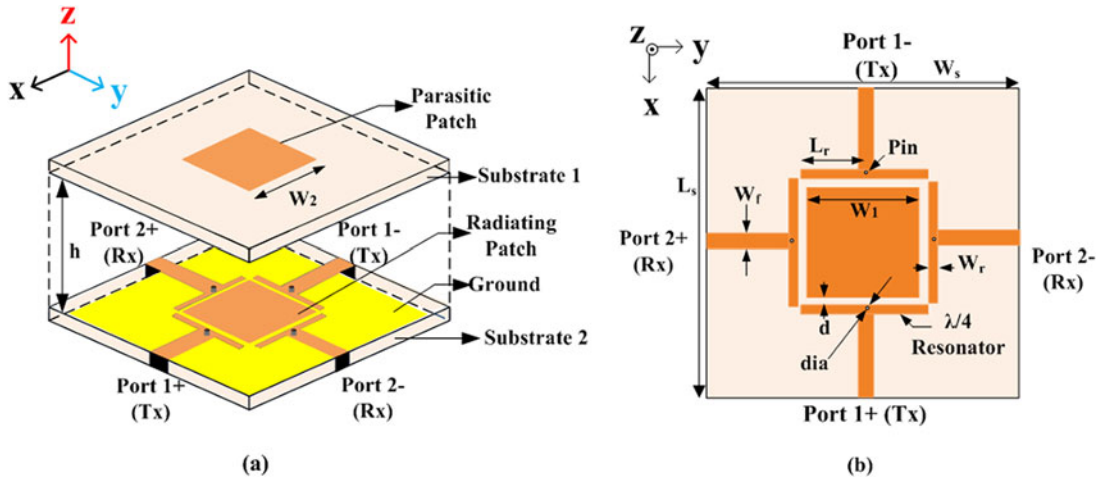


Figure 1. Proposed design with dimensions (in mm): (a) 3 d view, (b) top view of the bottom layer; $l_s = w_s = 100$, $W_1 = 37$, $W_2 = 38.2$, $l_r = 19.2$, $w_f = 4.5$, $w_r = 1.75$, $d = 1$, dia $= 0.3$, $h = 10$.

isolation is presented which can simultaneously suppress the higher-order modes. The major contributions of the proposed work are as follows:

1) To the best of the authors' knowledge, it is the first instance that dual differential capacitive feeding is proposed to implement a wideband IBFD antenna. This proposed feeding scheme has three major functions: (i) suppression of all second-order higher modes; (ii) creation of dual resonance behavior that can be transformed into wideband performance; (iii) enhancement of the inter-port isolation.

2) A mathematical model is presented to illustrate the effect of the double differential feeding on inter-port isolation.

3) To the best of the authors' knowledge, our proposed design is the first dual-polarized IBFD antenna that not only provides stellar full duplex performance (excellent inter-port isolation of greater than 58 dB over a wideband of 20% with significant peak gain of 9.1 dBi) but also suppresses all the spurious second-order higher modes. To suppress higher order modes, we have not employed any filtering circuits (LPF, BPF, BSF) that would increase the overall footprint, defected ground structures that have back radiations, or complex electronic bandgap structures.

4) The functionality of the proposed IBFD antenna is validated through the lumped equivalent circuit model. Moreover, the effectiveness of the proposed circuit model is validated through the parametric studies where the results obtained through circuit simulation closely resemble the EM simulation results.

Antenna configuration

The geometry of the proposed dual-polarized IBFD antenna is depicted in Fig. 1. It consists of two square microstrip patches which are built on Rogers RT 5880 substrate ($\varepsilon_r = 2.2$, tan $\delta = 0.0009$) of thickness 1.575 mm. The radiating patch (RP), built on substrate 2 is excited with dual differential feeding. Port 1+ and 1– are used as differential Tx ports while port 2+ and 2– are used as differential Rx ports. The feeding of RP through each differential port is done with the help of a pair of $\lambda/4$ resonators. A metallic pin having a diameter *dia* is placed at the center of the feedline which is shorted to the ground. The parasitic patch (PP) is placed above RP at a height of h on substrate 1. The ANSYS Electronics Desktop (HFSS) is used to design and optimize the structure.

Operational mechanism

Theoretically, a dual-polarized antenna where Tx and Rx have orthogonal polarization, has infinite isolation between the Tx and Rx port, a favorable condition for the IBFD system. Hence, at Stage 1, a dual-polarized square MPA is excited through direct feeding (insert-fed) at port 1 (Tx port) and port 2 (Rx port). The feeding arrangement shown in Fig. 2(a) ensures that the Tx port is orthogonally polarized with respect to the Rx port. This MPA resonates at a center frequency (f_0) of 2.5 GHz. It is evident from the S-parameters shown in Fig. 2(a) that in addition to fundamental mode TM_{10} (2.5 GHz) of this MPA, other higher order modes TM_{11} (3.55 GHz), TM_{20} (5.3 GHz) and TM_{22} (6.7 GHz) also exist.

Now, instead of direct feeding, capacitive feeding with the help of pairs of $\lambda/4$ resonators is provided to the MPA as shown in Stage 2 of Fig. 2(b). This configuration results in spurious-free characteristics as shown in Fig. 2(b). The suppression of higher-order modes is due to the presence of pairs of $\lambda/4$ resonators which are centrally loaded with a shorting pin [22].

The non-radiating $\lambda/4$ resonators along with a shorting pin exhibit resonance at discrete frequencies (nf_0, where $n = 1, 3, .. (2\,N + 1)$). The MPA is capacitively fed by $\lambda/4$ resonators, allowing energy transmission only at discrete frequencies where both the $\lambda/4$ resonators and MPA resonate simultaneously. Consequently, it is not possible to excite higher-order modes like TM_{11} and TM_{22} due to their resonance frequencies not being a multiple of f_0. For a square MPA, the resonant frequency of the TM_{11} and TM_{22} modes are $1.414\,f_0$ and $2.828\,f_0$, respectively.

In addition, the placement of a shorting pin at the center of a feedline results in the creation of an electric field null at the center plane. This configuration prevents the excitation of even order modes, such as TM_{20}, because the primary requirement to excite any even order mode is to have an electric field maximum in the central plane. Hence, this feeding mechanism shown in Stage 2 of Fig. 2(b) successfully suppresses the second-order modes which are the most critical in any communication system. In spite of good out-of-band performance, this configuration has inter-port

https://doi.org/10.1017/S1759078725000315 Published online by Cambridge University Press

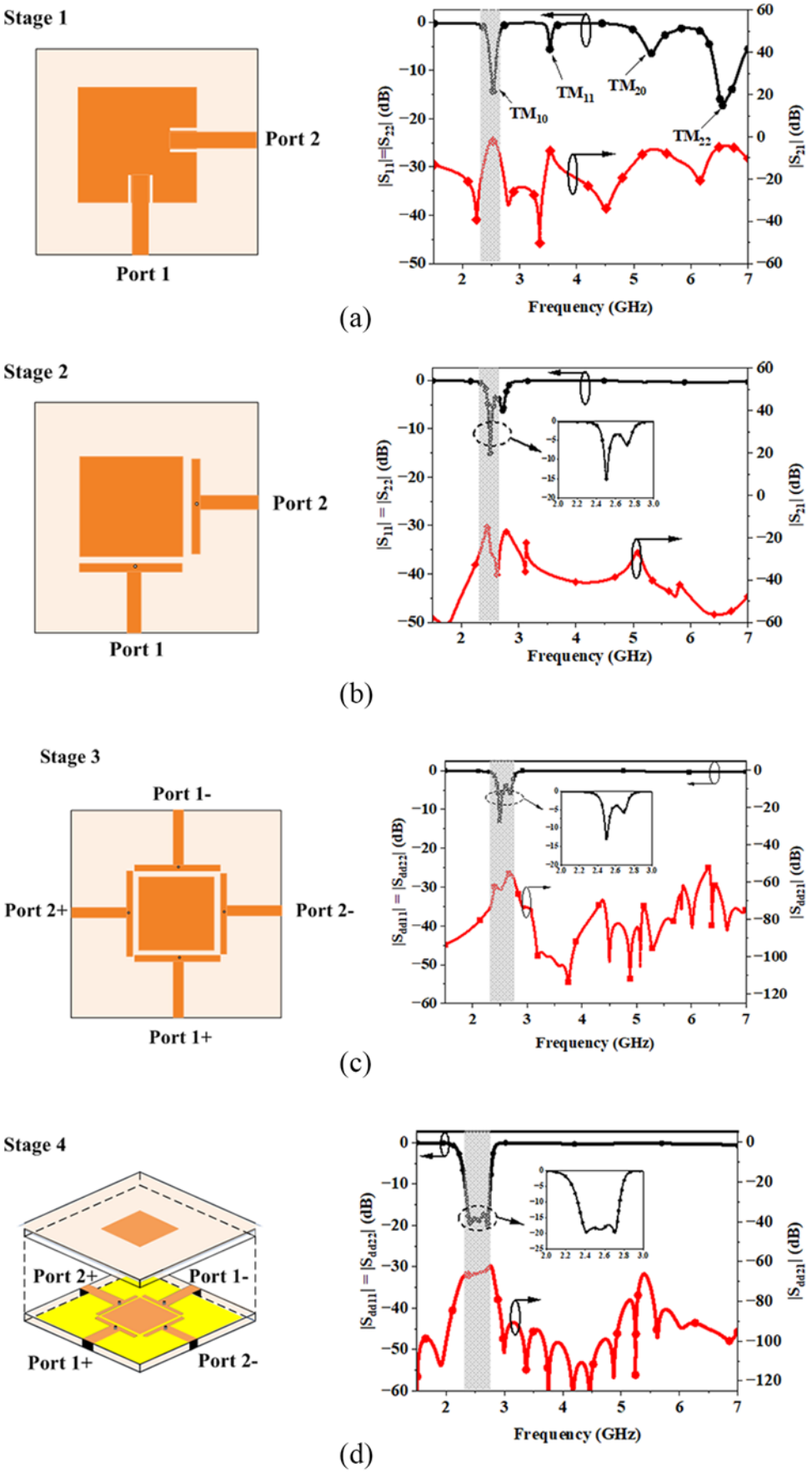


Figure 2. Design evolution of the proposed IBFD antenna and simulated s-parameters: (a) stage 1, (b) stage 2, (c) stage 3, (d) stage 4.

https://doi.org/10.1017/S1759078725000315 Published online by Cambridge University Press

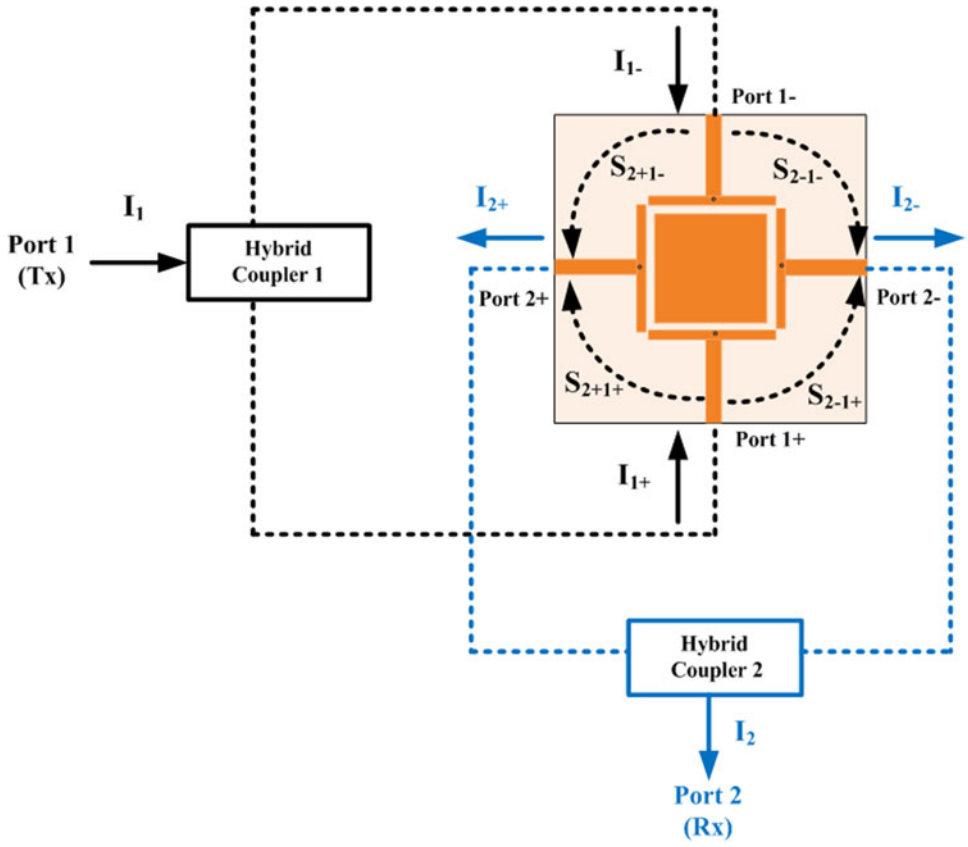


Figure 3. Stage 3 antenna having dual differential feedings realized using identical hybrid couplers.

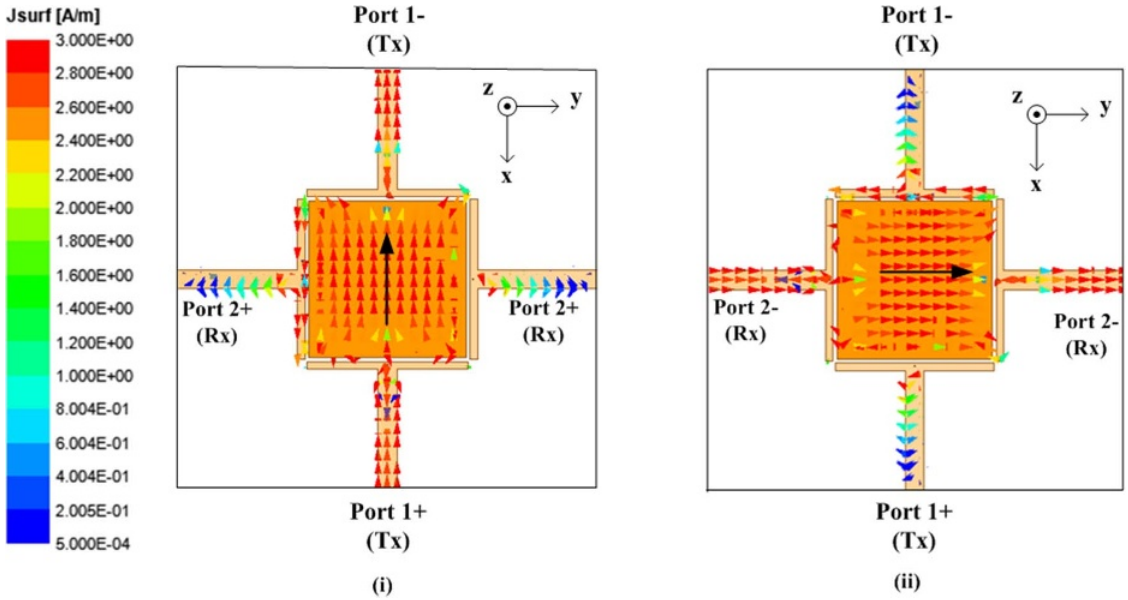


Figure 4. Distribution of surface current on the RP: (i) when differential Tx ports are excited and Rx ports are terminated with matched load, (ii) when differential Rx ports are excited and Tx ports are terminated with matched load.

isolation of 20 dB as shown in Fig. 2(b) which is insufficient for an IBFD system.

Further, to increase the inter-port isolation, differential feeding is provided to Tx and Rx ports which is presented as Stage 3 of Fig. 2(c). To understand the effect of this double differential feeding on inter-port isolation in more detail, a mathematical model is developed based on the analysis proposed in [13]. To apply the differential feed at both the Tx and Rx ports, two different hybrid couplers can be utilized as shown in Fig. 3. Here, the current at Port 1(Tx) and Port 2(Rx) is represented as I_1 and I_2 respectively.

https://doi.org/10.1017/S1759078725000315 Published online by Cambridge University Press

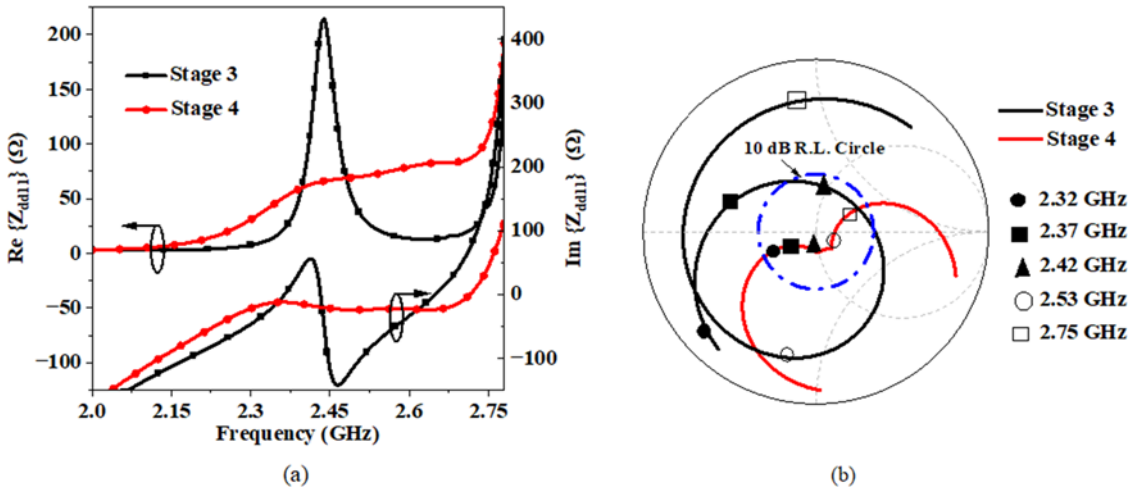


Figure 5. (a) Input impedance matching of stage 3 and stage 4 antenna, (b) Smith chart for impedance matching of stage 3 and stage 4 antenna.

Table 1. Key features of antennas of different stages

Stage	Type of feeding	Higher order modes suppression	Inter-port isolation	Impedance bandwidth
1	Insert feed	No	Very low	Narrow
2	Capacitive feed	Yes	Average	Narrow
3	Dual differential capacitive feed	Yes	Excellent	Narrow
4	Dual differential capacitive feed	Yes	Excellent	Large

Assuming the lossless hybrid coupler (H.C.), the output currents of H.C. 1 can be written as [24]

$$I_{1+} = \frac{1}{\sqrt{2}} I_1 \; and \; I_{1-} = -\frac{1}{\sqrt{2}} I_1 \tag{1}$$

As the differential port 2 is acting as a receiver port, different currents of H.C. 2 are related to each other as follows:

$$I_2 = \frac{1}{\sqrt{2}} (I_{2+} - I_{2-}) \tag{2}$$

Now, let the mutual coupling between port 2_+ (2-) and port 1_+(1-) is represented as S_{2+1+} ($S_{2\text{-}1\text{-}}$) and S_{2+1-} ($S_{2\text{-}1+}$) denotes the mutual coupling between port 2_+ (2–) and port 1_- (1+), then

$$I_{2+} = \frac{I_1}{\sqrt{2}} S_{2+1+} - \frac{1}{\sqrt{2}} I_1 \, S_{2+1-}$$

$$= \frac{I_1}{\sqrt{2}} [S_{2+1+} - S_{2+1-}] \tag{3}$$

$$I_{2-} = \frac{I_1}{\sqrt{2}} [S_{2-1+} - S_{2-1-}] \tag{4}$$

Now using eq. (2), (3) and (4), I_2 can be found as

$$I_2 = \frac{I_1}{2} [(S_{2+1+} - S_{2+1-}) - (S_{2-1+} - S_{2-1-})] \tag{5}$$

Therefore, based on Eq. (5) the relationship between the current of port 1 and 2 can be expressed as

$$\frac{I_1}{I_2} = \frac{2}{[(S_{2+1+} - S_{2+1-}) - (S_{2-1+} - S_{2-1-})]} \tag{6}$$

The proposed structure is symmetrical and considering the ideal differential operation performed by the hybrid couplers, the coupling between the orthogonally placed ports will be identical ($S_{2+1+} = S_{2+1-} = S_{2-1+} = S_{2-1-}$). Hence, it is evident from Eq. (6) that theoretically, infinite inter-port isolation is achieved using this configuration. Nevertheless, in practical use, significantly large inter-port isolation can be achieved by utilizing highly precise hybrid couplers that produce output signals with equal magnitude and out-of-phase characteristics [13].

It is important to note that double differential feeding deployed at Stage 3 is the major factor in achieving high inter-port isolation. The differential operation performed at the Tx port is responsible for suppressing any Tx-Rx leakage at Rx ports while the differential Rx ports subsequently suppress any further residual self-interference. Hence, the polarization diversity along with double differential feeding yields a high inter-port isolation better than 60 dB for the entire operating band as shown in Fig. 2(c).

The surface currents for Tx and Rx ports are shown in Fig. 4. It is evident that the surface current is oriented in the x (y) direction when differential Tx (Rx) ports are excited while the Rx (Tx) ports are terminated with a matched load. Even though high inter-port isolation is achieved at this design stage, the input matching of this stage is very poor, and impedance bandwidth ($|S_{11}| < -10\ dB$) is limited (~40 MHz).

Although the dual resonant behavior can be observed in antennas of Stage 2 and Stage 3 due to close proximity of $\lambda/4$ resonators with the RP (Fig. 2(b) and Fig. 2(c)), the input matching and impedance bandwidth are not sufficient. Therefore, a PP is placed above the RP in Stage 4 as shown in Fig. 2(d). The inclusion of PP

https://doi.org/10.1017/S1759078725000315 Published online by Cambridge University Press

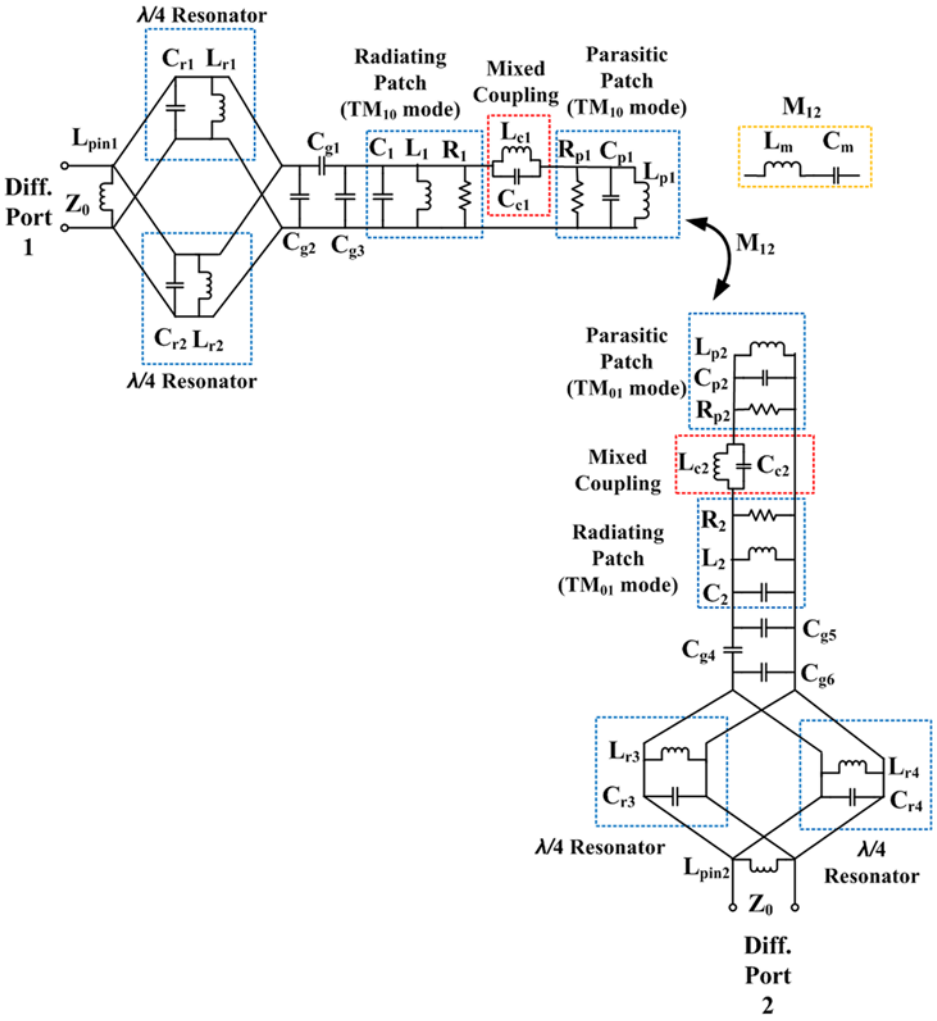


Figure 6. Lumped circuit equivalent model of the proposed antenna: $L_{pin1} = L_{pin2} = 72.5\ nH,\ L_{r1} = L_{r2} = L_{r3} = L_{r4} = 0.82\ nH,\ R_1 = R_2 = 328\ \Omega,\ C_{r1} = C_{r2} = C_{r3} = C_{r4} = 4.14\ pF,\ C_{g1} = C_{g4} = 1.77\ pF,\ C_{g2} = C_{g3} = C_{g5} = C_{g6} = 0.28\ pF, L_1 = L_2 = 0.95\ nH, C_1 = C_2 = 4.94\ pF, L_{c1} = L_{c2} = 0.23\ nH\ C_{c1} = C_{c2} = 12.86\ pF, R_{p1} = R_{p2} = 15\ \Omega, L_{p1} = L_{p2} = 0.28\ nH, C_{p1} = C_{p2} = 13.37\ pF.$

Table 2. Relationship of circuit elements used in ADS simulation with their counterpart in EM simulation

Element in circuit modelling	Element in EM modelling
Inductance of the shorting pin (L_{pin})	Dimensions (diameter) of shorting pin
Inductance of the resonator [L_{ri}, L_1, L_2, L_{p1}, L_{p2}]	Dimensions ($L \times W$) of the resonator [$\lambda/4$ resonator, RP(TM_{10}) RP (TM_{01}), PP (TM_{10}), PP (TM_{01})]
Capacitance of the resonator [C_{ri}, C_1, C_2, C_{p1}, C_{p2}]	Dimensions ($L \times W$) of the resonator [$\lambda/4$ resonator, RP(TM_{10}) RP (TM_{01}), PP (TM_{10}), PP (TM_{01})]
Equivalent capacitance formed by π network (C_{gi})	Gap (d) between the $\lambda/4$ resonator and RP
Inductance (L_c) and capacitance (C_c)	Mixed coupling facilitated by airgap (h) between RP and PP

has two-fold utility; first, it improves the input matching and then it also improves the impedance bandwidth. It is clear from the S-parameters shown in Fig. 2(d) that the Stage 4 antenna results in a high isolation of 65 dB over the entire operating frequency band (2.30–2.76 GHz).

To illustrate the improvement in input matching, the input impedances of Stage 3 and Stage 4 antennas are plotted in Fig. 5(a). It clearly shows that the Stage 3 antenna exhibits inadequate input matching over the desired operating band.

Now, the air gap between RP and PP facilitates the mixed coupling (electric and magnetic couplings) [23]. The effect of this mixed coupling can be observed in Fig. 5(a). This mixed coupling offers either inductive or capacitive impedance across different frequencies. Consequently, an air gap of varying height results in a substantial change in input impedance (the corresponding parametric study is shown in the Section V of this paper). Hence, with the placement of the PP above RP at an optimal height (h) of 10 mm, the imaginary part of Z_{dd11} of the Stage 4 antenna approaches zero, while the real component of Z_{dd11} is near to 50 Ω in the entire operating band. This demonstrates the significant improvement of input matching resulting from the strategic placement of PP at an optimal height of h.

https://doi.org/10.1017/S1759078725000315 Published online by Cambridge University Press

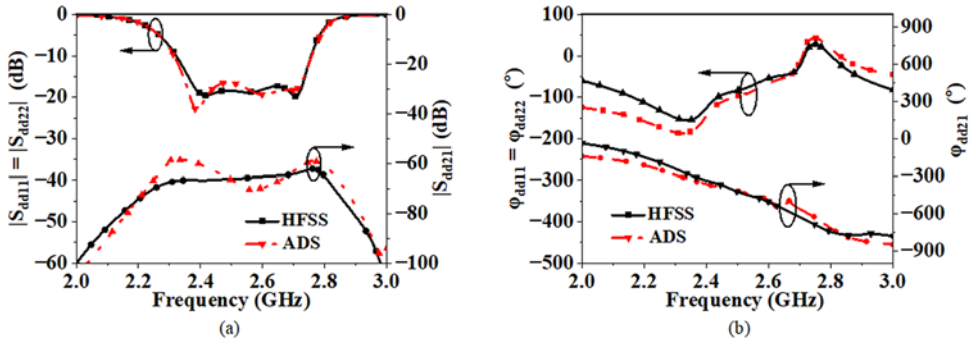


Figure 7. Comparison between EM simulation and circuit simulation of the proposed IBFD antenna: (a) magnitude plot (b) phase plot (unwrapped).

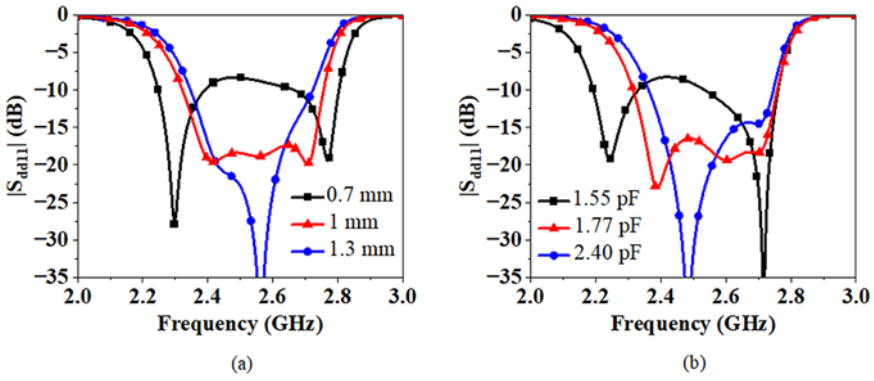


Figure 8. The effect of variations in (a) d (b) C_{g1} on the input reflection coefficient.

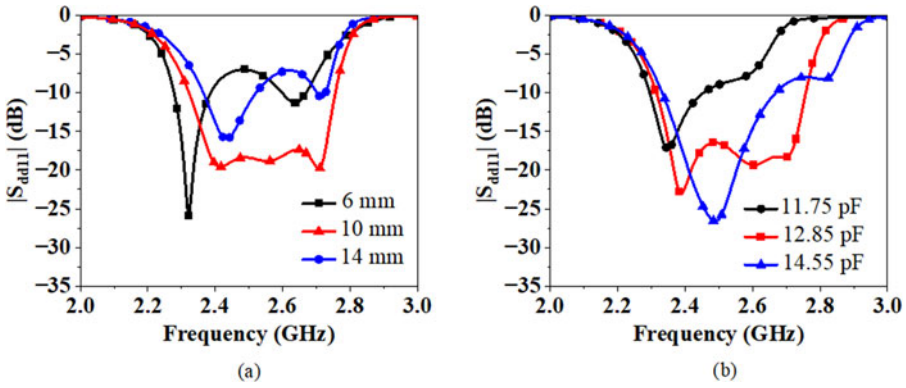


Figure 9. The effect of variations in (a) h (b) c_{c1} on the input reflection coefficient.

Additionally, the excellent input matching for the entire operating band is also reflected in the Smith Chart shown in Fig. 5(b). The impedance loop of the Stage 4 antenna evidently shifts towards the center of Smith chart which further confirms the substantial improvement in the matching performance due to PP being placed above the RP at a height of h.

Further, the judicious selection of the air gap (h) as 10 mm ($\sim$0.1 λ) facilitates the fundamental mode of PP to resonate with close proximity of the fundamental mode of RP [25]. Thus, the improved impedance bandwidth (460 MHz) is achieved because three different structures i.e. RP, PP, and $\lambda/4$ resonators are resonating in close vicinity. This Stage 4 antenna presented in Fig. 2(d) is our proposed IBFD antenna which has significantly large inter-port isolation over a wide frequency range. Also, this design successfully suppresses the higher-order modes which can be undesirable for other surrounding devices. Table 1 summarizes the key characteristics of antennas of different stages which clearly highlight the achievement of performance improvement of the proposed IBFD antenna in successive stages.

Equivalent lumped circuit model

To gain further insight into the proposed design, an equivalent lumped circuit model of the proposed dual-polarized IBFD antenna is presented in Fig. 6. The differential feeding is provided at both Tx and Rx ports through a microstrip line of characteristic impedance Z_0. The $\lambda/4$ resonator is modeled as a lossless LC circuit (L_{ri}, C_{ri}, i varying from 1 to 4). These pair of $\lambda/4$ resonators at the Tx and Rx ends are connected in parallel. At one end, these resonators are connected to shorting pins shown as L_{pin1} and L_{pin2} at the Tx and Rx ends, respectively. The gap between the feeding circuit and the RP is modeled as a π-network of three capacitances

https://doi.org/10.1017/S1759078725000315 Published online by Cambridge University Press

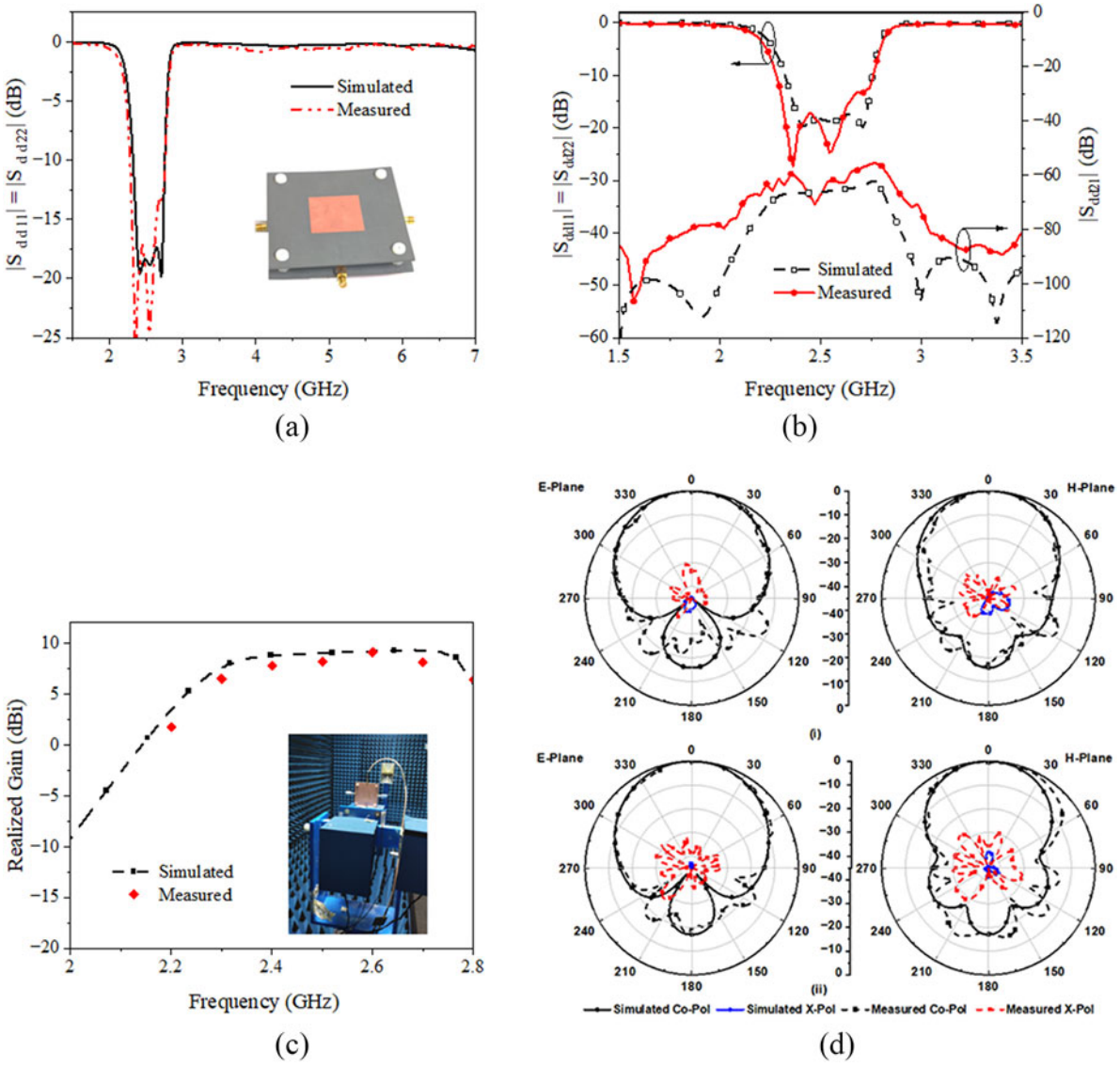


Figure 10. Measured results of the proposed IBFD antenna: (a) out of band performance, (b) reflection coefficients and inter-port isolation, (c) realized gain, (d) e-plane and h-plane radiation pattern of Tx port at (i) 2.4 Ghz (ii) 2.65 Ghz.

Table 3. Comparison of the proposed work with reported works

Ref.	Antenna structure	Fractional bandwidth (%)	Isolation (dB)	Shared/multi-radiator	Peak gain (dBi) (Tx/Rx)	Higher order mode suppression
[10]	MED	12.1	30	S	7/7	N.P.
[11]	MPA	4.1	59.5	S	N.P.	N.P.
[12]	MPA	2.1	70	S	4.2/4.2	N.P.
[13]	MPA	5	80	M	6/6	N.P.
[14]	MPA (2 design)	4.4 & 6.4	64 & 60	M	3.2/2.7 & 4.5/5.1	N.P.
[15]	DPA	5	34	S	8.7/8.7	N.P.
[16]	DPA	18.9	48	S	8.4/7.6	N.P.
[17]	DRA	4.4	50	S	8.9/8.1	N.P.
[18]	MPA	14.1	45	S	9.2/9.4	N.P.
[19]	MPA	4.1	54	S	3.3/2.4	N.P.
Prop.	**MPA**	**20**	**58**	**S**	**9.1/9.1**	**Yes**

DPA: Dielectric Patch Antenna, DRA: Dielectric Resonator Antenna, MED: Magneto Electric Dipole, N.P.: Not Provided

https://doi.org/10.1017/S1759078725000315 Published online by Cambridge University Press

C_{g1}, C_{g2}, C_{g3} and C_{g4}, C_{g5}, C_{g6} at Tx, and Rx ends, respectively. A combination of two parallel RLC circuits i.e. $R_1L_1C_1$ and $R_2L_2C_2$ is used to represent the RP when it is excited at differential Tx and Rx ports respectively. It is important to note that, the RP is operating under TM_{10} mode when excitation is provided at differential port 1 while the dominating mode of operation is TM_{01} when RP is excited at differential port 2. Similarly, the PP is presented as a parallel resonant circuit $R_{p1}L_{p1}C_{p1}$ and $R_{p2}L_{p2}C_{p2}$ when it is excited with differential port 1 and differential port 2, respectively. The resonant frequency (f_r) of each of the three resonators (RP, PP, and $\lambda/4$ resonators) can be calculated with the following relation:

$$f_r = \frac{1}{2\pi\sqrt{L_rC_r}} \tag{7}$$

where L_r and C_r represent the inductance and capacitance of the corresponding resonator.

Further, the mixed coupling (combination of electric and magnetic coupling) between the RP and PP is described as two parallel LC circuits (L_{C1}, C_{C1} and L_{C2}, C_{C2}). Finally, the mutual coupling between port 1 and port 2 is defined as M_{12} [26,27]. The working mechanism of various components used in the circuit model can be understood by comparing them with the actual elements utilized in the design of the IBFD antenna as tabulated in Table 2. In addition, the ADS software is utilized to design and validate the IBFD antenna equivalent circuit model. A comparison of simulated results obtained from ADS and HFSS is presented in Fig. 7, demonstrating a high level of consistency between the two.

Parametric analysis

To elucidate the impact of a few parameters on the performance of the proposed IBFD antenna, two parametric studies are performed. The first study illustrates the effect of variations in d (the gap between a pair of $\lambda/4$ resonators and RP) on the input reflection coefficient, as depicted in Fig. 8 (a). For a nominal value of d as 1 mm, all the three reflection poles due to RP, PP, and $\lambda/4$ resonators are combined to form a wideband antenna. When the value of d is increased to 1.3 mm, all the reflection poles are merged into a single reflection pole. Additionally, to validate the effectiveness of the proposed equivalent circuit model, one capacitor (C_{g1}) of the π network (the circuit equivalent of the gap (d) between a pair of $\lambda/4$ resonators and RP) is varied and its effect on input reflection coefficient is presented in Fig. 8 (b). It is evident from the Fig. 8 that both EM and circuit simulation leads to a similar conclusion. It is important to note that all parameters except d (C_{g1}) in EM (circuit) simulation are maintained constant in the aforementioned study.

Subsequently, in another parametric study, the effect of variations in the air gap (h) between RP and PP is analyzed, as shown in Fig. 9(a). It clearly demonstrates that the judicious selection of the airgap (h) as 10 mm is essential for the fundamental mode of PP to resonate with close proximity of the fundamental mode of RP which results in a good impedance matching over a wideband. Furthermore, as discussed in the previous section, the air gap between RP and PP facilitates the mixed coupling (electric and magnetic couplings) to improve the input matching. This can be further verified from the parametric variation in C_{c1} (parallel combination of C_{c1} and L_{c1} is circuit equivalent to mixed-coupling provided by the air gap between RP and PP) performed in circuit simulation as shown in Fig. 9(b). Hence, the prudent selection of air gap to obtain a good input matching over a wideband is justified with the help of EM and circuit simulation as illustrated in Fig. 9.

Results and discussion

To validate the functionality of the proposed design, a prototype is implemented using RT Duroid substrate, and the nylon screws are used to hold both substrate layers. The fabricated prototype is shown in the inset of Fig. 10(a). To provide differential signals at Tx and Rx ports, an ultra-wideband hybrid coupler proposed in [9] is used. The magnitude and phase imbalance between the two output ports of this coupler is within the limit ($\pm$0.2 dB and $\pm$2.5° respectively) over a large frequency range (1–10 GHz).

The S – parameters of the proposed dual-polarized IBFD antenna are measured using the Agilent Vector Network Analyzer [N5247A] and presented in Fig. 10. The improved out-of-band characteristic of the proposed structure can be observed in Fig. 10(a) which clearly shows that it possesses spurious free characteristics. Moreover, as shown in Fig. 10(b), the measured impedance bandwidth is ~20% (2.26–2.76 GHz) which is somewhat larger than the simulated results. Also, the measured isolation between Tx and Rx ports is better than 58 dB for the entire operating range (2.26–2.76 GHz) while in simulation it was obtained as 65 dB. The discrepancies between the simulated and measured results are due to cable losses and misalignment of the PP with respect to the driven patch.

The far-field parameters are measured in an anechoic chamber and the measured realized gain and radiation pattern are presented in Fig. 10(c) and Fig. 10(d), respectively. It is important to note that the proposed design is symmetrical with respect to Tx and Rx ports. Thus, the measured results are presented only for the differential Tx ports for brevity. The peak values of measured realized gain and radiation efficiency are 9.1 dBi and 92%, respectively. The cross-polarization better than −30 dB at boresight direction is measured at two operating frequencies 2.4 GHz and 2.65 GHz in both E and H planes. It is evident from the measured results that these are in close agreement with the simulated results.

A detailed comparison of the proposed structure with other reported dual-polarized IBFD antennas is presented in Table 3. The present design offers the largest fractional bandwidth ($|S_{11}| < -10$ dB) of 20% among all the designs listed in Table 3. Though the designs presented in [13] and [14] offer better interport isolation, these multi-radiator-based designs have very narrow bandwidths and possess lower values of peak gain. Moreover, none of the aforementioned IBFD structures provide suppression of higher-order modes. Hence, it can be concluded that the proposed shared radiator-based design provides significantly larger isolation of 58 dB over a wide bandwidth of 20% along with a large peak gain of 9.1 dBi and an excellent out-of-band performance.

Conclusion

In this paper, a wideband IBFD antenna is presented which is also capable of suppressing higher-order modes. The employability of $\lambda/4$ resonators along with PP in wideband operation. An isolation better than 58 dB between the Tx and Rx ports is achieved due to orthogonally polarized surface currents applied through a balanced feed to the Tx and Rx ports. The key attributes of the proposed design are simplicity in design, very high isolation over a large frequency band, and excellent out-of-band performance. Due to these features, the proposed design not only provides excellent performance in the ISM band but can also be scaled to higher frequency bands.

https://doi.org/10.1017/S1759078725000315 Published online by Cambridge University Press

Data availability statement. Data sharing is not applicable to this article as no data sets were generated or analyzed during this study.

Acknowledgements. The authors acknowledge the Department of Electronics and Communication Engineering, Indian Institute of Technology Roorkee, Uttarakhand, India, for providing equipment and laboratory facilities to carry out this work.

Funding statement. This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Competing interests. The authors report no conflict of interest.

References

1. **Sabharwal A, Schniter P, Guo D, Bliss DW, Rangarajan S and Wichman R** (2014) In-band full-duplex wireless: Challenges and opportunities. *IEEE Journal on Selected Areas in Communications* **32**(9), 1637–1652. doi:10.1109/JSAC.2014.2330193
2. **Kolodziej KE, Perry BT and Herd JS** (2019) In-Band Full-Duplex Technology: Techniques and Systems Survey. *IEEE Transactions on Microwave Theory & Techniques* **67**(7), 3025–3041. doi:10.1109/TMTT.2019.2896561
3. **King B, Xia J and Boumaiza S** (2018) Digitally assisted RF-analog self-interference cancellation for wideband full-duplex radios. *IEEE Transactions on Circuits and Systems II: Express Briefs* **65**(3), 336–340.
4. **Cheng X, Chen FJ, Zhang L, Gao H, Han JA, Han JY, Yu Y and Deng XJ** (2021) A closed-loop reconfigurable analog baseband circuitry with open-loop tunable notch filters to improve receiver Tx leakage and close-in blocker tolerance. *IEEE Transactions on Circuits and Systems II: Express Briefs*, **69**(3), 839–843.
5. **Prasannakumar PV, Elmansouri MA and Filipovic DS** (2018) Broadband reflector antenna with high isolation feed for full-duplex applications. *IEEE Transactions on Antennas and Propagation* **66**(5), 2281–2290. doi:10.1109/TAP.2018.2814224
6. **Everett E, Sahai A and Sabharwal A** (2014) Passive self-interference suppression for full-duplex infrastructure nodes. *IEEE Transactions on Wireless Communications* **13**(2), 680–694. doi:10.1109/TWC.2013.010214.130226
7. **Wu YZ, Zang HL, Li M and Behdad N** (2019) A compact, low profile simultaneous transmit and receive antenna with monopole-like radiation characteristics. *IEEE Antennas and Wireless Propagation Letters* **18**(4), 611–615. doi:10.1109/LAWP.2019.2898220
8. **Ha J, Elmansouri MA, Prasannakumar PV and Filipovic DS** (2017) Monostatic co-polarized full-duplex antenna with left- or right-hand circular polarization. *IEEE Transactions on Antennas and Propagation* **65**(10), 5103–5111. doi:10.1109/TAP.2017.2741064
9. **Ta SX and Trong NN** (2022) Analysis and Design of an Ultrawideband Dual-Polarized Antenna for IBFD Applications. *IEEE Transactions on Antennas and Propagation* **70**(11), 11121–11126. doi:10.1109/TAP.2022.3195558
10. **Zhang ZY and Wu KL** (2018) A wideband dual-polarized dielectric magnetoelectric dipole antenna. *IEEE Transactions on Antennas and Propagation* **66**(10), 5590–5591. doi:10.1109/TAP.2018.2859914
11. **Prasad DVS, Singh HV, Tripathi S and Paltani P** (2023) In-band full duplex co-radiator antenna with high isolation using differential feed and dual polarization for 2.4 GHz WLAN. *IEEE Antennas and Wireless Propagation Letters* **22**(11), 2740–2744. doi:10.1109/LAWP.2023.3297599
12. **Nawaz H, Ahmad N and Aslam J** (2021) A unidirectional, printed antenna with high inter-port isolation over wider bandwidth for 2.4 GHz full duplex applications. *IEEE Transactions on Antennas and Propagation* **69**(11), 7183–7191. doi:10.1109/TAP.2021.3069570
13. **Chaudhary G, Jeong J and Jeong Y** (2019) Differential Fed Antenna with High Self-Interference Cancellation for In-Band Full-Duplex Communication System. *IEEE Access* **7**, 45340–45348. doi:10.1109/ACCESS.2019.2909118
14. **Wang X-Y, Tang S-C, Yang LL and Chen J-X** (2020) Differential-fed dual-polarized dielectric patch antenna with gain enhancement based on higher order modes. *IEEE Antennas and Wireless Propagation Letters* **19**(3), 502–506. doi:10.1109/LAWP.2020.2964569
15. **Sangam RS, Poolakkal S, Palthiya R, Nallam N and Kshetrimayum RS** (2019) Dual-port, aperture coupled and tapered fed patch antenna for full-duplex ISM applications. *Microwave and Optical Technology Letters* **61**, 542–545. doi:10.1002/mop.31593
16. **Abedian M, Khalily M, Xiao P, Wang F, Tafazolli R and Kishk AA** (2023) MM-Wave High Isolated Dual Polarized Dielectric Resonator Antenna for In-Band Full-Duplex Systems. *IEEE Access* **11**, 38218–38225. doi:10.1109/ACCESS.2023.3266558
17. **Nawaz H, Basit MA and Niazi AU** (2021) A compact, dual-polarized patch antenna with improved Tx–Rx isolation for 2.4 GHz single frequency full-duplex applications. *International Journal of Microwave and Wireless Technologies* **13**, 266–273. doi:10.1017/S1759078720000537
18. **Chu C, Wang M, Wang J, Guo Y and Wu W** (2023) A New Design of Filtering Patch Antennas With Enhanced Bandwidth and Harmonic Suppression. *IEEE Transactions on Antennas and Propagation* **71**(7), 6120–6125. doi:10.1109/TAP.2023.3266059
19. **Yun Y, Liu Y, Wu Z and Lennerstad H** (2021) Harmonic suppression and size reduction of a low-profile CP patch antenna by using nondegenerate modes. *IEEE Antennas and Wireless Propagation Letters* **21**(3), 456–460. doi:10.1109/LAWP.2021.3135300
20. **Zhang J-D, Zhu L, Wu Q-S, Liu N-W and Wu W** (2016) A compact microstrip-fed patch antenna with enhanced bandwidth and harmonic suppression. *IEEE Transactions on Antennas and Propagation* **64**(12), 5030–5037. doi:10.1109/TAP.2016.2618539
21. **Huang F, Zhu L and Zhang H** (2022) Study on a multi-point differential feeding strategy for design of filtering patch antennas with stopband enhancement. *IEEE Transactions on Antennas and Propagation* **70**(12), 11293–11300. doi:10.1109/TAP.2022.3209184
22. **Pozar DM** (1998) Microwave Engineering. *Boston, MA, USA: Artech House*.
23. **Nishiyama E, Aikawa M and Egashira S** (2004) Stacked microstrip antenna for wideband and high gain. *IEE Proceedings-Microwaves, Antennas and Propagation* **151**(2), 143–148. doi:10.1049/ip-map:20040171
24. **Pradhan NC, Subramanian KS, Barik RK and Cheng QS** (2021) A shielded-QMSIW-based self-diplexing antenna for closely spaced bands and high isolation. *IEEE Antennas and Wireless Propagation Letters* **20**(12), 2382–2386. doi:10.1109/LAWP.2021.3112610
25. **Naseri H, PourMohammadi P, Melouki N, Ahmed F, Iqbal A and Denidni TA** (2023) Substrate integrated waveguide-based dual-polarized self-diplexing antenna array. *IEEE Transactions on Circuits and Systems II: Express Briefs*, **70**(8), 2839–2843.
26. **Zhang XK, Ke YH, Wang XY, Tang SC and Chen JX** (2023) Broadband Dual-Polarized Dielectric Patch Antenna With High Isolation for Full-Duplex Communication. *IEEE Antennas and Wireless Propagation Letters* **22**(4), 878–882. doi:10.1109/LAWP.2022.3227239
27. **Chen JH, Ye L, Liu T and Wu D-L** (2023) A Low-Profile Dual-Polarized Patch Antenna With Simple Feed and Multiple Decoupling Techniques. *IEEE Antennas and Wireless Propagation Letters* **22**(8), 1883–1887. doi:10.1109/LAWP.2023.3268344

https://doi.org/10.1017/S1759078725000315 Published online by Cambridge University Press



Sandeep Rana received his M.Tech. degree from Delhi Technological University, New Delhi, India, in 2014. He is presently enrolled as a full-time Ph.D. research scholar in the Department of Electronics & Communication Engineering at the Indian Institute of Technology Roorkee, Uttarakhand, India. He is on academic leave from his position as an Assistant Professor in the ECE Department at G. B. Pant Institute of Engineering and Technology (GBPIET), Uttarakhand, where he has been serving as a faculty member since 2019. Prior to his tenure at GBPIET, he served as a faculty member at Graphic Era Hill University, Dehradun, and Netaji Subhash Institute of Technology, New Delhi, India. He has nine publications till date including four journal and five conferences. His research interests revolve around the design of microwave circuits and antennas tailored for modern wireless communication applications.

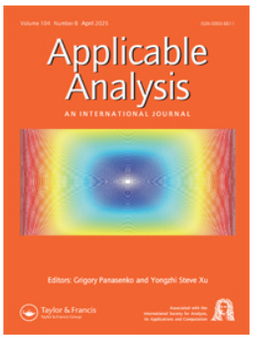


Amit Kumar is currently pursuing his Ph.D. from Electronics and Communication Engineering (ECE) Department, Indian Institute of Technology Roorkee, India. He completed his M.Tech in ECE Department at National Institute of Technology Warangal, India, in 2022. He has worked as research intern at Wipro, India during his M.Tech. His current areas of research include design of linear and efficient RF Power amplifiers, design of microwave circuits and antennas for V2X communication. He has five publications till date.



Akhilesh Mohan received the B.Tech. degree in Electronics Engineering from Kamla Nehru Institute of Technology, Sultanpur, India, in 2002, the M.Tech. degree and Ph.D. degree in Microwave Engineering from Indian Institute of Technology Kanpur, Uttar Pradesh, India, in 2004 and 2009, respectively. From 2009 to 2010, he worked as a Scientist at Space Applications Center, Indian Space Research Organization, Ahmedabad, India. From 2010 to 2013, he was a faculty member with the Department of Electrical Engineering at Indian Institute of Technology Jodhpur, Rajasthan, India. From 2013 to 2020, he was a faculty member with the Department of Electronics and Electrical Communication Engineering at Indian Institute of Technology Kharagpur, West Bengal, India. Since 2021, he has been a faculty member with the Department of Electronics and Communication Engineering at Indian Institute of Technology Roorkee, Uttarakhand, India. Dr. Mohan has authored and co-authored more than 150 peer reviewed international journals and conference papers. His research interests include the design of microwave filters, antennas, and absorbers for wireless communication systems.

https://doi.org/10.1017/S1759078725000315 Published online by Cambridge University Press



Applicable Analysis
An International Journal

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/gapa20

Bernstein type semi-exponential operators

Kanita & Naokant Deo

To cite this article: Kanita & Naokant Deo (01 Apr 2025): Bernstein type semi-exponential operators, Applicable Analysis, DOI: 10.1080/00036811.2025.2480642

To link to this article: https://doi.org/10.1080/00036811.2025.2480642

Published online: 01 Apr 2025.

Submit your article to this journal

Article views: 22

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=gapa20

APPLICABLE ANALYSIS
https://doi.org/10.1080/00036811.2025.2480642

Taylor & Francis
Taylor & Francis Group

Check for updates

Bernstein type semi-exponential operators

Kanita and Naokant Deo

Department of Applied Mathematics, Delhi Technological University, Delhi, India

ABSTRACT

The concept of semi-exponential operators is an extension of the exponential operators in approximation theory. The analysis begins by considering the partial differential equation,

$$\frac{\partial}{\partial x}W_\beta(r,x,t) = \frac{r(t-x)}{p(x)}W_\beta(r,x,t) - \beta W_\beta(r,x,t)$$

with domain of definition Ω, for the kernel of sequence of operators

$$S_r(f;x) = \sum_{k\in\Omega} W_\beta\left(r,x,\frac{k}{r}\right) f\left(\frac{k}{r}\right)$$

while satisfying the normalization condition $\sum_{k\in\Omega} W_\beta\left(r,x,\frac{k}{r}\right) = 1$. Carrying forward this notion, we examine these operators specifically for the function $p(x) = x(1-x)$ and focus our attention on the semi-exponential Bernstein operators. Our primary focus is to explore the properties and characteristics of these Bernstein type semi-exponential operators. The paper delves into deriving a recurrence relation for the central moments of these operators, while also establishing a moment generating function for the same. Finally, we analyze the approximation of these operators using the Voronovskaya-type asymptotic result. The objective of this analysis is to characterize functions by the degree of approximation.

ARTICLE HISTORY
Received 12 October 2023
Accepted 13 March 2025

COMMUNICATED BY
T. Suslina

KEYWORDS
Semi-exponential operators; Bernstein polynomials; moment generating function; modulus of continuity

XXXX CLASSIFICATION

1. Introduction

In their influential study, Ismail and May were the first to propose the idea of exponential operators [1], wherein they defined the exponential-type operators,

$$S_\lambda(f(u);t) = \int_\Omega W(\lambda,t,u)f(u)\mathrm{d}u, \tag{1}$$

where Ω is the domain of definition and W is a positive function, known as the kernel of S_λ, satisfying the homogeneous partial differential equation

$$\frac{\partial}{\partial t}W(\lambda,t,u) = \frac{\lambda}{p(t)}W(\lambda,t,u)(u-t), \tag{2}$$

CONTACT Naokant Deo naokantdeo@dce.ac.in; Kanita kanitatewari@gmail.com

© 2025 Informa UK Limited, trading as Taylor & Francis Group

such that $p(t)$ is both analytic and positive over the domain Ω. The kernel $W(\lambda, t, u)$ satisfies the normalization condition

$$S_\lambda\,(1,t) = \int_\Omega W\,(\lambda,t,u)\,\mathrm{d}u = 1. \tag{3}$$

Assigning particular values to the function $p(t)$ in Equation (2), various operators can be obtained. For instance, the Bernstein polynomials can be obtained by considering the function $p(t) = t(1-t)$. Similarly, the operators of Szász-Mirakjan can be derived by the function $p(t) = t$, the Post–Widder operators by $p(t) = t^2$ and the Gauss–Weierstrass operators by $p(t) = 1$, among others.

For quite a few years following Ismail and May, no new work had been done on exponential operators up until the year 2005. It was during this time that A. Tyliba and E. Wachnicki [2] introduced the concept of semi-exponential operators and they expanded upon the findings of Ismail and May by considering a sequence of operators, denoted as V_r^β, such that $V_r^\beta(t,x) \neq x$. In Equation (2), they introduced a non-negative real parameter β and thus worked on a new homogeneous partial different equation, thus obtaining a new family of operators, known as the semi-exponential operators, defined in the following way:

$$V_r^\beta\,\big(f(t);x\big) = \int_\Omega W_\beta\,(r,x,t)\,f(t)\mathrm{d}t, \tag{4}$$

where Ω is the domain of definition. Here, we denote the kernel by $W_\beta(r,x,t)$, which satisfies:

(1) The homogeneous partial differential equation

$$\frac{\partial}{\partial x}W_\beta(r,x,t) = \frac{r}{p(x)}\,(t-x)\,W_\beta(r,x,t) - \beta W_\beta(r,x,t). \tag{5}$$

(2) The normalization condition

$$V_r^\beta\,(1,x) = \int_\Omega W_\beta\,(r,x,t)\,\mathrm{d}t = 1. \tag{6}$$

For the case $\beta = 0$, the operators V_r^β can simply be regarded as the operators of the exponential type. Recently, Herzog [3] conducted a study on the approximation properties of exponential-type operators when acting on functions from exponential weighted spaces. A range of several new semi-exponential operators were captured by Abel et al. recently [4], each corresponding to a different function $p(x)$. Further, in the past few years, numerous investigations have been carried out concerning the modification and preservation of exponential-type operators [5–14].

In all the definitions mentioned above, the domain of each operator is defined as Ω, which varies depending on the specific operator. This paper will address the case in which the polynomial function $p(x) = x(1-x)$ is considered and the function $f(k/r)$ is discretely varied over the interval $[0,\infty)$. Due to the discrete nature of the distribution, the aforementioned definitions and normalization condition will vary when summed over the discrete range of k from 0 to ∞. Solving the partial differential Equation (5) for $p(x) = x(1-x)$, and satisfying the normalization condition (6), we get the kernel of the semi-exponential Bernstein operators, $W_\beta(r,k,x)$ as follows,

$$W_\beta(r,k,x) = e^{-\beta x}x^k(1-x)^{r-k}\sum_{l=0}^{\infty}{}^{r-l}c_{k-l}\frac{\beta^l}{l!}, \tag{7}$$

where ${}^rc_k = \frac{r!}{k!(r-k)!}$ is the binomial coefficient. Solving the partial differential Equation (2) corresponding to exponential operators we get the kernel of the classical Bernstein operators, known as

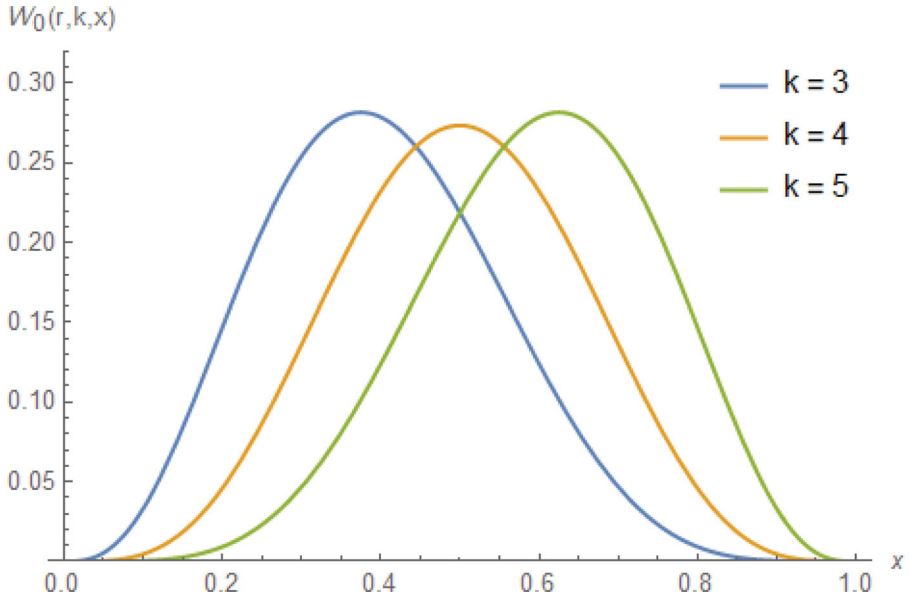


Figure 1. Graphs of the kernel of semi-exponential Bernstein operators for $r = 8$, $k = 3, 4, 5$ and $\beta = 0$.

the Bernstein polynomials, defined as $b_{r,k}(x) = {}^{r}c_k x^k (1-x)^{r-k}$. We see that the relation between the kernel of the semi-exponential Bernstein operators (7) and the kernel of the exponential Bernstein operators is given by

$$W_0(r,k,x) = b_{r,k}(x). \tag{8}$$

In this paper we focus on a whole new class of kernels $W_\beta(r,k,x)$ depending upon r and k and the parameter β, unlike the classical Bernstein polynomials $b_{r,k}(x)$ which solely depend upon r and k. To further analyze these functions and determine the effect of the parameters on $W_\beta(r,k,x)$, we graph them for different values of r, k and β.

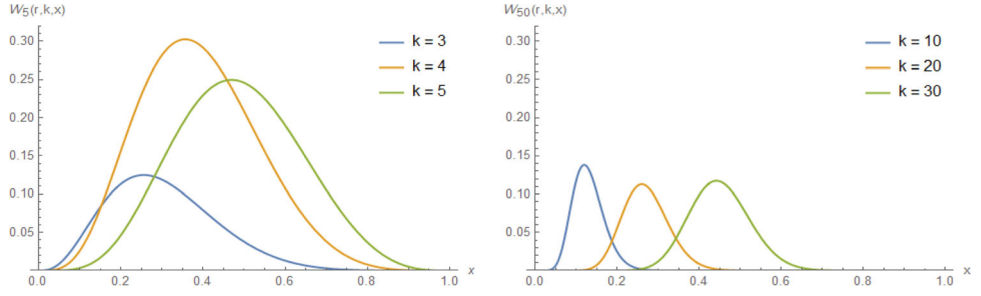
Figure 1 validates expression (8) as it is well known that the Bernstein polynomials are symmetric in the interval $[0,1]$, that is, $b_{r,k}(1-x) = b_{r,r-k}(x)$. However, from Figure 2, we can see that the kernel of the semi-exponential operators for $\beta > 0$ does not satisfy this property. Thus, we can conclude that unlike the kernels of classical Bernstein operators and its other modifications, such as the λ-Bernstein operators [15] and α-Bernstein [16], the kernel of semi-exponential operators defined by (4) for $\beta > 0$, are not symmetric over $[0,\infty]$.

Employing $W_\beta(r,k,x)$ to the definition of semi-exponential operators (4), we get the Bernstein type semi-exponential operators:

$$B_r^\beta(f;x) = e^{-\beta x}\sum_{k=0}^{\infty} x^k (1-x)^{r-k} \sum_{l=0}^{\infty} {}^{r-l}c_{k-l}\frac{\beta^l}{l!} f\left(\frac{k}{r}\right), \tag{9}$$

where ${}^{r}c_k$ is the binomial coefficient. The above operators can be demonstrated and derived as shown by Abel et al. [4] in their paper, where they introduced several new semi-exponential operators, including the semi-exponential Bernstein operators (9). Furthermore, they introduced the following alternate form of these operators (10), which is used in this paper to compute the moments and central moments of these operators,

$$B_r^\beta(f;x) = e^{-\beta x}\sum_{l=0}^{\infty} \frac{\beta^l}{l!} x^l (1-x)^{r-l} \sum_{j=0}^{\infty} {}^{r-l}c_j \left(\frac{x}{1-x}\right)^j f\left(\frac{l+j}{r}\right). \tag{10}$$



(a) Graphs of the kernel of semi-exponential Bernstein operators for $r = 8$, $k = 3, 4, 5$ and $\beta = 5$.

(b) Graphs of the kernel of semi-exponential Bernstein operators for $r = 40$, $k = 10, 20, 30$ and $\beta = 50$.

Figure 2. Graphs of the kernel of semi-exponential Bernstein operators for different values of r and k and $\beta > 0$. (a) Graphs of the kernel of semi-exponential Bernstein operators for $r = 8$, $k = 3, 4, 5$ and $\beta = 5$ and (b) Graphs of the kernel of semi-exponential Bernstein operators for $r = 40$, $k = 10, 20, 30$ and $\beta = 50$.

Based on the definition and derivation of the semi-exponential Bernstein polynomials (4), it can be stated that the operators denoted as B_r^β map a function f on $[0, \infty)$ to a function $B_r^\beta f$ defined on $[0, \frac{1}{2})$, whenever the sum is convergent.

2. Preliminaries

In the following section, we present the properties of the operators mentioned in (10), which are utilized in the subsequent sections of this paper to establish the main theorems.

Lemma 2.1: *Let $B_r^\beta(.;x)$ be the Bernstein type semi-exponential operators. Then for $x \in [0, \frac{1}{2})$ the following equalities hold:*

(i) $B_r^\beta(1;x) = 1$

(ii) $B_r^\beta(t;x) = x + \beta\frac{x(1-x)}{r}$

(iii) $B_r^\beta(t^2;x) = x^2 + \frac{x(1-x)+2\beta x^2(1-x)}{r} + \frac{\beta x(1-3x+2x^2)+\beta^2x^2(1-x)^2}{r^2}$

Proof: Applying the negative binomial expansion in (10) we get,

$$B_r^\beta(1;x) = e^{-\beta x}\sum_{l=0}^{\infty}\frac{\beta^l}{l!}x^l(1-x)^{r-l}\sum_{j=0}^{\infty}{}^{r-l}c_j\left(\frac{x}{1-x}\right)^j$$

$$= e^{-\beta x}\sum_{l=0}^{\infty}\frac{\beta^l}{l!}x^l(1-x)^{r-l}\left(1+\frac{x}{1-x}\right)^{r-l} = 1.$$

$$B_r^\beta(t;x) = e^{-\beta x}\sum_{l=0}^{\infty}\frac{\beta^l}{l!}x^l(1-x)^{r-l}\sum_{j=0}^{\infty}{}^{r-l}c_j\left(\frac{x}{1-x}\right)^j\left(\frac{l+j}{r}\right)$$

$$= e^{-\beta x}\left[\sum_{l=0}^{\infty}\frac{\beta^l}{l!}x^l(1-x)^{r-l}\sum_{j=0}^{\infty}{}^{r-l}c_j\left(\frac{x}{1-x}\right)^j\frac{l}{r}\right.$$

$$\left.+\sum_{l=0}^{\infty}\frac{\beta^l}{l!}x^l(1-x)^{r-l}\sum_{j=0}^{\infty}{}^{r-l}c_j\left(\frac{x}{1-x}\right)^j\frac{j}{r}\right]$$

$$= e^{-\beta x}\left[\sum_{l=0}^{\infty}\frac{\beta^l}{l!}x^l\frac{l}{r}+\sum_{l=0}^{\infty}\frac{\beta^l}{l!}x^l(1-x)^{r-l}\left(\frac{r-l}{r}\right)\sum_{j=0}^{\infty}{}^{r-l-1}c_j\left(\frac{x}{1-x}\right)^{j+1}\right]$$

$$= e^{-\beta x}\left[\sum_{l=0}^{\infty}\frac{\beta^l}{l!}x^l\frac{l}{r}+x\sum_{l=0}^{\infty}\frac{\beta^l}{l!}x^l\left(1-\frac{l}{r}\right)\right]$$

$$= e^{-\beta x}\left[\frac{\beta x}{r}e^{\beta x}+x\left(1-\frac{\beta x}{r}\right)e^{\beta x}\right]$$

$$= x+\frac{\beta x}{r}(1-x)$$

and

$$B_r^{\beta}\left(t^2;x\right) = e^{-\beta x}\sum_{l=0}^{\infty}\frac{\beta^l}{l!}x^l(1-x)^{r-l}\sum_{j=0}^{\infty}{}^{r-l}c_j\left(\frac{x}{1-x}\right)^j\left(\frac{l^2+2lj+j^2}{r^2}\right)$$

$$= e^{-\beta x}\left[\sum_{l=0}^{\infty}\frac{\beta^l}{l!}x^l\frac{l^2}{r^2}+2\sum_{l=0}^{\infty}\frac{\beta^l}{l!}\frac{l}{r}x^l(1-x)^{r-l}\sum_{j=0}^{\infty}{}^{r-l}c_j\frac{j}{r}\left(\frac{x}{1-x}\right)^j\right.$$

$$\left.+\sum_{l=0}^{\infty}\frac{\beta^l}{l!}x^l(1-x)^{r-l}\sum_{j=0}^{\infty}{}^{r-l}c_j\left(\frac{x}{1-x}\right)^j\left(\frac{j}{r}\right)^2\right]$$

$$= e^{-\beta x}\left[\sum_{1}+\sum_{2}+\sum_{3}\right].$$

Solving the above three summations, we will arrive at,

$$\sum_{1} = e^{\beta x}\frac{\beta x}{r^2}(1+\beta x)$$

$$\sum_{2} = e^{\beta x}\frac{2\beta x^2}{r}\left(1-\frac{1}{r}-\frac{\beta x}{r}\right)$$

$$\sum_{3} = e^{\beta x}\left(\frac{\beta^2x^4}{r^2}-\frac{\beta x^2}{r^2}-2\left(\frac{r-1}{r^2}\right)\beta x^3+\left(\frac{r-1}{r}\right)x^2+\frac{x}{r}\right).$$

Substituting these values in $B_r^{\beta}(t^2;x)$ we get the desired moment. ■

Now, let us denote the ς th central moment as,

$$\mu_{r,\varsigma}(x) = e^{-\beta x}\sum_{k=0}^{\infty}A(r,k,\beta)\,x^k(1-x)^{r-k}\left(\frac{k}{r}-x\right)^{\varsigma}, \tag{11}$$

where $A(r,k,\beta)=\sum_{l=0}^{\infty}{}^{r-l}c_{k-l}\frac{\beta^l}{l!}$.

Theorem 2.2: *Based on the aforementioned definition, a recurrence relation can be derived for the central moments of semi-exponential Bernstein operators,*

$$\begin{aligned}&\mu_{r,1}(x) = \frac{\beta}{r}x(1-x)\\ &r\mu_{r,\varsigma+1}(x) = x(1-x)\left[\mu'_{r,\varsigma}(x)+\beta\mu_{r,\varsigma}(x)+\varsigma\,\mu_{r,\varsigma-1}(x)\right];\quad \varsigma\geq 1.\end{aligned} \tag{12}$$

Theorem 2.3: *$A_\varsigma(r,x) = r^\varsigma \mu_{r,\varsigma}(x)$ is a polynomial of degree $[\frac{\varsigma}{2}]$ in r, where [.] is the greatest integer function. Also, the coefficient of r^ς in*

(i) *$A_{2\varsigma}(r,x)$ is $1\cdot 3\cdot 5\ldots(2\varsigma-1)(x(1-x))^\varsigma$, and*
(ii) *$A_{2\varsigma+1}(r,x)$ is $c_1\beta(x(1-x))^{\varsigma+1} + c_2(x(1-x))^\varsigma(1-2x)$*

where c_1, c_2 are constants and $c_1 = 1\cdot 3\cdot 5\ldots(2\varsigma+1)$.

Proof: We have $A_\varsigma(r,x) = r^\varsigma \mu_{r,\varsigma}(x)$. Therefore, we can write,

$$A_0(r,x) = 1,$$
$$A_1(r,x) = \beta x(1-x)$$

and

$$A_{\varsigma+1}(r,x) = x(1-x)\left(\frac{\mathrm{d}}{\mathrm{d}x}A_\varsigma(r,x) + \beta A_\varsigma(r,x) + r\varsigma A_\varsigma(r,x)\right), \quad \varsigma \geq 1. \tag{13}$$

By repeatedly substituting the value of $A_\varsigma(r,x)$ in Equation (13), it can be demonstrated that $A_{2\varsigma}(r,x)$ and $A_{2\varsigma+1}(r,x)$ are polynomials in r of degree ς. It follows that, the highest powered term appearing in their expression is r^ς with its coefficient as:

$$1\cdot 3\cdot 5\ldots(2\varsigma-1)(x(1-x))^\varsigma \quad \text{for } A_{2\varsigma}(r,x)$$
$$c_1\beta(x(1-x))^{\varsigma+1} + c_2(x(1-x))^\varsigma(1-2x) \quad \text{for } A_{2\varsigma+1}(r,x)$$

where c_1, c_2 are constants and $c_1 = 1\cdot 3\cdot 5\ldots(2\varsigma+1)$. ■

Remark 2.1: From Theorem 2.3 we have,

$$A_\varsigma(r,x) = O\left(r^{[\varsigma/2]}\right).$$

It follows

$$\mu_{r,\varsigma}(x) = O\left(r^{-\left[\frac{\varsigma+1}{2}\right]}\right),$$

where [.] is the greatest integer function.

One way to calculate the central moments of semi-exponential Bernstein operators is through the recurrence relation stated above in Theorem 2.2. An alternative method for calculating the central moments involves utilizing the exponential moment-generating function (MGF). Our next objective is to derive an equation for the central moment generating function of the semi-exponential operators (10).

Consider the recurrence relation (12),

$$r\mu_{r,\varsigma+1}(x) = x(1-x)\left[\mu'_{r,\varsigma}(x) + \beta\mu_{r,\varsigma}(x) + \varsigma\mu_{r,\varsigma-1}(x)\right].$$

It follows

$$\frac{r}{x(1-x)}B_r^\beta\left((t-x)^{\varsigma+1};x\right) = \frac{\mathrm{d}}{\mathrm{d}x}B_r^\beta\left((t-x)^\varsigma;x\right) + \beta B_r^\beta\left((t-x)^\varsigma;x\right) + \varsigma B_r^\beta\left((t-x)^{\varsigma-1};x\right). \tag{14}$$

Now, the exponential moment generating function of $B_r^{\beta}((t-x)^{\varsigma};x)$ is defined as,

$$\begin{aligned} G(x,z) &= B_r^{\beta}\left(e^{z(t-x)};x\right) \\ &= \sum_{\varsigma=0}^{\infty}\frac{z^{\varsigma}}{\varsigma!}B_r^{\beta}\left((t-x)^{\varsigma};x\right). \end{aligned}$$

Multiplying Equation (14) by $\frac{z^{\varsigma}}{\varsigma!}$ and taking summation over $\varsigma = 0, 1, \ldots \infty$, we get,

$$\frac{r}{x(1-x)}\sum_{\varsigma=1}^{\infty}\frac{z^{\varsigma-1}}{(\varsigma-1)!}B_r^{\beta}\left((t-x)^{\varsigma};x\right) = \frac{\partial}{\partial x}G(x,z) + \beta G(x,z) + zG(x,z).$$

It follows

$$\frac{r}{x(1-x)}\frac{\partial}{\partial z}G(x,z) = \frac{\partial}{\partial x}G(x,z) + (\beta+z)G(x,z)$$

and therefore

$$\frac{r}{x(1-x)}\frac{\partial G}{\partial z} - \frac{\partial G}{\partial x} = (\beta+z)G.$$

For the above PDE, solving the auxiliary equation: $\frac{x(1-x)}{r}\mathrm{d}z = -\mathrm{d}x = \frac{1}{(\beta+z)G}\mathrm{d}G$, we get

$$z = -r\ln\left(\frac{x}{1-x}\right) + c_1 \tag{15}$$

and

$$\begin{aligned} \ln G + c' &= -\beta x - c_1 x + r\int_{\alpha}^{x}\ln\left(\frac{\theta}{1-\theta}\right)\mathrm{d}\theta \\ &= -\beta x - c_1 x + r\int_{\theta=\alpha}^{x}\int_{\vartheta=\alpha}^{\theta}\frac{1}{\vartheta(1-\vartheta)}\mathrm{d}\vartheta\,\mathrm{d}\theta \\ &= -\beta x - c_1 x + r\int_{\vartheta=\alpha}^{x}\int_{\theta=\vartheta}^{x}\frac{1}{\vartheta(1-\vartheta)}\mathrm{d}\theta\mathrm{d}\vartheta \\ &= -\beta x - c_1 x + r\left[x\int_{\alpha}^{x}\frac{1}{\vartheta(1-\vartheta)}\mathrm{d}\vartheta - \left\{\int_{\alpha}^{x}\frac{1}{(1-\vartheta)}\right\}\mathrm{d}\vartheta\right], \end{aligned}$$

where c_1 and c' are arbitrary constants. It follows

$$G = \phi(c_1)\exp\{r\ln(1-x) - (\beta+z)x\},$$

where ϕ is an arbitrary function whose value we need to calculate. Now, by substituting the value of c_1 obtained from the first characteristic curve (15), we get

$$G(x,z) = \phi\left(z + r\ln\left(\frac{x}{1-x}\right)\right)\exp\{r\ln(1-x) - (\beta+z)x\}. \tag{16}$$

Putting the initial condition, $G(x,0) = 1$, we arrive at

$$\phi\left(r\ln\left(\frac{x}{1-x}\right)\right) = \exp\{\beta x - r\ln(1-x)\}$$

It follows

$$\phi\left(ry\right) = \exp\left\{\beta\frac{e^{y}}{1+e^{y}} - r\ln\left(1-\frac{e^{y}}{1+e^{y}}\right)\right\}$$
$$= \exp\left\{\beta\frac{e^{y}}{1+e^{y}} + r\ln\left(1+e^{y}\right)\right\}.$$

Thus

$$\phi\left(t\right) = \exp\left\{\beta\frac{e^{t/r}}{1+e^{t/r}} + r\ln\left(1+e^{t/r}\right)\right\}.$$

Substituting this value of the function ϕ in Equation (16), we arrive at the final solution, thus obtaining our moment generating function,

$$G\left(x,z\right) = \left(1-x+xe^{z/r}\right)^{r}\exp\left\{-zx + \frac{\beta x(1-x)(e^{z/r}-1)}{1-x+xe^{z/r}}\right\}. \tag{17}$$

3. Asysmptotic results

Theorem 3.1: *Let $R > 0$ and a, b be such that $0 < a, b < \infty$. Then for a sufficiently large r, we have,*

$$\left\|B_r^{\beta}\left(e^{R|t|};x\right)\right\|_{C[a,b]} < \infty,$$

$\|f\|_{C[a,b]} = \sup_{x\in[a,b]}|f(x)|$ represents the supremum norm and $C[a,b]$ is the space of all real-valued continuous functions on $[a,b]$ endowed with the supremum norm.

Proof: As per the definition of semi-exponential operators and Equation (17), we can write,

$$B_r^{\beta}\left(e^{Rt};x\right) = e^{Rx}\sum_{\varsigma=0}^{\infty}\frac{R^{\varsigma}}{\varsigma\,!}B_r^{\beta}\left((t-x)^{\varsigma};x\right)$$
$$= \left(1-x+xe^{R/r}\right)^{r}\exp\left\{\frac{\beta x(1-x)(e^{R/r}-1)}{1-x+xe^{R/r}}\right\}.$$

Taking the limit $r\to\infty$, we get

$$\lim_{r\to\infty}B_r^{\beta}\left(e^{Rt};x\right) = \lim_{r\to\infty}\left(1-x+xe^{R/r}\right)^{r} = e^{Rx}.$$

Hence, $\lim_{r\to\infty}B_r^{\beta}(e^{Rt};x) = e^{Rx}$ uniformly in $C[a,b]$. Thus, for a sufficiently large r, we get $\|B_r^{\beta}(e^{Rt};x)\|_{C[a,b]} < \infty$. And similarly, we can show that $\|B_r^{\beta}(e^{-Rt};x)\|_{C[a,b]} < \infty$.

Hence,

$$\|B_r^{\beta}(e^{R|t|};x)\|_{C[a,b]} \le \|B_r^{\beta}(e^{Rt};x)\|_{C[a,b]} + \|B_r^{\beta}(e^{-Rt};x)\|_{C[a,b]}.$$

Thus,

$$\left\|B_r^{\beta}\left(e^{R|t|};x\right)\right\|_{C[a,b]} < \infty.$$

■

Theorem 3.2: *Let $x \in (0,\infty)$. Let f be a continuous function such that $f''(x)$ exists. Then,*

$$\lim_{r\to\infty} r\left[B_r^{\beta}\left(f(t);x\right)-f(x)\right]=\beta x(1-x)f'(x)+\frac{1}{2}x(1-x)f''(x).$$

Proof: Using Taylor's expansion, we have

$$f(t)=f(x)+(t-x)f'(x)+\frac{1}{2}(t-x)^2f''(x)+\varepsilon(t,x)(t-x)^2,$$

where $\lim_{t\to x}\varepsilon(t,x)=0$. Taking $t=k/r$, we can write

$$\begin{aligned}B_r^{\beta}\left(f(t);x\right)-f(x)&=f'(x)B_r^{\beta}\left((t-x);x\right)+\frac{f''(x)}{2}B_r^{\beta}\left((t-x)^2;x\right)\\&\quad+B_r^{\beta}\left(\varepsilon(t,x)(t-x)^2;x\right)\\&=\frac{\beta x(1-x)}{r}f'(x)+\left[\frac{x(1-x)}{r}+\frac{\beta x(1-x)(1-x-x^2)}{r^2}\right]\frac{f''(x)}{2}\\&\quad+\sum_{k=0}^{\infty}W_{\beta}(r,x,k)\,\varepsilon\left(\frac{k}{r},x\right)\left(\frac{k}{r}-x\right)^2,\end{aligned}\tag{18}$$

where $\lim_{t\to x}\varepsilon(t,x)=0$.

Now, we will prove that

$$r\sum_{k=0}^{\infty}W_{\beta}(r,x,k)\,\varepsilon\left(\frac{k}{r},x\right)\left(\frac{k}{r}-x\right)^2\to 0\quad\text{as } r\to\infty.\tag{19}$$

Using Holder's inequality,

$$\begin{aligned}&r\sum_{k=0}^{\infty}W_{\beta}(r,x,k)\,\varepsilon\left(\frac{k}{r},x\right)\left(\frac{k}{r}-x\right)^2\\&\le\left(\sum_{k=0}^{\infty}W_{\beta}(r,x,k)\,\varepsilon^3\left(\frac{k}{r},x\right)\right)^{1/3}\left(r^{3/2}\sum_{k=0}^{\infty}W_{\beta}(r,x,k)\left(\frac{k}{r}-x\right)^3\right)^{2/3}.\end{aligned}$$

Since $\lim_{t\to x}\varepsilon(t,x)=0$, we get that $\sum_{k=0}^{\infty}W_{\beta}(r,x,k)\varepsilon^3(\frac{k}{r},x)\to 0$, as $r\to\infty$.

Also, from Remark 2.1, we have

$$\begin{aligned}\left(r^{3/2}\sum_{k=0}^{\infty}W_{\beta}(r,x,k)\left(\frac{k}{r}-x\right)^3\right)^{2/3}&=\left(r^{3/2}\mu_{r,3}(x)\right)^{2/3}\\&=\left(r^{3/2}O\left(\frac{1}{r^2}\right)\right)^{2/3}\\&=O\left(\frac{1}{r^{1/3}}\right)\to 0.\end{aligned}$$

It follows that the limit (19) is true. Thus, using (18) we get the conclusion. ■

4. Categorizing functions according to their level of approximation

The focus of this section is to categorize functions in a manner such that

$$\left\| B_r^{\beta}(f;x) - f(x) \right\| \leq O\left(r^{-\alpha/2}\right).$$

Definition 4.1: Let $f : \langle a,b\rangle \to \mathbb{R}$ and $h > 0$, where $\langle a,b\rangle$ stands for all possibilities of the open/closed intervals (a,b), $[a,b]$, $(a,b]$ and $[a,b)$. Then the modulus of continuity of the function f, of order one is defined as

$$\omega\left(f;h,a,b\right) = \sup_{\substack{t,t+\delta\in\langle a,b\rangle \\ |\delta|\leq h}} \left|f(t) - f(t+\delta)\right|$$

and the modulus of continuity of second order is defined as

$$\omega_2\left(f;h,a,b\right) = \sup_{\substack{t,t+2\delta,\in\langle a,b\rangle \\ |\delta|\leq h}} \left|f(t) - 2f(t+\delta) + f(t+2\delta)\right|,$$

where $\langle a,b\rangle$ collectively denotes the intervals (a,b) or $[a,b]$ or $[a,b)$ etc.

We know, $C[0,\infty)$ denotes the set of all real-valued continuous functions on $[0,\infty)$. In this section, we define the space $(C_R, \|.\|_{C_R})$ as follows:

$$C_R = \left\{ f \in C\left[0,\infty\right) : \exists M > 0 \forall t \in [0,\infty), \left|f(t)\right| \leq Me^{R|t|} \right\},$$

and

$$\left\|f\right\|_{C_R} = \sup\left\{ e^{-R|t|}\left|f(t)\right| : t \in [0,\infty) \right\}.$$

Lemma 4.2: *Suppose that ς and η are positive numbers, $R \geq 0$ and $[a,b] \subset [0,\infty)$. Then,*

$$\sum_{\mathrm{T}_r} W_{\beta}\left(r,x,\frac{k}{r}\right) e^{R\frac{k}{r}} = O\left(r^{-\varsigma}\right)$$

uniformly on $[a,b]$ as $r \to \infty$, where

$$\mathrm{T}_r = \left\{ k : \left|\frac{k}{r} - x\right| \geq \eta; k = 0,1,\ldots \right\}.$$

Proof: From the definition of T_r and Theorem 2.3, we have that

$$\sum_{\mathrm{T}_r} W_{\beta}\left(r,x,\frac{k}{r}\right) \leq r^{-4\varsigma}\eta^{-4\varsigma} A_{4\varsigma}\left(r,x\right).$$

Also, by Theorem 3.1, as $r \to \infty$, we get

$$\sum_{\mathrm{T}_r} W_{\beta}\left(r,x,\frac{k}{r}\right) e^{2R\frac{k}{r}} < \infty.$$

Thus, by Cauchy–Schwarz's inequality and Remark 2.1,

$$\sum_{\mathrm{T}_r} W_{\beta}\left(r,x,\frac{k}{r}\right) e^{R\frac{k}{r}} \leq \left(\sum_{\mathrm{T}_r} W_{\beta}\left(r,x,\frac{k}{r}\right)\right)^{1/2} \left(\sum_{\mathrm{T}_r} W_{\beta}\left(r,x,\frac{k}{r}\right) e^{2R\frac{k}{r}}\right)^{1/2}$$

$$\leq \left(r^{-4\varsigma}\eta^{-4\varsigma}A_{4\varsigma}(r,x)\right)^{1/2}\left(B_r^{\beta}(e^{2R\frac{k}{r}};x)\right)^{1/2}$$
$$= O\left(r^{-\varsigma}\right) \quad \text{as } r\to\infty.$$

This completes the proof. ■

Theorem 4.3: *Suppose that $f \in C_R$ and $0 < a < a_1 < b_1 < b < \infty$. Then, for every $\varsigma > 0$ there exists a constant K_ς, such that*

$$\left\|B_r^{\beta}(f,x)-f(x)\right\|_{C_{[a_1,b_1]}} \leq K_\varsigma\left[r^{-1/2}\omega\left(f;r^{-1/2},a,b\right)+\omega_2\left(f;r^{-1/2},a,b\right)+r^{-\varsigma}\|f\|_{C_R}\right].$$

Proof: Let $\delta > 0$. Define,

$$g_\delta(x)=\frac{1}{2\delta^2}\int_{-\delta/2}^{\delta/2}\int_{-\delta/2}^{\delta/2}\left[f(x+u+v)+f(x-u-v)\right]\mathrm{d}u\,\mathrm{d}v.$$

By the linearity of our operators B_r^{β}, it follows that

$$\begin{aligned}&\left\|B_r^{\beta}(f;x)-f(x)\right\|_{C_{[a_1,b_1]}}\\ &\leq \left\|B_r^{\beta}(f-g_\delta;x)\right\|_{C_{[a_1,b_1]}}+\left\|B_r^{\beta}(g_\delta;x)-g_\delta(x)\right\|_{C_{[a_1,b_1]}}+\|g_\delta-f\|_{C_{[a_1,b_1]}}\\ &= I_1+I_2+I_3. \end{aligned} \tag{20}$$

For I_1, we consider

$$\left|B_r^{\beta}(f;x)-B_r^{\beta}(g_\delta;x)\right|=\sum_{k=0}^{\infty}W_\beta\left(r,x,\frac{k}{r}\right)\left|f(k/r)-g_\delta(k/r)\right|.$$

Let us define,

$$\Gamma_r=\left\{k:\left|\frac{k}{r}-x\right|<\eta;k=0,1,\ldots\right\}$$

and

$$\mathrm{T}_r=\left\{k:\left|\frac{k}{r}-x\right|\geq\eta;k=0,1,\ldots\right\}.$$

Then,

$$\begin{aligned}&\left|B_r^{\beta}(f;x)-B_r^{\beta}(g_\delta;x)\right|\\ &\leq\sum_{\Gamma_r}W_\beta\left(r,x,\frac{k}{r}\right)\left|f(k/r)-g_\delta(k/r)\right|+\sum_{\mathrm{T}_r}W_\beta\left(r,x,\frac{k}{r}\right)\left|f(k/r)-g_\delta(k/r)\right|\\ &= L_1+L_2.\end{aligned}$$

For L_1, $\frac{k}{r}\in[x-\eta,x+\eta]\subset[a_1-\eta,b_1+\eta]$. Hence, we can state that

$$L_1\leq\frac{1}{2}\omega_2\left(f;\delta,a,b\right).$$

Now, since $f-g_\delta\in C_R$, it follows $|(f-g_\delta)(\frac{k}{r})|e^{-R\frac{k}{r}}\leq 4\|f\|_{C_R}$.

And thus, for a constant M, using Lemma 4.2, we can write L_2 as

$$\begin{aligned} L_2 &= \left| \sum_{\mathrm{T}_r} W_\beta \left(r, x, \frac{k}{r} \right) e^{R\frac{k}{r}} \left| f - g_\delta \right| (k/r)\, e^{-R\frac{k}{r}} \right| \\ &\leq 4Mr^{-\varsigma} \left\| f \right\|_{C_R}. \end{aligned}$$

Hence,

$$\left\| B_r^\beta (f;x) - B_r^\beta (g_\delta;x) \right\|_{C_{[a_1,b_1]}} \leq \frac{1}{2}\omega_2 \left(f;\delta,a,b \right) + 4Mr^{-\varsigma} \left\| f \right\|_{C_R},$$

which gives the following estimate of I_1:

$$I_1 \leq M_1 \left(\omega_2 \left(f;\delta,a,b \right) + r^{-\varsigma} \left\| f \right\|_{C_R} \right). \tag{21}$$

Now, to estimate I_2, we consider the first and second derivative of $g_\delta(x)$

$$\begin{aligned} {g_\delta}'(x) = \frac{1}{2\delta^2} \int_{-\delta/2}^{\delta/2} \left[f\left(x+u+\frac{\delta}{2} \right) - f\left(x+u-\frac{\delta}{2} \right) \right] \mathrm{d}u \\ - \frac{1}{2\delta^2} \int_{-\delta/2}^{\delta/2} \left[f\left(x-u+\frac{\delta}{2} \right) - f\left(x-u-\frac{\delta}{2} \right) \right] \mathrm{d}u. \end{aligned}$$

Hence, for $x \in [a_1 - \eta, b_1 + \eta]$, we have

$$\left\| {g_\delta}'(x) \right\|_{C_{[a_1-\eta,b_1+\eta]}} \leq \frac{1}{\delta}\omega \left(f;\delta,a,b \right).$$

Moreover,

$${g_\delta}''(x) = \frac{1}{\delta^2} \left[f\left(x-\delta \right) - 2f(x) + f\left(x+\delta \right) \right],$$

and

$$\left\| {g_\delta}''(x) \right\|_{C_{[a_1-\eta,b_1+\eta]}} \leq \frac{1}{\delta^2}\omega_2 \left(f;\delta,a,b \right).$$

By the definition of semi-exponential operators and Taylor's expansion, we have

$$\begin{aligned} &\left| B_r^\beta \left(g_\delta;x \right) - g_\delta \left(x \right) \right| \\ &= \left| \sum_{k=0}^{\infty} W_\beta \left(r, x, \frac{k}{r} \right) \left[g_\delta \left(k/r \right) - g_\delta \left(x \right) \right] \right| \\ &\leq \left| \sum_{k=0}^{\infty} W_\beta \left(r, x, \frac{k}{r} \right) \left(\frac{k}{r} - x \right) {g_\delta}' \left(x \right) \right| + \left| \sum_{\Gamma_r} W_\beta \left(r, x, \frac{k}{r} \right) \left(\frac{k}{r} - x \right)^2 {g_\delta}'' \left(x \right) \right| \\ &\quad + \left| \sum_{\mathrm{T}_r} W_\beta \left(r, x, \frac{k}{r} \right) \left(\frac{k}{r} - x \right)^2 {g_\delta}'' \left(x \right) \right| \\ &= S_1 + S_2 + S_3. \end{aligned}$$

For $x \in [a_1 - \eta, b_1 + \eta]$ and some constant M, we obtain,

$$S_1 = \left| g_\delta{}'(x)\, r^{-1} A_1(r,x) \right| \le \frac{M r^{-1}}{\delta} \omega\left(f;\delta,a,b\right).$$

Whereas, for $x \in [a_1, b_1]$ and some constant K,

$$S_2 = \left| g_\delta{}''(x)\, r^{-2} A_2(r,x) \right| \le \frac{K r^{-1}}{\delta^2} \omega_2\left(f;\delta,a,b\right).$$

Following Theorem 2.3 and Lemma 4.2, we can write,

$$\begin{aligned} S_3 &= \left| \sum_{\mathrm{T}_r} W_\beta\left(r,x,\frac{k}{r}\right)\left(\frac{k}{r}-x\right)^2 g_\delta{}''(x) \right| \\ &\le 4\left\|f\right\|_{C_R} \left| \left(\sum_{\mathrm{T}_r} W_\beta\left(r,x,\frac{k}{r}\right) e^{2r} \right)^{1/2} \left(r^{-4} A_4(r,x)\right)^{1/2} \right| \\ &\le L r^{-\varsigma} \left\|f\right\|_{C_R}. \end{aligned}$$

Thus, finally there exists M_2, such that

$$I_2 \le M_2 \left[\frac{r^{-1}}{\delta} \omega\left(f;\delta,a,b\right) + \frac{r^{-1}}{\delta^2} \omega_2\left(f;\delta,a,b\right) + r^{-\varsigma} \left\|f\right\|_{C_R} \right]. \tag{22}$$

To estimate I_3, let $\eta = \frac{1}{2}\min\{a_1 - a, b - b_1\}$.

Then, for $x \in [a_1 - \eta, b_1 + \eta]$, we have

$$\left| f(x) - g_\delta(x) \right| = \left| \frac{1}{2\delta^2} \int_{-\delta/2}^{\delta/2} \int_{-\delta/2}^{\delta/2} \left[f(x+u+v) - 2f(x) + f(x-u-v) \right] \mathrm{d}u\, \mathrm{d}v \right|.$$

By the limits of u and v, we have $|u+v| \le \delta$. Thus, using the definition of second order modulus of continuity,

$$\begin{aligned} \left| f(x) - g_\delta(x) \right| &\le \frac{1}{2\delta^2} \int_{-\delta/2}^{\delta/2} \int_{-\delta/2}^{\delta/2} \omega_2\left(f;\delta,a,b\right) \mathrm{d}u\, \mathrm{d}v \\ &= \frac{1}{2} \omega_2\left(f;\delta,a,b\right). \end{aligned}$$

It follows that

$$I_3 \le \frac{1}{2} \omega_2\left(f;\delta,a,b\right). \tag{23}$$

Putting $\delta = r^{-1/2}$ in estimation (21), (22), (23) and using (20) we get the conclusion. ■

Particularly, this theorem highlights that, if $\omega_2(f;h,a,b) = O(h^\alpha)$ then we can say that $\|B_r^\beta(f;x) - f(x)\|_{C[a_1,b_1]} = O(r^{-\alpha/2})$. The converse of the above statement is also true and is proved next. But first let us define the Zygmund class of functions.

Definition 4.4: Let $0 < \alpha < 2$. The Zygmund class, $Lip^*\alpha$ is defined as

$$Lip^*(\alpha;a,b) = \left\{ f : \omega_2\left(f;h,a,b\right) \le M h^\alpha \right\}.$$

Theorem 4.5: *Let $0 < a < a_1 < b_1 < b < \infty$, $0 < \alpha < 2$ and $f \in C_R$. Then the following are all equivalent statements:*

(i) $f \in Lip^*(\alpha; a_1, b_1)$
(ii) $\|B_r^\beta(f;x) - f(x)\|_{C[a_1,b_1]} = O(r^{-\alpha/2})$

Proof: We have already seen that (*i*) implies (*ii*). We are only left to prove that (*ii*) implies (*i*). Let,

$$\left\| B_r^\beta\left(f;x\right) - f(x) \right\|_{C[a_1,b_1]} \leq M_1 r^{-\alpha/2}.$$

We need to find an upper bound for the first order modulus of continuity of f.

Let $x, y \in [0,\infty)$ be such that $|y - x| \leq h$. Then,

$$\begin{aligned}
&\left|f\left(y\right) - f\left(x\right)\right| \\
&\quad \leq \left|f\left(y\right) - B_r^\beta\left(f;y\right)\right| + \left|B_r^\beta\left(f;y\right) - B_r^\beta\left(f;x\right)\right| + \left|B_r^\beta\left(f;x\right) - f\left(x\right)\right| \\
&\quad \leq 2M_1 r^{-\alpha/2} + P.
\end{aligned}$$

To calculate P, let us use the PDE (5) and examine $|(B_r^\beta)'(f;x)|$. We have

$$\begin{aligned}
\left|\left(B_r^\beta\right)'\left(f;x\right)\right| &= \left|\sum_{k=0}^{\infty} \frac{\partial}{\partial x} W_\beta\left(r,x,k/r\right) f\left(\frac{k}{r}\right)\right| \\
&= \left|\frac{r}{x(1-x)} \sum_{k=0}^{\infty} W_\beta\left(r,x,k/r\right)\left(\frac{k}{r} - x\right) f\left(\frac{k}{r}\right) - \beta \sum_{k=0}^{\infty} W_\beta\left(r,x,k/r\right) f\left(\frac{k}{r}\right)\right| \\
&\leq \left|\frac{r}{x(1-x)} \sum_{k=0}^{\infty} W_\beta\left(r,x,k/r\right)\left(\frac{k}{r} - x\right)\left(f\left(\frac{k}{r}\right) - f\left(x\right)\right)\right| \\
&\quad + \left|\frac{r}{x(1-x)} \sum_{k=0}^{\infty} W_\beta\left(r,x,k/r\right)\left(\frac{k}{r} - x\right) f\left(x\right) - \beta \sum_{k=0}^{\infty} W_\beta\left(r,x,k/r\right) f\left(\frac{k}{r}\right)\right|.
\end{aligned}$$

Also, from Theorem 2.3, we get

$$\begin{aligned}
\sum_{k=0}^{\infty} W_\beta\left(r,x,k/r\right)\left(\frac{k}{r} - x\right) f\left(x\right) &= \frac{A_1\left(r,x\right)}{r} f\left(x\right) \\
&= \frac{\beta}{r} x\left(1-x\right) f\left(x\right).
\end{aligned}$$

Using the property $\omega(\lambda\delta) \leq [1+\lambda]\omega(\delta)$, we can write

$$\omega\left(\left|\frac{k}{r} - x\right|\right) \leq \left[1 + r\left|\frac{k}{r} - x\right|\right] \omega\left(r^{-1}\right). \tag{24}$$

Thus,

$$\left|\left(B_r^\beta\right)'\left(f;x\right)\right| \leq \frac{r}{x(1-x)} \sum_{k=0}^{\infty} W_\beta\left(r,x,k/r\right)\left|\frac{k}{r} - x\right| \omega\left(f;\left|\frac{k}{r} - x\right|, a, b\right)$$

$$
\begin{aligned}
&+\left|\beta f(x)-\beta\sum_{k=0}^{\infty}W_\beta(r,x,k/r)f\left(\frac{k}{r}\right)\right|\\
&\le\frac{r}{x(1-x)}\sum_{k=0}^{\infty}W_\beta(r,x,k/r)\left|\frac{k}{r}-x\right|\omega\left(f;\left|\frac{k}{r}-x\right|,a,b\right)+\beta M_1r^{-\alpha/2}\\
&\le\frac{r}{x(1-x)}\sum_{k=0}^{\infty}W_\beta(r,x,k/r)\left|\frac{k}{r}-x\right|\left[1+r\left|\frac{k}{r}-x\right|\right]\omega\left(f;\left|\frac{1}{r}\right|,a,b\right)\\
&\quad+\beta M_1r^{-\alpha/2}\\
&\le rM_2\omega\left(f;\left|\frac{1}{r}\right|,a,b\right)\sum_{k=0}^{\infty}W_\beta(r,x,k/r)\left|\frac{k}{r}-x\right|\left[1+r\left|\frac{k}{r}-x\right|\right]\\
&\quad+\beta M_1r^{-\alpha/2}\\
&\le rM_2\omega\left(f;\left|\frac{1}{r}\right|,a,b\right)\left[A_1(r,x)+\frac{A_2(r,x)}{r}\right]+\beta M_1r^{-\alpha/2}\\
&\le M_3\left[r\omega\left(f;\left|\frac{1}{r}\right|,a,b\right)+r^{-\alpha/2}\right].
\end{aligned}
$$

Now, for $|y-x|\le h$,

$$
\begin{aligned}
P&=\left|B_r^\beta(f;y)-B_r^\beta(f;x)\right|\\
&=\left|\int_x^y B_r^{\beta'}(f;\xi)\,\mathrm{d}\xi\right|\\
&\le\int_x^y\left|B_r^{\beta'}(f;\xi)\right|\mathrm{d}\xi\\
&\le\int_x^y M_3\left[r\omega\left(f;\left|\frac{1}{r}\right|,a,b\right)+r^{-\alpha/2}\right]\mathrm{d}\xi\\
&=M_3r\omega\left(f;\left|\frac{1}{r}\right|,a,b\right)|y-x|+M_3r^{-\alpha/2}|y-x|\\
&\le M_3rh\omega\left(f;\left|\frac{1}{r}\right|,a,b\right)+M_3r^{-\alpha/2}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\omega(f;h,a,b)&\le|f(y)-f(x)|\\
&\le M'\left[rh\omega\left(f;\left|\frac{1}{r}\right|,a,b\right)+r^{-\alpha/2}\right].
\end{aligned}
$$

Next we will prove, if $\omega(f;h,a,b)\le M'[rh\omega(f;|\frac{1}{r}|,a,b)+r^{-\alpha/2}]$ then, for $0<h\le1$ and $r>1$, we have that $f\in Lip^*(\alpha;a,b)$.

Let $K>1$ be such that $2M'<K^{1-\alpha}$. Choose $M''=\max\{\omega(f;1,a,b),2M'K^\alpha\}$ and define $h_m=K^{-m}$ for $m=1,2,\ldots$.

Using the principle of mathematical induction, we can easily show that for every positive integer m,

$$\omega(f;h_m,a,b)\le M''h_m^\alpha.$$

Now, for every $0 < h \leq 1$, there exists an integer $m > 0$ such that $h_m < h \leq h_{m-1}$.

Thus we can imply that,

$$\begin{aligned}\omega\left(f;h,a,b\right) &\leq \omega\left(f;h_{m-1},a,b\right)\\ &\leq M''h_{m-1}^{\alpha}\\ &= M''K^{\alpha}h_m^{\alpha}\\ &< M'''h^{\alpha}.\end{aligned}$$

Since,

$$\begin{aligned}\left|f\left(t\right)-2f\left(t+\delta\right)+f\left(t+2\delta\right)\right| &\leq \left|f\left(t\right)-f\left(t+\delta\right)\right|+\left|f\left(t+\delta\right)-f\left(t+2\delta\right)\right|\\ &\leq 2\omega\left(f;h,a,b\right)\\ &\leq 2M'''h^{\alpha}.\end{aligned}$$

We get,

$$\omega_2\left(f;h,a,b\right) \leq Ch^{\alpha}.$$

Thus, $f \in Lip^{*}(\alpha;a,b)$. With this, we have shown that the conclusion is true. ■

Acknowledgments

The first author would like to express gratitude to Delhi Technological University for their support. Both the authors are also indebted to the editor and the anonymous reviewers for their constructive comments and suggestions, which led to an improvement in the original manuscript.

Data availability statement

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Authors' contributions

Kanita was involved in conceptualization, methodology and in the original draft preparation. Naokant Deo contributed to review editing, investigation and supervision. All authors have read and agreed to the final manuscript.

Consent for publication

Both the authors have given their consent to the final version that has been submitted.

Disclosure statement

There are no competing interests. The authors have no relevant competing interests to disclose, either financial or otherwise.

References

[1] Ismail M EH, May CP. On a family of approximation operators. J Math Anal Appl. 1978;63(2): 446–462. doi: 10.1016/0022-247X(78)90090-2
[2] Tyliba A, Eugeniusz W. On some class of exponential type operators. Comment Math. 2005;45(1):57–70.
[3] Herzog M. Semi-exponential operators. Symmetry. 2021;13(4):637. doi: 10.3390/sym13040637
[4] Abel U, Gupta V, Sisodia M. Some new semi-exponential operators. Rev Real Acad Ciencias Exactas, Físicas Nat Ser A Mat. 2022;116(2):87. doi: 10.1007/s13398-022-01228-2

[5] Aral A, Limmam ML, Ozsarac F. Approximation properties of Szász–Mirakyan–Kantorovich type operators. Math Methods Appl Sci. 2019;42(16):5233–5240. doi: 10.1002/mma.v42.16
[6] Deo N, Muhammad Aslam N, Mohammad Arif S. On approximation by a class of new Bernstein type operators. Appl Math Comput. 2008;201(1–2):604–612.
[7] Gupta V, Ali A, Firat O. On semi-exponential Gauss–Weierstrass operators. Appl Math Comput. 2022;12(5):111.
[8] Gupta V, Ruchi G. Convergence estimates for some composition operators. Constr Math Anal. 2024;7(2):69–76.
[9] Lipi KM, Deo N. General family of exponential operators. Filomat. 2020;34(12):4043–4060. doi: 10.2298/FIL2012043L
[10] Lipi KM, Deo N. On modification of certain exponential type operators preserving constant and e^{-x}. Bull Malays Math Sci Soc. 2021;44:3269–3284. doi: 10.1007/s40840-021-01100-3
[11] Mishra NS, Deo N. Kantorovich variant of Ismail–May operators. Iran J Sci Technol Trans A Sci. 2020;44(3):739–748. doi: 10.1007/s40995-020-00863-x
[12] Mishra Nav S, Naokant D. On the preservation of functions with exponential growth by modified Ismail–May operators. Math Methods Appl Sci. 2021;44(11):9012–9025.
[13] Ozsarac F, Acar T. Reconstruction of Baskakov operators preserving some exponential functions. Math Methods Appl Sci. 2019;42(16):5124–5132. doi: 10.1002/mma.v42.16
[14] Ozsarac F, Gupta V, Aral A. Approximation by some Baskakov–Kantorovich exponential-type operators. Bull Iran Math Soc. 2022;48(1):227–241. doi: 10.1007/s41980-020-00513-3
[15] Cai Q-B, Bo-Yong L, Guorong Z. Approximation properties of λ-Bernstein operators. J Inequal Appl. 2018;2018(61):1–11.
[16] Chen X, Tan J, Liu Z, et al. Approximation of functions by a new family of generalized Bernstein operators. J Math Anal Appl. 2017;450(1):244–261. doi: 10.1016/j.jmaa.2016.12.075

Check for updates

Concurrency and Computation: Practice and Experience

WILEY

RESEARCH ARTICLE

Building Scalable and Quantum Attack Resistant Authenticated Message Delivery System for Internet of Vehicles With Blockchain Consensus Mechanism

Rahul Singh[1] | Laxminarayan Das[1] | Dharminder Chaudhary[2]

[1]Department of Applied Mathematics, Delhi Technological University, Delhi, India | [2]Department of Computer Science & Engineering, Amrita School of Computing, Amrita Vishwa Vidyapeetham, Chennai, India

Correspondence: Dharminder Chaudhary (manndharminder999@gmail.com)

Received: 1 October 2024 | **Revised:** 27 February 2025 | **Accepted:** 20 March 2025

Keywords: aggregate signature | authentication | blockchain technology | internet of vehicles | privacy | security

ABSTRACT

The goal of intelligent transportation systems is becoming more and more realized with the use of the Internet of Vehicles (IoV) and the rapid advancement of processing and communication technologies. Nonetheless, a lot of Internet of Vehicles applications depend on a central processing and storage unit as well as wireless transmission mediators. This may result in exorbitant expenses and delays, as well as the disclosure of real data. We suggest the Vehicle-Based Quantum-Safe Blockchain Consensus (VBQBC) algorithm as a solution to these problems and an enhancement of the effectiveness of data storage, authentic processing, and data sharing in the Internet of Vehicles. The proposed VBQBC algorithm uses a consensus algorithm and blockchain technology to assure authentic communication between cars, overcoming the shortcomings and constraints of current state-of-the-art systems. This algorithm uses ring learning with errors, and short integer solution assumptions in the construction of aggregation signatures to provide authenticity in blockchain technology. This aggregation technique reduces the size of data required for verification and improves scalability. The algorithm also incorporates a quantum-safe authentication procedure as well as a key distribution and request process, which are demonstrated when vehicles move between different zones. This allows blockchain-based systems to maintain their security, scalability, and efficiency even in the face of future cryptographic problems. The aggregation of signatures size for the proposed framework to the number of signatures to be aggregated N varies between 2^{10} (63.48 kb) and 2^{20} (131.34 kb). For aggregation of 2^{10} signatures, the proposed framework has signature size 6.3 kb, and the aggregated size of signatures is 64 kb. In simulation findings, our suggested VBQBC algorithm outperformed previous techniques in terms of authentication delay, key processing time, attack detection rate, throughput, and packet loss rate.

1 | Introduction

The concept of aggregated signatures was first introduced by Boneh et al. [1] in 2003. The aggregated signatures they proposed merge multiple signatures tied to different messages into a single signature, enabling the authentication of several message-signature pairs belonging to different users in one process. This innovation minimizes the storage space needed for signatures, reduces the network bandwidth needed for data transmission, and decreases the number of signature verification processes required. Aggregated signature technology serves not only as a method for signature compression but also as a technique for batch processing of signatures, catering to the demands for rapid data authentication in today's big data landscape,

© 2025 John Wiley & Sons Ltd.

particularly in areas like the Internet of Things, cloud computing, and blockchain technology. The aggregation of signature offers numerous possible applications, such as signing contracts, establishing trust routes, distributing certificate authorities, and utilizing blockchain technology [2, 3]. Blockchain technology brings transformative advantages to the Internet of Vehicles (IoV) by addressing critical challenges in data security, communication, and trust management within the vehicular ecosystem.

The IoV is a network where vehicles, infrastructure, and other entities interact to enable autonomous driving, efficient traffic management, and advanced vehicular services. The IoV represents a significant advancement in vehicular technology, offering numerous benefits such as enhanced safety, optimized traffic management, and improved driving experiences. The widespread adoption of (IoV) has revolutionized Intelligent Transportation Systems (ITS) in the coming years by enabling V2V (vehicle to vehicle) communication, and V2I (vehicle to infrastructure) communication. It directly addresses the increasing demand for smarter and efficient intelligent transportation system [4–7]. With the rapid growth of smart vehicles, the volume of data exchange will increase exponentially. The vast amount of data generated by IoV systems poses significant data management challenges.

Blockchain eliminates the need for centralized servers, ensuring that vehicle-related data and services are distributed across a network of nodes [8]. This technology helps to coordinate vehicle movement in smart cities by securely sharing data about traffic flow and congestion. This technology also enhances IoV by providing a secure, decentralized, and trustworthy framework for data exchange and collaboration among the participants. But, with the advent of quantum computing algorithm, traditional cryptographic schemes like RSA or ECC-based aggregate signatures algorithms are vulnerable to quantum attacks [9]. Lattice-based cryptographic techniques and aggregate signature algorithms are being explored for quantum-resistant solutions. Post-quantum cryptography, such as lattice-based algorithms, can future-proof IoV systems against quantum computing threats. Therefore, we need a quantum-safe blockchain-based distributed, scalable, flexible, and adaptable architecture to get rid of this big data exchange, and its storage. The scalability concerns include maintaining low latency in communication, managing network congestion, and ensuring reliable data transmission across the network [10].

1.1 | Motivation

The IoV envisions a highly connected ecosystem where vehicles, infrastructure, pedestrians, and other entities interact seamlessly to enhance safety, efficiency, and user experience. However, the realization of IoV comes with significant challenges related to security, privacy, scalability, and trust. Blockchain technology provides a robust solution to these issues, serving as a foundation for secure and decentralized IoV communication. The motivation for integrating blockchain technology into IoV stems from its ability to address critical challenges in security, trust, and efficiency while enabling innovative use cases like autonomous payments, and decentralized data sharing. By leveraging blockchain, IoV can realize its full potential as a secure, scalable, and collaborative transportation ecosystem. Although, there are many blockchain based solutions for IoV, but they are based on number theoretic cryptosystems. Shor's algorithm invents a significant theoretical threat to existing cryptosystems based on integer factorization and discrete logarithm. While practical quantum computer is capable of running Shor's algorithm at a threatening scale do not yet exist, the rapid advancement in quantum computing necessitates proactive measures. Transitioning to quantum-resistant cryptography and preparing infrastructure for post-quantum security are crucial steps to mitigate the potential future risks posed by quantum computing advancements. Post-quantum cryptography, such as lattice-based algorithms, can future-proof IoV systems against quantum computing threats. Moreover, there are some lattice-based quantum-safe blockchain frameworks exist, but either they are not scalable or vulnerable to well-known attacks. Therefore, we need a quantum-safe blockchain-based distributed, scalable, flexible, and adaptable architecture to get rid of this big data exchange, and its storage. The scalability concerns include maintaining low latency in communication, managing network congestion, and ensuring reliable data transmission across the network.

1.2 | Contribution

In the proposed framework, we have proposed a quantum-safe blockchain-based distributed, scalable, flexible, and adaptable architecture to get rid of this big data exchange, and its storage. This framework also solves scalability concerns include maintaining low latency in communication, managing network congestion, and ensuring reliable data transmission across the network.

- The proposed algorithm utilizes blockchain technology to provide a secure authentication scheme with low latency. In the performance section, we have given details of parameters and comparison of the proposed framework with related protocols in terms of blocks, and read latency. The proposed algorithm possesses low read latency, it deploy aggregation of signatures, which are faster than other existing ones, results approximately 60.5% faster reading time for blocks.
- The proposed aggregation technique produces single compact signature, while ensuring authenticity of communication. This technique is proven to be secure in the model "existential unforgeability under chosen-message and chosen identity attack" means if M-LWE and M-SIS assumptions are hard then proposed aggregation of signature framework is secure.
- The proposed framework uses M-LWE and M-SIS assumptions those are more standard lattice based problems and increase both security and efficiency due to simple algebraic operations (additions and multiplications). The proposed framework avoids linear compression idea that is frequently used to existing versions.
- The concept is based on zone-based architecture, and aggregation of quantum safe signatures to provide better latency, attack detection rate, and enhanced authentication in ITS. Every authentication procedure between vehicles and controllers will only be permitted in the designated zones in

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70071 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

order to minimize overhead and the number of nodes. This greatly optimizes processing time when vehicles join or depart.

- Finally, the blockchain based implementation of the proposed aggregation of signatures illustrates its practical application in IoV environment.

2 | Related Work

To address data security and efficiency issues in the Internet of Vehicles (IoV), a novel blockchain-based consensus algorithm was proposed in [11]. This solution calls for the creation of a vehicle-centered consensus mechanism. It also introduces a strategy for managing vehicle migration between regions, with the goal of reducing the overhead associated with frequent re-authentication during regional transitions. However, the proposed scheme lacks a formal security proof. Additionally, as the speed of the moving car is higher, some issues appear connected with the stability of the wireless connection, the safety and the reliability of the verification and migration steps. Some of these limitations may greatly impact the general efficiency and reliability of the solution. Due to the high speed of vehicles, constant checking with various types of roadside units is unavoidable, resulting in significant computation and communication costs. These increased demands may cause congestion in interactions between vehicles and roadside units, reducing overall system performance and efficiency. On a similar basis, Roy et al. [12] established the blockchain decentralized storage system and created BACHP-IoV, a lightweight access control protocol based on ECDSA (elliptic curve based digital signature algorithm). The effectiveness of this protocol in the security domain was later confirmed using the Real or Random Model. However, the following study did not address the security of message transmission between vehicles, leaving an unresolved issue. Yang et al. [13] presented an authentication architecture for vehicles that is flexible to multi-domain scenarios and addresses the cross-domain authentication problem. With the use of blockchain technology, it became possible to share information across many domains of management with decentralized trust. To address the issue of anonymity and traceability at the same time, the framework built an authentication solution, a pseudonym allocation mechanism based on ECDLP. A new dual-chain architecture described in [14] is used to handle vehicle authentication and message exchange in the Internet of Vehicles. It utilizes two blockchains: Car registration uses the Registration Blockchain (RBC), whereas message management prefers the Message Blockchain (MBC). The system prevents and mitigates fraud by combining the Proof of Authority (POA) and Proof of Participation (POP) consensus procedures. Nonetheless, the issue of performance enhancement for the dual-chain system was outside the purview of the study, allowing for a suggestion of an improvement. To address identity authentication difficulties in the Internet of Vehicles (IoV) using 6G technology, the scheme [15] introduced a multi-domain authentication solution with conditional privacy protection. The system splits the Blockchain-assisted multi-domain network into management and geographical domains, and it proposes an identity signature authentication approach that does not require bilinear pairing. Furthermore, Tao et al. [16] proposed a dynamic method for message verification that employs the Chinese Remainder Theorem to ensure message integrity. They also presented a blockchain approach for protecting automotive data on a large scale while maintaining privacy, as well as a smart contract enabling car traceability during postmortem investigations. Thus, the scheme's great efficiency was obtained by utilizing elliptic curve point multiplication to explain the computation burden of the RSUs. However, it failed to defend against quantum algorithm concerns, indicating that the system is not secure in the future.

But all of the above discussed frameworks are based on number-theoretic assumption like RSA/DLP which are not secure against quantum attacks. A post-quantum blockchain is a forward-looking approach to securing blockchain ecosystems against future quantum threats. It blends advanced cryptographic techniques from lattice theory, and secure hashing to create a resilient and scalable platform for decentralized applications. In [17], authors have proposed a lattice based approach to enforce anonymous authentication for VANET with vulnerable security and privacy prerequisites. Again a novel DiLSe method for safe distribution of data in IoV was introduced in [18], and it uses lattice theory for secure distribution, deep learning for content management, fault-tolerant data replication, access control, and Blockchain supported key exchange. In the same year 2023, Kaur et al. [19] introduced an authenticated architecture for transmitting sensitive data and a significant transformation using blockchain, but they lacks in blockchain testing. In the view of edge-based computing, Liu et al. [20] proposed a group-based blockchain framework. This idea is capable to provide decentralized identity based on secret information and a dynamic representation for IoV system, but it lacks the concept of shared driving. In the year 2021, Li et al. [21] proposed a post quantum secure lattice-Based aggregation of signatures, but communication cost is high, and public/private key size is bigger, consensus algorithm is missing. In [22], Wen et al. developed the first upper lattice ring signature scheme with forced revocation and proved it to possess the theoretical properties such as anonymity, unforgeability and revocability via random oracle model. This method has been used with IoV to design the first PQC secure against chosen plain text attack that solely relies on ring signatures. However, they face several challenges including high costs of computing and communication in the network environment as well as the stringent mobility and security constraints of vehicular networks. In 2024, Prajapat et al. [23] proposed a new NTRU lattice-based aggregation of signatures, but the size of signature increases during aggregation. Chen et al. [24] proposed protocol incorporates identity authentication functionality and effectively integrates the cross-domain authentication characteristics of the Blockchain. Initially, the Trusted Authority (TA) uploads the registration information of vehicles onto the Blockchain to verify the authenticity of a vehicle's identity by examining the data stored in the Blockchain. This is an authentication protocol, but it does not supports aggregation and batch verification process. Most of these methods fail to provide a decentralized nature of blockchain in authentication [19, 20, 25]. Private blockchains [19, 20, 25] have been used to solve these issues of privacy. Therefore, especially for vehicular communication, the proposed framework helps to resolve vehicular network authentication related challenges, specifically for use in autonomous vehicle-to-vehicle communication. This paper contains analysis on post-quantum blockchain [21, 23, 26–29] and attempted to implement them in authenticated vehicle-to-vehicle communication.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70071 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

3 | Preliminaries

Learning with Errors (LWE) is a fundamental assumption in cryptography, and it forms the basis for many post-quantum cryptographic encryption/signature schemes, including some signature that could potentially be used for aggregate signatures in blockchain technology.

The Learning with Errors assumption is to find a secret vector $\mathbf{s}$ from noisy linear equations. Specifically, for a given matrix $\mathbf{A}$ and a vector $\mathbf{c}$, the aim is to finding $\mathbf{s}$ such that:

$$\mathbf{A} \cdot \mathbf{s} \equiv \mathbf{c} (\mathrm{mod}\, q)$$

where, $\mathbf{A}$ is a rectangular $m \times n$ matrix with entries from a finite ring $\mathbb{Z}_q$ (integers modulo prime q), $\mathbf{s}$ is an n-dimensional tuple, $\mathbf{c}$ is an m-dimensional tuple, and q is a prime modulus parameter.

- *Noise Addition*: The tuple $\mathbf{c}$ includes a noisy tuple $\mathbf{e}$ such that $\mathbf{c} = \mathbf{A} \cdot \mathbf{s} + \mathbf{e}$, where $\mathbf{e}$ is a small random vector. The noise $\mathbf{e}$ makes the assumption computationally hard to solve without knowledge of $\mathbf{s}$.
- *Hardness Assumption*: The security of LWE-based cryptosystem depends upon the assumption that it is computationally hard to distinguish $\mathbf{c}$ from uniformly random vectors in $\mathbb{Z}_q^m$, even with known matrix $\mathbf{A}$.
- *Parameter Choices*: The security of LWE depends on choosing suitable parameters m, n, and q. In general, n is dimension of $\mathbf{s}$ chosen to be larger than m is number of equations, and $q = 2^n$ is a large prime.

3.1 | Learning With Errors (LWE) Assumption

The Learning With Errors (LWE) assumption is a lattice based fundamental assumption in cryptography.

Definition 1. Let us consider parameters $n, m, q \in \mathbb{N}$, then $\mathrm{LWE}_{n,m,q}$ assumption is given n samples (a_i, b_i) where $a_i \leftarrow \mathbb{Z}_q^n$ and $b_i = \langle a_i, s \rangle + e_i \mod q$, with $s \leftarrow \mathbb{Z}_q^n$ and $e_i \leftarrow \chi$, distinguish whether $b_i \leftarrow \mathbb{Z}_q$ or $b_i \leftarrow \mathbb{Z}_q$ with non-negligible advantage.

Here, $\langle \cdot, \cdot \rangle$ denotes the general inner product, χ represents a noise distribution over $\mathbb{Z}_q$.

3.2 | Short Integer Solution (SIS) Assumption

The Short Integer Solution (SIS) assumption is another lattice based fundamental problem in cryptography.

Definition 2. Let us consider parameters $n, m, q \in \mathbb{N}$, then $\mathrm{SIS}_{n,m,q}$ assumption is given a uniformly random matrix $A \leftarrow \mathbb{Z}_q^{n \times m}$, find a non-zero vector $\mathbf{s} \in \mathbb{Z}^m$ satisfying:

- $A \cdot \mathbf{s} \equiv \mathbf{0} \pmod q$,
- $\| \mathbf{s} \| \leq \beta$,

where, $\| \cdot \|$ denotes the ℓ_2-norm and β is a parameter.

3.3 | Module Learning With Errors (M-LWE) Problem

The Module Learning With Errors (M-LWE) assumption is another variant of the Learning With Errors (LWE) assumption, which is a lattice based foundational problem in modern cryptography. Here's a description of what M-LWE involves:

Definition 3. Let us consider $k, \ell, \eta \in \mathbb{N}$, then $\mathrm{M\text{-}LWE}_{k,\ell,\eta}$ is defined for a random matrix $A \leftarrow \mathbb{R}_q^{k \times \ell}$ and a vector $t \in \mathbb{R}_q^k$, distinguish whether:

1. $t \leftarrow \mathbb{R}_q^k$, or
2. $t = \left[A | I_\ell\right] \cdot \mathbf{s}$,

where, $\mathbf{s} \leftarrow S_{\ell+k}^{\eta}$.

The distribution $S_{\ell+k}^{\eta}$ denotes vectors of length $\ell + k$ with entries from $\mathbb{Z}_q$ and a small error term from discrete Gaussian.

3.4 | Module Short Integer Solution (M-SIS) Problem

The Module Short Integer Solution (M-SIS) assumption is another fundamental problem in the field of lattice-based cryptography, closely related to the Learning With Errors (LWE) assumption. Here's a detailed description of what the M-SIS problem entails:

Definition 4. Let us consider parameters $k, \ell, b \in \mathbb{N}$, then $\mathrm{M\text{-}SIS}_{k,\ell,b}$ is defined for uniform randomly chosen matrix $A \leftarrow \mathbb{R}_q^{k \times \ell}$, find a non-zero small $\mathbf{s} \in \mathbb{R}_q^{k+\ell}$ satisfying:

- $\| \mathbf{s} \|_2 \leq b$,
- $\left[A | I_\ell\right] \cdot \mathbf{s} = \mathbf{0} \in \mathbb{R}_q^k$.

Here, I_ℓ represents the $\ell \times \ell$ identity matrix.

3.5 | Preliminaries on Aggregate Signatures

Aggregation of signatures involves the process to combine multiple signatures into a single compact signature. Here, we have discussed some mathematical notations, concepts and definitions related to aggregation of signatures:

1. *Digital Signature Algorithm*: A digital signature algorithm is a basic cryptographic method used to provide authenticity, integrity, and non-repudiation of messages or documents transferred on public channel. Here, we have discussed that how a digital signature algorithm typically works and some of its key components:
   - *Key Generation Process*: $\mathrm{Gen}(1^\kappa) \rightarrow (s\kappa, p\kappa)$, where $s\kappa$, and $p\kappa$ are private and public keys respectively.
   - *Signature Algorithm*: $\mathrm{Sign}_{s\kappa}(m) \rightarrow \sigma$, where σ is the signature on the message m using private key $s\kappa$.
   - *Verification Algorithm*: $\mathrm{Verify}_{p\kappa}(m, \sigma) \rightarrow \{\mathrm{True}, \mathrm{False}\}$, does the verification of σ for message m using public key $p\kappa$.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70071 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ALGORITHM 1 | Aggregation of signatures.

0: **procedure** AGGREGATE-SIGNATURE-GENERATION (Messages $\{m_1, m_2, \ldots, m_n\}$, Keys $\{sk_1, sk_2, \ldots, sk_n\}$)
1: Initializes $\sigma_{agg} \leftarrow$ Identity of the signature group
1: **for** $i = 1 \rightarrow n$ **do**
2: Computes $\sigma_i \leftarrow \text{Sign}(m_i, sk_i)$
3: $\sigma_{aggs} \leftarrow \sigma_{aggs} \oplus \sigma_i$
3: **end for**
4: **return** σ_{aggs}
4: **end procedure** = 0

2. *Aggregate Signature Algorithm*: An aggregate signature Algorithm 1 extends the above definitions to aggregate multiple signatures:
   - *Aggregate key generation*: $\text{Aggr} - \text{Gen}(1^\kappa, n) \rightarrow (s\kappa_{\text{aggr}}, p\kappa_{\text{aggr}})$, where n is the number of signers.
   - *Aggregate signing algorithm*: $\text{AggrSign}_{s\kappa_{\text{aggr}}}(m_1, \ldots, m_n) \rightarrow \sigma_{\text{aggr}}$, generates an aggregate signature σ_{aggr} for messages $m_1, \ldots, m_n$ using private key $s\kappa_{\text{aggr}}$.
   - *Aggregate verification algorithm*: $\text{AggrVerify}_{p\kappa_{\text{aggr}}}(m_1, \ldots, m_n, \sigma_{\text{aggr}}) \rightarrow \{\text{True}, \text{False}\}$, verifies σ_{aggr} for messages $m_1, \ldots, m_n$ using aggregate public key $p\kappa_{\text{aggr}}$.

3.6 | Network Model

The proposed blockchain solution comprises the following types of entities.

1. *Peer-to-Peer Communication*: The blockchain based networking operates on a peer-to-peer basis, providing decentralized functionalities. Any node can participate in the network and establish connections with other nodes.
2. *Nodes*: The nodes represent users within the network. Depending on the application, these nodes can vary in their roles.
3. *Consensus*: To reach an agreement on transactions, a consensus mechanism is executed. This ensures that decisions are transparent and safeguards against compromised activities, such as decision manipulation and Byzantine faults.

3.7 | Working of Aggregate Signature in Blockchain

An aggregation of signature algorithm is a cryptographic technique that allows multiple signers to combine multiple signatures into a single one. This aggregated signature is verified against a combined set of messages and public keys, rather than verifying each single signature individually. Aggregation of signatures in blockchain is executing multiple transaction signatures into a single compact aggregated signature. This execution process generally follows these steps:

1. *Signature Collection*: Transactions with individual signatures are collected from the network.
2. *Aggregation Algorithm*: An aggregation algorithm computes a consolidated signature (often using techniques like BLS or Schnorr aggregation) for the entire block of transactions.
3. *Verification*: The aggregated signature is included in the block header. Nodes can verify the entire block's validity using this single aggregated signature, rather than verifying each transaction individually.

4 | Proposed System Architecture

As shown in Figure 1, the proposed architecture divides vehicles into multiple divisions/zones. Each of division equipped with a designated controller to observe and supervises all of the activities occurring, including vehicle authentication, movement, and vehicle-to-vehicle internal communication with a distributed ledger powered by blockchain technology. While the vehicles are subject to restrictions on their energy consumption and computational power, the controllers in this scenario, while the Roadside Units (RSUs), are considered to have no such restrictions. The divisions have been designed to facilitate communication between various controllers, RSU-to-vehicle, vehicle-to-vehicle, and RSU-to-blockchain. Furthermore, the blockchain-enabled P2P (peer to peer) networking allows the divisions to safely interact with one another. Given the fast speeds at which vehicles are traveling, scalability has a critical issue in the authentication process [8]. As seen in Figure 1, our suggested architecture employs a structure of hierarchy to overcome this. A controller is in charge of overseeing a coverage division, which consists of several levels of modules. Identified as R, controllers engage with the public blockchain and exchange messages with one another inside a network.

Figure 2 illustrates the blockchain architecture for the vehicular network. Authentication functions, secret keys, and blockchain transaction blocks are all part of the system. All vehicle information is included in transactions, such as the vehicle ID, MAC address, zone number, and route and direction metadata. Public and private keys are created by R for each vehicle at the time of registration. The private key can be used to sign/authenticate transmitted data between vehicle in their respective divisions. The private key can also be used to migrate across divisions and communicate with other in those divisions.

4.1 | Blockchain-Based Controller

In particular, Figure 3 illustrates the internal workings of the controller and how it is connected to the cars. The ledger, denoted by L, maintains and archives the data about the cars. The duties assigned to the controller can be taken from the mechanism described in [2]. Since the vehicles can register in any zone for safe communication and migration while in motion, we gave the public blockchain some thought. As a result, it aids in protecting the vehicles' privacy [3]. The vehicles are able to connect to the network and distribute either private or public keys for the public blockchain.

The Hash functions/signatures, private keys, and public blockchain transactions are examples of the blocks that make up

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70071 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

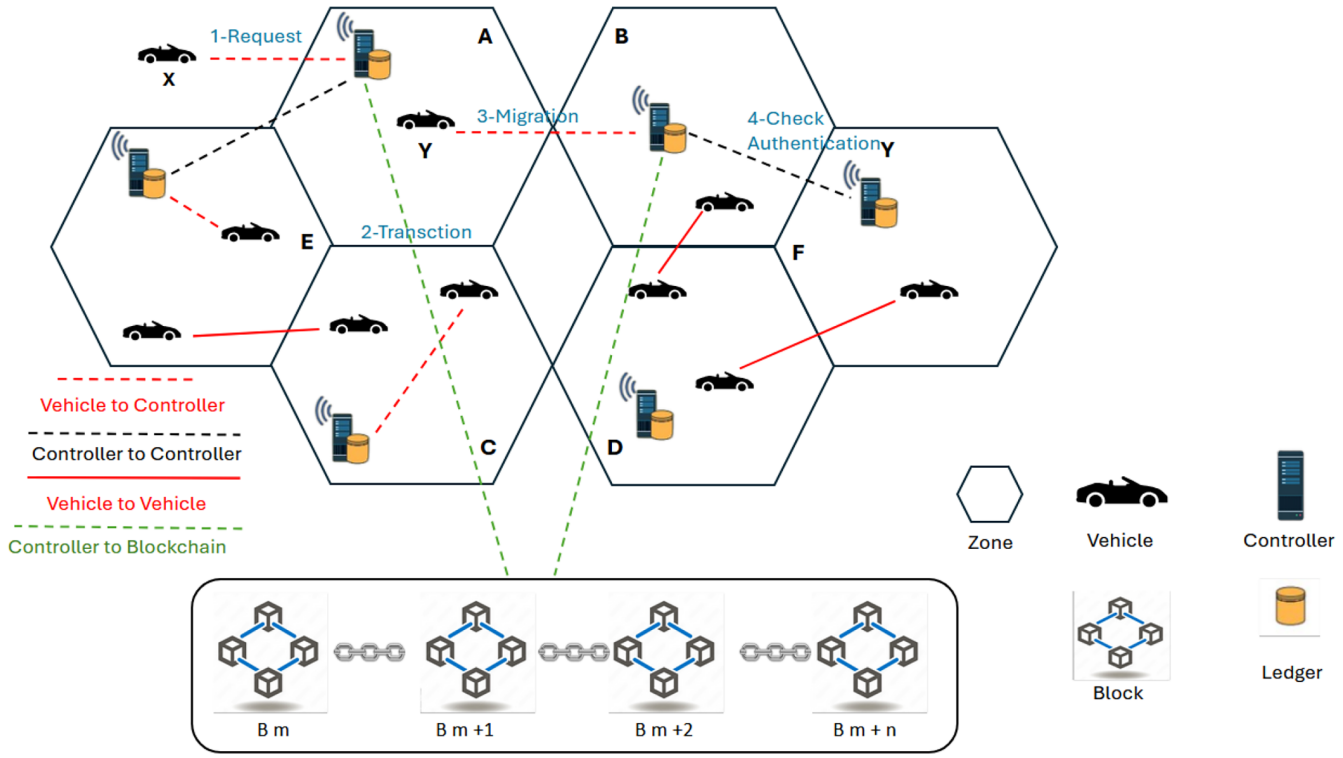


FIGURE 1 | The proposed zone-based framework for the IoV using blockchain.

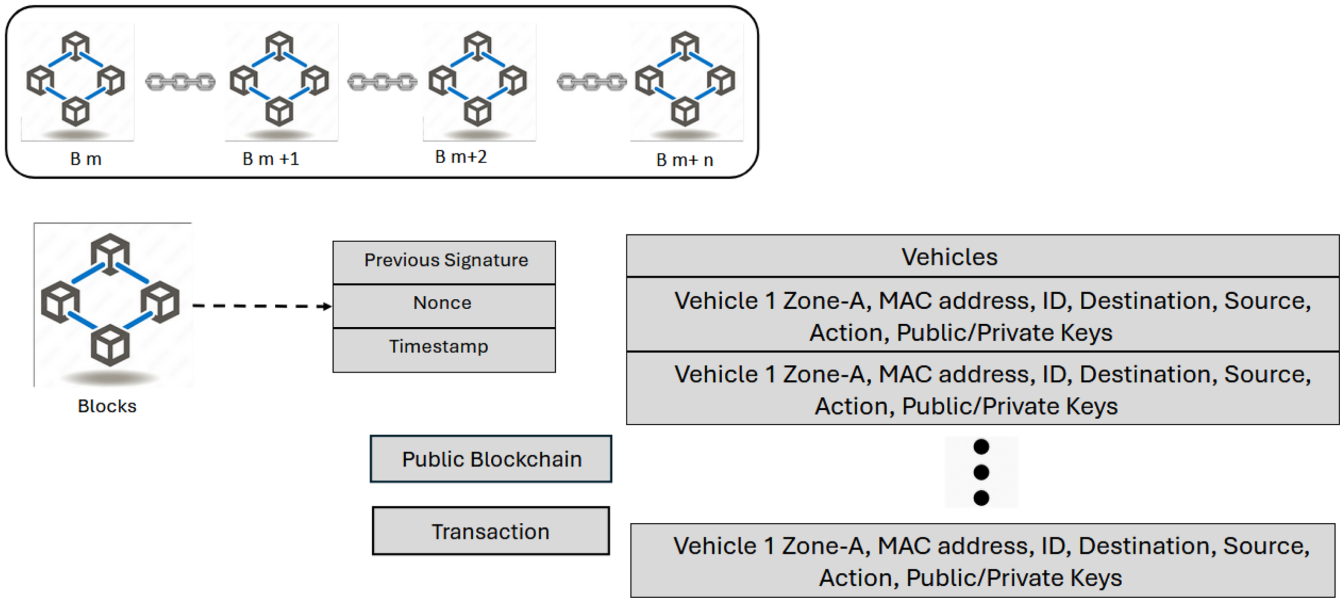


FIGURE 2 | Blockchain components and public blockchain-based communication in IoV.

the blockchain. Vehicle information, such as a car's distinct ID and MAC address obtained while driving, are contained in the transactions. On the other hand, the information is stored in a vector that the system generates and represents as V_{info}. To access their respective divisions, the vehicles can receive the public or private keys by using V_{info} and a legitimate ID. Additionally, they can use V_{info} to migrate to other vehicles the information. It is mentioned that the consensus process, a low latency, secure proof of work protocol, is responsible for managing the controllers. Due to constraints and shortcomings in the present literature, including centralized systems and problems with real-world voting, we provide a vehicle-based quantum secure blockchain consensus (VBQBC) method while maintaining the current consensus algorithm. The transactions in each division are endorsed and checked by our suggested VBQBC. Every authentication procedure between vehicles and controllers will only be permitted in the designated zones in order to minimize overhead and the number of nodes (Algorithm 2).

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70071 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

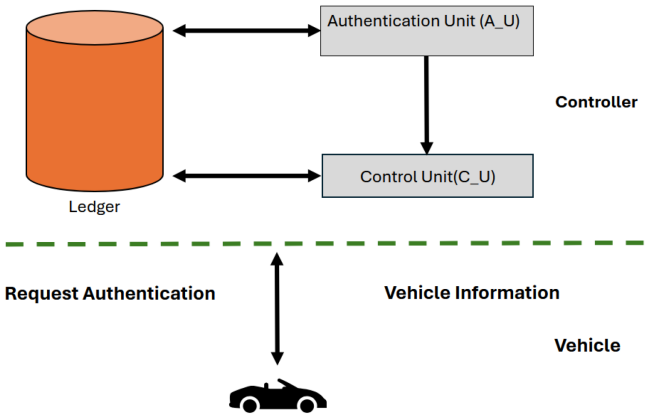


FIGURE 3 | The working flow between controller and vehicle in blockchain-based IoV.

ALGORITHM 2 | Vehicle-based quantum-safe blockchain consensus (VBQBC) algorithm.

1: **Input**: CON (.) = Division/Zone Data (Block-Header, Block-Version, Timestamps, Transactions)
2: Generate Agents Group for Mining
3: **For each** CON (.) **do**
4: Send information vehicles
5: Computes transactions
6: Compute functions
7: Sign block header
8: **while** Counts_Division **do**
9: Mine the block
10: Verification of header
11: Signature = endorses || nonce; endorse contains header, pre-header, timestamps, transactions
12: R = Sign($Sign$||Nonce)
13: Coordinates among Divisions
14: Extracts Nonce = getNonce($Sign$)
15: Transactions in Information
16: nonce++
17: **end while**
18: Blocks (Zone(j) = Information) Mining process successful
19: Share information of Division = 0

4.2 | Secure and Authentication Process

We take into account scattered divisions/zones among the network's automobiles using blockchain technology in the safe authentication procedure. Each division is responsible for managing the blockchain process and safeguarding the cars' data transmission while they are operating. Nevertheless, there is a situation in which cars can move between zones without having to authenticate again, saving time and computing resources. As can be seen from Figure 1, in order to encrypt the data, the vehicles need to register with the controller in each division. Each vehicle is given unique keys and authenticated data in this manner. Subsequently, the controller simultaneously sends the authenticated data to all other controllers and the blockchain. Accordingly, each vehicle receives the public and private keys from the controller R. Additionally, information regarding the adversary's vehicle might be sent by the controller. By sharing this data via a shared ledger, the blockchain enables controllers in other divisions to deny the rogue cars' incoming requests.

ALGORITHM 3 | Vehicle migrations between divisions.

1: Vehicle registers with controllers and Divisions
2: Performs necessary authentications
3: Sends public/private keys
4: Use Signing
5: $Vehicle_j$ = Receive(Sign)
6: Joins the Division
7: **if** Vehicle gets registered with the controller **then**
8: Authentication = Successful
9: Computes the mobility-rate
10: **else**
11: Computes the migration-rate
12: **end if**
13: **if** Movement is requested **then**
14: Joins the Division
15: **else**
16: Block
17: **end if**
18: **while** Authentication==Successful **do**
19: **if** Mobility/Migration == Successful **then**
20: Authenticates in the Division
21: $Controller_j$: Information
22: Updates the Division's information
23: Migrates from current Division to requested Division
24: **else**
25: $Controller_j$: Blockchain updates
26: **end if**
27: **end while**
28: **while** Authentication == Not Successful **do**
29: Next Division CON(.) receives the data
30: Next Division CON(.) verifies the data
31: **end while** = 0

4.3 | Key Distribution and Requesting Process

Unique vehicle attributes, including position, identity, direction, and public and private keys, are identified by the authentication unit (A_U). These properties are generated by the controller unit (C_U). For security purposes, the data is added explicitly to the decentralized ledger through C_U. Every piece of data pertaining to the cars is registered and uploaded to the blockchain, where it is stored in a ledger. Other controllers must also have access to this data for estimate purposes. The private key, κ_{pri} can be used for signature on transactions between two vehicles, but the public key, κ_{pub}, is available to both vehicle and controller. The paired keys K_{pri} and K_{pub} have a mathematical relationship. Only the matching K_{pri} can be used to sign the message that can be verified using K_{pub}.

4.4 | Migration and Authentication Process

The procedure for migrating across divisions is intended for secure communication and necessitates the least amount of latency during the process of authentication, as seen in Algorithm 3. The system controller shortens the time needed

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70071 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

for key transfer between zones by facilitating key management inside divisions. The main goal is to optimize the transfer method by effectively managing the controllers' keys inside the divisions. By regulating the amount of time needed for key execution in vehicle authorization using the blockchain and public/private keys, the goal is to reduce the amount of third-party contact during transactions. The suggested architecture notifies the controller of the zone next to it if one of its controllers malfunctions. As a result, the cellular structure makes sure that in the zones around the malfunctioning controller, the fewest neighboring controllers are situated. The controller in charge of handling the incident will be chosen from among these nearby zones whose vehicle count is lowest.

1. *Key-Generation*: Key-generation algorithms are fundamental components of cryptographic systems, responsible for generating keys used in encryption/decryption, and signature processes. This algorithm ensures the creation of secure, random, and unique keys that are resistant to cryptographic attacks. This takes a random matrix $\overline{A}$, $s \xleftarrow{\$} S^{\eta}_{\ell+k}$, $\tau \coloneqq \overline{A}s \bmod \rho$, and generates key-pair containing $s\kappa \coloneqq s$, and $p\kappa \coloneqq \tau$ respectively.

2. *Block Generation using Sequential Aggregate Signature*: This phase considers a challenge space C, and a message space $\mathcal{M}' = \{0,1\}^{l}$, and a collision resistant hashing $H : \{0,1\}^{*} \to C$. This algorithm starts with $i = 1$, $\varrho_0 = \emptyset$ and $\sigma_0 = (0,0)$. Furthers the algorithm **SeqSigns** (Sequential Aggregate Signature) takes input $s\kappa_i, m_i, \varrho_{i-1}, \sigma_{i-1}$, and computes $\left(\tilde{\vartheta}_{i-1}, z_1, \ldots, z_{i-1}\right) \coloneqq \sigma_{i-1}$, $s_i \coloneqq s\kappa_i$, $\tau_i \coloneqq \overline{A}s_i \bmod \rho$, $\varrho_i \coloneqq \varrho_{i-1} \| (\tau_i, m_i)$. This samples $y_i \leftarrow D_{\ell+k}$, $\vartheta_i \coloneqq \overline{A}y_i \bmod \rho$, $\tilde{\vartheta}_i \coloneqq \tilde{\vartheta}_{i-1} \oplus \vartheta_i$, $c_i \coloneqq H\left(\tilde{\vartheta}_i, \varrho_i, z_{i-1}, vid_i\right)$, $z_i \coloneqq c_i \cdot s_i + y_i$, $z_i \coloneqq \text{RejSamps}\left(z_i, c_i \cdot s_i\right)$, and returns $\sigma_i \coloneqq \left(\tilde{\vartheta}_i, z_1, \ldots, z_i\right)$ see Algorithm 4.

3. *Block Verification using Sequential Aggregate Signature*: Sequential aggregate signatures are a cryptographic construct that allows multiple signatures from different signers on different messages to be combined into a single signature. This single signature can then be verified to authenticate all the messages. The primary advantage is efficiency, as it reduces the size of the overall signature and the complexity of verification. The algorithm **SeqVerify** (Sequential Aggregate Signature Verification) takes input $< \varrho_N, \sigma_N >$, and $(\tau_1, m_1) \| \ldots \| (\tau_N, m_N) \coloneqq \varrho_N$, and $\left(\tilde{\vartheta}_N, z_1, \ldots, z_N\right) \coloneqq \sigma_N$. Furthers the algorithm computes $z_0 \coloneqq 1$ if $\exists\ i$ satisfying τ_i has no inverse of any coefficient. For $i = N, \ldots, 1$, and $\| z_i \|_2 > B$, returns 0, else $\varrho_i \coloneqq (\tau_1, m_1) \| \ldots \| (\tau_i, m_i)$, $c_i \coloneqq H\left(\tilde{\vartheta}_i, \varrho_i, z_{i-1}, vid_i\right)$, $\vartheta_i \coloneqq \overline{A}z_i - c_i\tau_i \bmod \rho$, $\tilde{\vartheta}_{i-1} \coloneqq \tilde{\vartheta}_i \oplus \vartheta_i$, and if $\tilde{\vartheta}_1 = \vartheta_1$, then returns 1 and accepts the signature (see Algorithm 4).

5 | Security Analysis

Blockchain aggregation of signatures is a cryptographic technique that enables multiple signatures to be combined into a single compact signature. This can provide significant efficiency gains in blockchain technology based systems by reducing the amount of data that needs to be stored and transmitted. Here's an

ALGORITHM 4 | Description of the proposed sequential aggregation of signatures.

Cipher-Space: $C \coloneqq \left\{ c \in R : \| c \|_2 = \sqrt{\kappa} \right\}$
Message-Space: $\mathcal{M}' = \{0,1\}^{l}$
Collision Resistant Hashing: $H : \{0,1\}^{*} \to C$
Starts: $i = 1$, $\varrho_0 = \emptyset$ and $\sigma_0 = (0,0)$
Setup: as in Algorithm 1
Gens $\overline{A}$
$\tau \coloneqq 0$
while τ without inverse of coefficients **do**
  $s \xleftarrow{\$} S^{\eta}_{\ell+k}$
  $\tau \coloneqq \overline{A}s \bmod \rho$
end while
$s\kappa \coloneqq s$
$p\kappa \coloneqq \tau$
$(s\kappa, p\kappa)$
$\left(\tilde{\vartheta}_{i-1}, z_1, \ldots, z_{i-1}\right) \coloneqq \sigma_{i-1}$
$s_i \coloneqq s\kappa_i$
$\tau_i \coloneqq \overline{A}s_i \bmod \rho$
$\varrho_i \coloneqq \varrho_{i-1} \| (\tau_i, m_i)$
$z_i \coloneqq \perp$
while $z_i = \perp$ **do**
  $y_i \leftarrow D_{\ell+k}$
  $\vartheta_i \coloneqq \overline{A}y_i \bmod \rho$
  $\tilde{\vartheta}_i \coloneqq \tilde{\vartheta}_{i-1} \oplus \vartheta_i$
  $c_i \coloneqq H\left(\tilde{\vartheta}_i, \varrho_i, z_{i-1}, vid_i\right)$
  $z_i \coloneqq c_i \cdot s_i + y_i$
  $z_i \coloneqq \text{RejSamps}\left(z_i, c_i \cdot s_i\right)$
end while
$\sigma_i \coloneqq \left(\tilde{\vartheta}_i, z_1, \ldots, z_i\right)$

return σ_i
SeqVerifies ϱ_N, σ_N
$(\tau_1, m_1) \| \ldots \| (\tau_N, m_N) \coloneqq \varrho_N$
$\left(\tilde{\vartheta}_N, z_1, \ldots, z_N\right) \coloneqq \sigma_N$
$z_0 \coloneqq 1$
if ∃ i satisfying τ_i where element not invertible **then**

  return 0
end if
for $i = N, \ldots, 1$ **do**
  if $\| z_i \|_2 > B$ **then**

    return 0
  end if
  $\varrho_i \coloneqq (\tau_1, m_1) \| \ldots \| (\tau_i, m_i)$
  $c_i \coloneqq H\left(\tilde{\vartheta}_i, \varrho_i, z_{i-1}, vid_i\right)$
  $\vartheta_i \coloneqq \overline{A}z_i - c_i\tau_i \bmod \rho$
  $\tilde{\vartheta}_{i-1} \coloneqq \tilde{\vartheta}_i \oplus \vartheta_i$
end for
if $\tilde{\vartheta}_1 = \vartheta_1$ **then**

  return 1
else
  return 0
end if
=0

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70071 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

analysis of the security attributes, and mathematical correctness of blockchain aggregate signatures:

Theorem 1. *The scheme aggregation of signatures* = `Setup, Gens, SeqSigns, SeqVerifies` *as described in Algorithm 3 is mathematical correct.*

Proof. In this theorem, we have shown that, if an ith signature $\sigma_i = (\tilde{\vartheta}_i, z_1, \ldots, z_i)$ for $1 \le i < N$ is correct, then $(i+1)$ indexed signature $\sigma_{i+1} = (\tilde{\vartheta}_{i+1}, z_1, \ldots, z_{i+1})$ is also mathematical correct. The value z_{i+1} is computed correctly by signer, then it generates $\overline{A} \cdot z_{i+1} - \tau_{i+1} \cdot c_{i+1} = \vartheta_{i+1}$. Therefore, $\tilde{\vartheta}_i$ is recovered with the help of $\tilde{\vartheta}_{i+1} - \vartheta_{i+1}$ and the $\sigma_i = (\tilde{\vartheta}_i, z_1, \ldots, z_i)$ is used to do verification by the mathematical induction.

Now, let us assume $i = 1$, then it is used to generate $\| z_1 \|_\infty \le B$ and $\overline{A} \cdot z_1 = \tau_1 \cdot c_1 + \tilde{\vartheta}_1$ due to the linearity of mathematical algebra over matrices multiplication over $\mathbb{R}_q$. □

Theorem 2. *Let* $\kappa, \ell, n, \rho, \eta, \gamma, l \in \mathbb{N}$, *where* $n = 2^i$, ρ *is a prime*, p_{inv} *is the probability that* $\overline{A}$ *can find minimum one inverse of coefficient over* $\mathbb{R}_\rho$, *where* s *is uniform random sample from* $S_\eta^{\ell+\kappa}$ *and* $\overline{A} = [A | I_\kappa]$ *with* A *uniform random sample from* $\mathbb{R}_\rho^{\kappa \times \ell}$, *respectively. If the single signature is Unforgeable, then the proposed sequential aggregation in Algorithm 4, is secure.*

Proof. In cryptography, security proof often uses a series of games to show that an adversary's success probability in breaking a scheme can be reduced to an adversary breaking an underlying assumption, typically a hard problem. Here's a security proof for the proposed aggregation of signatures using a series of games G_0, G_1, G_2, G_3, and G_4, respectively. □

First, the broad concepts of the reduction $\mathcal{B}$ are outlined. You may find the full explanation of $\mathcal{B}$ in Algorithm 5. For an unforgeable game (UF-CMA game), OSeqSigns and H (resp. OSigns and H') stand for the random oracle as well as the signing oracle. After obtaining the challenge public key τ^* and the public parameter A, $\mathcal{B}$ verifies that $\tau^* \in \mathbb{R}_\rho^k$ comprises a minimum of one invertible element. In that case, $\mathcal{B}$ transfers (A, τ^*) to $\mathcal{A}$. In response to an inquiry, OSeqSigns asks OSigns to sign a uniformly selected m and programs H to produce the value c that is returned by an external random oracle H'. Since a forgery filed by $\mathcal{A}$ may subsequently utilize the same m_i, we are unable to simply pass m_i to OSigns in this case. Afterward, $\mathcal{B}$ loses the UF-CMA game since their submission of the forgery with respect to m_i will be invalid. The simulation of responses to H queries is the core element of the reduction. The forgery tuple's key list L_N includes (t_i, m_i) such that $\tau_i = \tau^*$. Afterward, in order for $(m, (\vartheta_i z_i))$ to be considered a legitimate forgery within the UF-CMA game, any arbitrary message m and the accompanying ϑ_i must be transmitted to H' by $\mathcal{B}$. It is noteworthy that z_i^{-1} is used in this extraction procedure each time $(\tilde{\vartheta}_i, \varrho_i, z_i^{-1})$ is queried from H. $\mathcal{B}$ can extract $\vartheta_i = \tilde{\vartheta}_i - \tilde{\vartheta}_{i-1}$ by using z_i^{-1} as an intuitive look-up key to retrieve the previously aggregated $\tilde{\vartheta}_{i-1}$. To be more precise, we build a number of hybrid games from the initial FH-UF-CMA game in order to arrive at the one that is employed in the final reduction $\mathcal{B}$. The probability that $G_i(\mathcal{A})$ stops when output is one, and it is indicated by $\Pr[G_i(\mathcal{A})]$.

ALGORITHM 5 | Reduction to UF-CMA VBQBC.

The UF-CMA H'.
The UF-CMA OSigns.
Consider $\tilde{\vartheta}_0 = 0$, A query to H with $\tau_1, \ldots, \tau_i$.
Process $B^{\text{OSigns},H'}(A, \vec{t}^*)$
  $Q \coloneqq \emptyset$; $M \coloneqq \emptyset$
  if t^* not having inverse **then**
    badinverse $\coloneqq$ true
  $(\sigma_N, \varrho_N) \leftarrow A^{\text{OSeqSigns},H}(A, t^*)$
  $(\tau_1, m_1) \| \ldots \| (\tau_N, m_N) \coloneqq \varrho_N$
  $(\tilde{\vartheta}_N, z_1, \ldots, z_N) \coloneqq \sigma_N$
  $I \coloneqq \{i \in [N] : \tau_i = t^* \wedge (m_i, \varrho_{i-1}) \notin Q\}$
  if SeqVerifies$(\sigma_N, \varrho_N) = 1 \wedge |I| \neq 0$ **then**
    Derives $\tilde{\vartheta}_{N-1}, \ldots, \tilde{\vartheta}_1$ **as in** SeqVerifies
    if $\exists i \in [N]$ **satisfying** $\text{ZT}[\tilde{\vartheta}_i, \varrho_i, z_{i-1}] = \perp$ **then**
      badord $\coloneqq$ true
    if $\exists i^* \in I$ **satisfying** $m \coloneqq \text{MT}[\tilde{\vartheta}_{i^*}, \varrho_{i^*}, z_{i^*-1}] \neq \perp$ **then**
      $\vartheta_{i^*} \coloneqq \tilde{\vartheta}_{i^*} - \tilde{\vartheta}_{i^*-1}$
      returns $(m, \vartheta_{i^*}, z_{i^*})$
  else
    badmcol $\coloneqq$ true
Process OSeqSigns$(m_i, \varrho_{i-1}, \sigma_{i-1})$
  $Q \coloneqq Q \cup \{(m_i, \varrho_{i-1})\}$
  $(\tilde{\vartheta}_{i-1}, z_1, \ldots, z_{i-1}) \coloneqq \sigma_{i-1}$
  $m \xleftarrow{\$} M$
  if $\exists X$ **satisfying** $\text{MT}[X] = m$ **then**
    $\text{MT}[X] \coloneqq \perp$
  $M \coloneqq M \cup \{m\}$, $(\vartheta, z) \leftarrow \text{OSigns}(m)$
  $c \coloneqq H'(\vartheta, m, t^*)$, $\varrho_i \coloneqq \varrho_{i-1} \| (t^*, m_i)$
  if $\tilde{\vartheta}_i \coloneqq \tilde{\vartheta}_{i-1} + \vartheta \bmod q$, $z_i \coloneqq z$
  if $\text{HT}[\tilde{\vartheta}_i, \varrho_i, z_{i-1}] \neq \perp$ **then**
    baducols $\coloneqq$ true
  else
    $\text{HT}[\tilde{\vartheta}_i, \varrho_i, z_{i-1}] \coloneqq c$
  $\sigma_i \coloneqq (\tilde{\vartheta}_i, z_1, \ldots, z_i)$
  returns σ_i
Process $H(\tilde{\vartheta}_i, \varrho_i, z_{i-1})$
  if $\text{HT}[\tilde{\vartheta}_i, \varrho_i, z_{i-1}] \neq \perp$ **then**
    returns $\text{HT}[\tilde{\vartheta}_i, \varrho_i, z_{i-1}]$
  $(\tau_1, m_1) \| \ldots \| (\tau_i, m_i) \coloneqq \varrho_i$
  if $i = 1$ **or** $\exists X \coloneqq (\tilde{\vartheta}_{i-1}, \varrho_{i-1}, z_{i-2})$ **satisfying** $\text{HT}[X] \neq \perp \wedge \text{ZT}[X] = A\overline{z}_{i-1} \bmod q$ **then**, $\vartheta_i \coloneqq \tilde{\vartheta}_i - \tilde{\vartheta}_{i-1} \bmod q$
  if $\tau_i = t^*$ **then**, $m \xleftarrow{\$} M$, $c_i \coloneqq H'(\vartheta_i, m, t^*)$
    if $m \notin M$ **then**
      $\text{MT}[\tilde{\vartheta}_i, \varrho_i, z_{i-1}] \coloneqq m$
    else
      $c_i \xleftarrow{\$} \text{CH}$
  if $\exists X' \coloneqq (\tilde{\vartheta}\prime_i, \varrho_i, z'_{i-1})$ **satisfying** $\text{ZT}[X'] = \vartheta_i + c_i\tau_i \bmod q$ **then**
    badzcols $\coloneqq$ true
    $\text{ZT}[\tilde{\vartheta}_i, \varrho_i, z_{i-1}] \coloneqq \vartheta_i + c_i\tau_i \bmod q$
  else
    $c_i \xleftarrow{\$} \text{CH}$
  $\text{HT}[\tilde{\vartheta}_i, \varrho_i, z_{i-1}] \coloneqq c_i$, **returns** c_i

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70071 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

- G_0: This game starts with initializes a blank table HT. When HT's entry is non-empty, the random oracle HH returns $HT[X]$ upon receiving the query with input X; otherwise, it returns c, assigns $HT[X] \coloneqq c$, and choose a random sample $c \in Ch$ with $\Pr\left[G_0(A)\right] = \mathrm{Adv}^{\mathrm{UF}}_{\mathrm{VBQBC}}(A)$.
- G_1: The game is pretty similar to G_0, only the difference is *OSeqSigns* implements Oracle list $HT\left[\tilde{\vartheta}_i, \varrho_i, \mathbf{z}_{i-1}\right] \coloneqq c_i$ if rejection sample phase successful, and samples uniform $c_i \in CH$ rather than making a call $c_i = H\left(\tilde{\vartheta}_i, \varrho_i, \mathbf{z}_{i-1}, vid_i\right)$ after $\tilde{\vartheta}_i$ is computed. The game gets terminated by declaring $\mathrm{bad}_{ucol} = \mathrm{true}$ if $HT\left[\tilde{\vartheta}_i, \varrho_i, \mathbf{z}_{i-1}\right]$ has been set. Thus, $|\Pr\left[G_0(A)\right] - \Pr\left[G_1(A)\right]| \leq \Pr\left[\mathrm{bad}_{ucols}\right]$.

TABLE 1 | Experiments, summaries, and metrics.

Parameters	Values
Transactions	[1; 200; 200]
Executed tests	60
Confidence-interval	82%
Price	3
Average TPS (transactions per day)	3.01296296
Average mining per transaction	1.244 s
Time for single hashing	1.42 s

TABLE 2 | Modification among three groups of blockchain.

Parameter	Transaction-1	Transaction-2	Transaction-3
Block-interval	10 min	10 min	1 min
Block-size	1 MB	1 MB	1 MB
Difficulty-adjustment-interval	2012 blocks	62 blocks	Adaptable-limit
Past-blocks	2012	62	Difficulty-adjustment

- G_2: The game is pretty similar to G_1, only the difference is simulating the responses to random oracle queries $H\left(\tilde{\vartheta}_i, \varrho_i, \mathbf{z}_{i-1}\right)$. Initialize a new empty ZT table. If $i = 1$ or $X \coloneqq \left(\tilde{\vartheta}_{i-1}, \varrho_{i-1}, \mathbf{z}_{i-2}\right)$ exists and $ZT[X] = \overline{\mathbf{A}}\mathbf{z}_{i-1} \bmod q$, extracts $\vartheta_i \coloneqq \tilde{\vartheta}_i - \tilde{\vartheta}_{i-1}$, samples random uniform $c_i \in Ch$, and assigns $ZT\left[\tilde{\vartheta}_i, \varrho_i, \mathbf{z}_{i-1}\right] \coloneqq \vartheta_i + c_i\mathbf{t}_i$. If an entry $X' \coloneqq \left(\tilde{\vartheta}_{i-1}, \varrho_{i-1}, \mathbf{z}_{i-2}\right)$ already exists and $ZT\left[X'\right] = \vartheta_i + c_i\mathbf{t}_i$, the game terminates by setting $\mathrm{bad}_{zcol} = \mathrm{true}$. Hence, $|\Pr\left[G_1(A)\right] - \Pr\left[G_2(A)\right]| \leq \Pr\left[\mathrm{bad}_{zcols}\right]$.
- G_3: The game is pretty similar to G_2, only the difference is that *OSeqSigns* and HH initializes random empty table MT and an empty set M. Upon receiving a query, *OSeqSigns* samples random message $m \in M$ and assigns m to M. After extracting ϑ_i as previously queried, H samples $m \in M$ and rejects $\mathrm{bad}_{mcols} = \mathrm{true}$ if $m \in M$ whenever it receives a query with input $\left(\tilde{\vartheta}_i, \varrho_i, \mathbf{z}_{i-1}\right)$, then $|\Pr\left[G_2(A)\right] - \Pr\left[G_3(A)\right]| \leq \Pr\left[\mathrm{bad}_{mcols}\right]$.
- G_4: The game is pretty similar to G_3, only the difference is that when $\mathcal{A}$ outputs a legal signature-history pair $\left(\varrho_N, \left(\tilde{\vartheta}_N, \mathbf{z}_1, \ldots, \mathbf{z}_N\right)\right)$, the game validates the ZT entries as follows. Let $\tilde{\vartheta}_{N-1}, \ldots, \tilde{\vartheta}_1$ be the values obtained during the *SeqVerifies* run. The game terminates by setting $\mathrm{bad}_{ords} = \mathrm{true}$ if for random chosen $i \in [N]$, the value for $ZT\left[\tilde{\vartheta}_i, \varrho_i, \mathbf{z}_{i-1}\right]$ is undefined, then $|\Pr\left[G_3(A)\right] - \Pr\left[G_4(A)\right]| \leq \Pr\left[\mathrm{bad}_{ords}\right]$.

Assuming that A wins G_4, the reduction B mentioned in Algorithm 4 is obtained in this way. Upon receiving an *OSeqSigns* query, B queries the UF-CMA game's *OSigns* with a uniformly chosen $m \in M$. B then retrieves ϑ_i and $\mathbf{z}_i$, programs HT using c_i returned by the external $H'\left(\tilde{\vartheta}_i, \tau^*, m\right)$. Therefore, $\vartheta_i^* = \tilde{\vartheta}_i^* - \tilde{\vartheta}_{i^*-1}$ is extracted by querying $H\left(\tilde{\vartheta}_i^*, \varrho_i^*, \mathbf{z}_{i^*-1}\right)$. As long as $\mathrm{bad}_{zcols} = \mathrm{bad}_{ords} = \mathrm{false}$, then queries are processed in right order.

```
# Define a simple transaction structure
class Transaction:
    def __init__(self, sender, recipient, amount, message, public_key, signature):
        self.sender = sender
        self.recipient = recipient
        self.amount = amount
        self.message = message
        self.public_key = public_key
        self.signature = signature

# Generate key pairs
def generate_key_pair():
    private_key = bls.PrivateKey()
    public_key = private_key.get_public_key()
    return private_key, public_key
```

FIGURE 4 | Screenshot of executed code for aggregate signature in blockchain network.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70071 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

6 | Performance

This section illustrates a simplified implementation in Python using aggregation of signatures. Note that this is a conceptual example and might require additional cryptographic libraries and support with a blockchain framework. To integrate with blockchain, first updates the node software to handle number of transactions with aggregation of signatures, and it implements logic to create, verify, and propagate aggregation of signatures transactions. It also ensures that the block generation process handles and validates blocks containing transactions with aggregation of signatures, and implements block validation logic that verifies aggregation signatures efficiently. It depends upon the consensus mechanism (Proof of Stake, Proof of Work) integrating aggregation of signatures verification process to ensure the authenticity in the blockchain. Aggregation of signatures is a cryptographic innovation that can significantly enhance the efficiency and performance of blockchain systems. By enabling multiple signatures to be combined into a single, shorter signature, aggregate signatures can reduce the storage and bandwidth requirements of blockchain networks. This analysis covers various performance aspects of aggregate signature-based blockchains, including efficiency, scalability, authenticity, throughput, and practical implementation. Additionally, this section contains an analysis on results of existing algorithm, and the proposed algorithm. The setup details are available in the Tables 1 and 2. The Figures 4–6 contains screenshots for compiled code.

We have performed an analysis based on number of vehicles versus authentication process delay in microseconds. The authentication process delay is measured in terms of time needed

```
# Sign a message
def sign_message(private_key, message):
    message_hash = sha256(message.encode()).digest()
    signature = bls.Sign(private_key, message_hash)
    return signature

# Verify a single signature
def verify_signature(public_key, message, signature):
    message_hash = sha256(message.encode()).digest()
    return bls.Verify(public_key, message_hash, signature)

# Aggregate multiple signatures
def aggregate_signatures(signatures):
    return bls.Aggregate(signatures)
```

FIGURE 5 | Screenshot of executed code for aggregate signature in blockchain network.

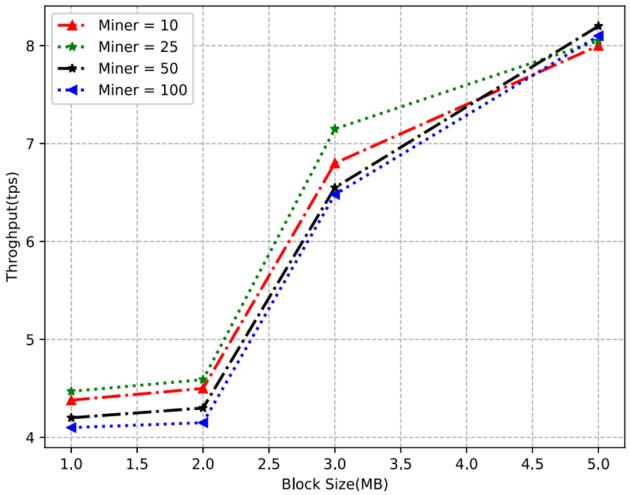


FIGURE 7 | Read latency versus number of blocks comparison among related, and the proposed frameworks.

```
sig1 = sign_message(priv_key1, message1)
sig2 = sign_message(priv_key2, message2)

tx1 = Transaction("Alice", "Bob", 10, message1, pub_key1, sig1)
tx2 = Transaction("Charlie", "Dave", 20, message2, pub_key2, sig2)

# Aggregate signatures
aggregate_sig = aggregate_signatures([sig1, sig2])

# Verify individual signatures
assert verify_signature(pub_key1, message1, sig1)
assert verify_signature(pub_key2, message2, sig2)

# Verify aggregate signature
assert verify_aggregate_signature([pub_key1, pub_key2], [message1, message2],

print("All signatures verified successfully!")
```

FIGURE 6 | Screenshot of executed code for aggregate signature in blockchain network.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70071 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

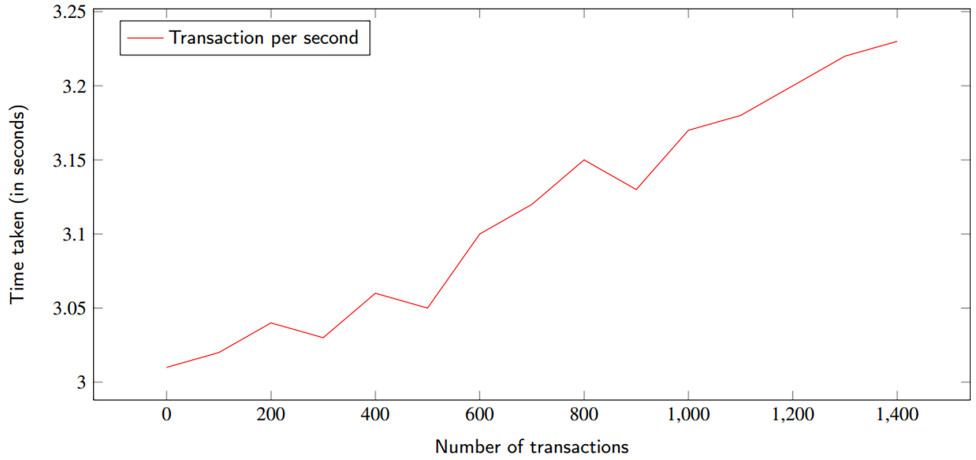


FIGURE 8 | Transaction latency versus number of blocks comparison among related, and the proposed frameworks.

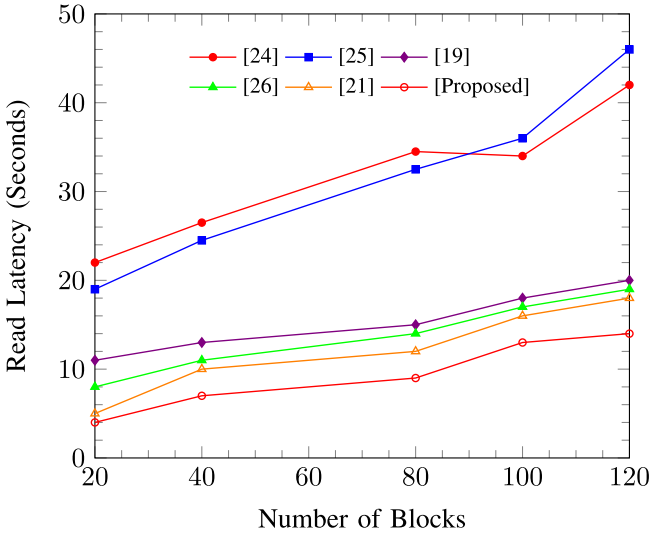


FIGURE 9 | An illustration of time taken and number of transaction for the proposed framework.

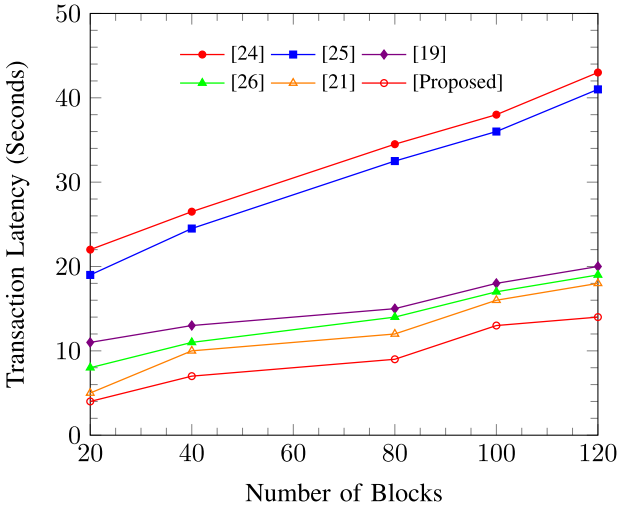


FIGURE 10 | Throughput versus block-size comparison for proposed framework.

for authentication when a vehicle leave a division, and it joins a new division. This process needs to authenticate vehicle again while moving changing zones. In contrast, the proposed VBQBC algorithm eliminates re-authentication process of vehicles, and the consensus mechanism uses aggregation of signatures for the authenticity of transactions, and generating valid blocks to be added in the blockchain. This section also contains analysis on blockchain parameters, likewise block sizes and miners count, on the proposed algorithm bandwidths and throughput. The number of executed transactions varies with block sizes, and it is responsible for the throughput consumed by the proposed algorithm. We have done a performance analysis based on read latency versus number of blocks comparison among related, and the proposed frameworks (see Figure 7). Read latency refers to the time it takes to retrieve data from a storage system or database after a request is initiated. It is a critical performance metric in systems such as distributed databases, memory, and storage networks. In this figure, X-axis denotes number of blocks and Y-axis denotes read latency in seconds. Furthers the Figure 8 illustrates transaction latency versus number of blocks comparison among related, and the proposed frameworks. Transaction latency is the time elapsed from when a transaction is initiated to when it is confirmed and becomes part of the blockchain. In this figure, the X-axis denotes number of blocks whereas Y-axis denotes transaction latency in seconds. Both figures ensures the efficiency of the proposed framework in terms of read and transaction latency.

TABLE 3 | Comparison of the suggested frameworks' computing costs against similar ones.

Designs	Key-generations	Signatures	Verification
Kiktenko et al. [26]	$2\tau_p + 2\tau_* + 2n\tau_l$	$2\tau_p + 2n\tau_* + 4\tau_h + 2n\tau_a$	$n\tau_p + 4n\tau_* + 2n\tau_h + n\tau_a$
Li et al. [27]	$4\tau_p + 4\tau_* + 2n\tau_l$	$4\tau_p + 2n\tau_* + 2\tau_h + 4\tau_a$	$n\tau_p + 4n\tau_* + 2n\tau_h + n\tau_a$
Li et al. [21]	$4\tau_p + 4\tau_* + n\tau_l$	$2\tau_p + 2n\tau_* + 4\tau_h + 2n\tau_a$	$n\tau_p + 4n\tau_* + n\tau_h + 2n\tau_a$
Saha et al. [28]	$2\tau_p + 3\tau_* + \tau_l$	$2\tau_p + 4\tau_* + 2\tau_h + \tau_a + n\tau_l$	$\tau_p + 2\tau_h + (n+2)\tau_l$
Prajapat et al. [23]	$\tau_p + 3\tau_*$	$2\tau_p + 4\tau_* + 2\tau_h + \tau_a + n\tau_l$	$\tau_p + 2\tau_h + (n+2)\tau_l$
Proposed	$\tau_p + 3\tau_*$	$2\tau_p + 2\tau_* + 2\tau_h + \tau_a + n\tau_l$	$\tau_p + 2\tau_h + n\tau_l$

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70071 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

In the Figure 7, we have compared the proposed algorithm with other relevant algorithms in terms of blocks (denoted on X-axis) versus read latency in seconds (denoted on Y-axis). The proposed algorithm possesses low read latency, it deploy aggregation of signatures, which are faster than other existing ones, results approximately 60.5% faster reading time for blocks. In the Figure 8, we have compared the proposed algorithm with other relevant algorithms in terms of blocks (denoted on X-axis) versus transaction latency in seconds (denoted on Y-axis). Figure 9 illustrates the number of transactions occurring per second. The X-axis represents the increasing number of open transactions, while the Y-axis shows the time taken per second as the number of transactions grows. The graph indicates a noticeable decrease in time after reaching a certain threshold of transactions. Furthers, Figure 10 illustrates the analysis with respect to throughput in transactions per second denoted on Y-axis versus block sizes increase from one megabyte to 5 megabytes. This experiment is performed with varying number of miners 10, 25, 50, and 100, respectively. As the number of miners increases, then consensus execution gets expedited among number of controllers. The Table 3 contains a comparison of the suggested frameworks' computing costs against similar ones in terms of computational costs.

This section contains six algorithms for implementation. The Algorithm 6 is denotes the process for adding a new authority to the contract, Algorithm 7 plays the role of verifier for existence of authority in the contract, Algorithm 8 executes the process of taking consent from a vehicle for data sharing, Algorithm 9 verifies consents, Algorithm 10 permits to have an access vehicle data, and Algorithm 11 validates transactions and generates new block to be added to the blockchain.

ALGORITHM 6 | Adds provider.

1: **Inputs**: Authority address _provAddress, Authority-Name _name, Authority-Role _role
2: **Outputs**: None
3: **New-Authority** ← **Vehicles-Authority**(_name, _role)
4: **Authority**[_provAddress] ← **New-Authority**

ALGORITHM 7 | Checks-authority-existence.

1: **Inputs**: Authority-Address _provAdd
2: **Outputs**: **Successful** if Authority exists, **Unsuccessful** Else
3: **if** authority[_provAdds].role is non empty **then**
4: **return Successful**
5: **else**
6: **return Unsuccessful**
7: **end if** = 0

ALGORITHM 8 | Consents.

1: **Inputs**: authority-address _provAdds, Vehicle-Address _vehicleAddress
2: **Outputs**: None
3: **Vehicle-Consents**[_vehicleAddress][_provAdds] ← **true**
=0

ALGORITHM 9 | Check-vehicle-consents.

1: **Input**: Authority-address _provAdds, vehicle-address _truckAddress
2: **Output**: **True** if truck consent exists, **False** otherwise
3: **if** vehicle-Consents[_vehicle-Address][_provAdds] is **Successful then**
4: **return Successful**
5: **else**
6: **return Unsuccessful**
7: **end if** = 0

ALGORITHM 10 | Access-vehicle-data.

1: **Inputs**: Authority-address _provAdds, Vehicle-address _vehicle-Address
2: **Outputs**: Access-Vehicle-Data
3: **if** vehicleConsents[_vehicleAddress][msg.sender] is **Successful then**
4: Access-Vehicle-Data
5: **else**
6: **Outputs Error**: Vehicle consent missing for accessing data
7: **end if** = 0

ALGORITHM 11 | Proof-of-work mining.

1: **Inputs**: Previous-Block-Hashing *previousHashes*, Data block *data*
2: **Outputs**: Nonce *nonce*, Current-Hashing *currentHash*
3: *nonce* ← 0
4: *difficulty* ← 5 {Adjusting difficulty needed}
5: *prefixes* ← new strings of *difficulty* zeros
6: **while** successful **do**
7: *hashInputs* ← SPHINCS+(abi.encodePacked (*previousHashing*, *data*, *nonce*))
8: *hash-Bytes* ← bytes(*hashingInput*)
9: *validPrefixes* ← successful
10: **for** $i \leftarrow 0$ **to** *difficulty* − 1 **do**
11: **if** *hashing-Bytes[i]* ≠ *prefix[i]* **then**
12: *validPrefixes* ← Unsuccessful
13: **breaks**
14: **end if**
15: **end if**
16: **if** *validPrefixes* **then**
17: *currentHashing* ← *hInputs*
18: **breaks**
19: **end if**
20: *nonce* ← *nonce* + 1
21: **end while** = 0

7 | Conclusion

In this paper, we have presented a Vehicle-Based Quantum Secure Blockchain Consensus (VBQBC) method to address the security and effectiveness issues raised by the massive amount of data in the Internet of Vehicles (IoV). Our system is decentralized, which improves authentication while reducing the drawbacks of centralization, such as high computation costs

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70071 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and message delay. Authentication is an important part of the Internet of Vehicles, and the VBQBC algorithm uses blockchain technology to provide authenticated communication between vehicles. Implementation of aggregate signatures in blockchain technology offers several advantages, and this technique helps in improving scalability, efficiency, and security. Aggregation of signatures allow multiple signatures to be combined into a single compact signature that leads to reduced size of transactions and blocks on the blockchain, leading to lower storage and bandwidth requirements. Simulation findings show that our VBQBC method surpasses current cutting-edge algorithms in terms of authentication latency, key processing time, and attack detection rate. Using post-quantum aggregation of signatures in blockchain technology represents a forward-thinking approach to security, particularly in the era of quantum computing.

Ethics Statement

This article does not contain any studies with animals performed by any of the authors. We have used ChatGPT/Grammarily for the improvement of English only. The technology is used under human oversight and control, and authors have carefully review and edit the results.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

References

1. D. Boneh, C. Gentry, B. Lynn, and H. Shacham, "Aggregate and Verifiably Encrypted Signatures From Bilinear Maps," in *Advances in Cryptology—EUROCRYPT 2003: International Conference on the Theory and Applications of Cryptographic Techniques, Warsaw, Poland, May 4–8, 2003 Proceedings 22* (Springer, 2003), 416–432.

2. D. X. Li, Y. Lu, and L. Li, "Embedding Blockchain Technology Into IOT for Security: A Survey," *IEEE Internet of Things Journal* 8, no. 13 (2021): 10452–10473.

3. K. Yu, L. Tan, M. Aloqaily, H. Yang, and Y. Jararweh, "Blockchain-Enhanced Data Sharing With Traceable and Direct Revocation in IIOT," *IEEE Transactions on Industrial Informatics* 17, no. 11 (2021): 7669–7678.

4. B. Ji, X. Zhang, S. Mumtaz, et al., "Survey on the Internet of Vehicles: Network Architectures and Applications," *IEEE Communications Standards Magazine* 4, no. 1 (2020): 34–41.

5. M. Waqas, Y. Niu, Y. Li, et al., "A Comprehensive Survey on Mobility-Aware d2d Communications: Principles, Practice and Challenges," *IEEE Communications Surveys & Tutorials* 22, no. 3 (2019): 1863–1886.

6. D. Chaudhary, P. Santhi, M. S. P. Durgarao, A. Padmavathi, M. M. Hassan, and B. F. Alkhamees, "Module Lattice Based Post Quantum Secure Blockchain Empowered Authentication Framework for Autonomous Truck Platooning," *IEEE Access* 1 (2024): 1–13.

7. D. K. Yadav, H. Sahu, D. Chaudhary, and C.-C. Lee, "Module Lattice Based Post Quantum Secure Blockchain Empowered Vehicle to Vehicle Communication in the Internet of Vehicles," *Computers and Electrical Engineering* 117 (2024): 109245.

8. R. Jabbar, N. Fetais, M. Kharbeche, M. Krichen, K. Barkaoui, and M. Shinoy, "Blockchain for the Internet of Vehicles: How to Use Blockchain to Secure Vehicle-to-Everything (v2x) Communication and Payment?," *IEEE Sensors Journal* 21, no. 14 (2021): 15807–15823.

9. W. Peter, "Shor. Introduction to Quantum Algorithms," in *Proceedings of Symposia in Applied Mathematics. Introduction to Quantum Algorithms*, vol. 58 (American Mathematical Society, 2002), 143–160.

10. T. Limbasiya, D. Das, and S. K. Das, "Mcomiov: Secure and Energy-Efficient Message Communication Protocols for Internet of Vehicles," *IEEE/ACM Transactions on Networking* 29, no. 3 (2021): 1349–1361.

11. S. Tu, H. Yu, A. Badshah, M. Waqas, Z. Halim, and I. Ahmad, "Secure Internet of Vehicles (IOV) With Decentralized Consensus Blockchain Mechanism," *IEEE Transactions on Vehicular Technology* 72, no. 9 (2023): 11227–11236.

12. S. Roy, S. Nandi, R. Maheshwari, S. Shetty, A. K. Das, and P. Lorenz, "Blockchain-Based Efficient Access Control With Handover Policy in IOV-Enabled Intelligent Transportation System," *IEEE Transactions on Vehicular Technology* 1 (2023): 1–15.

13. Y. Yang, L. Wei, J. Wu, C. Long, and B. Li, "A Blockchain-Based Multidomain Authentication Scheme for Conditional Privacy Preserving in Vehicular Ad-Hoc Network," *IEEE Internet of Things Journal* 9, no. 11 (2021): 8078–8090.

14. R. Tandon, A. Verma, and P. K. Gupta, "D-Blac: A Dual Blockchain-Based Decentralized Architecture for Authentication and Communication in Vanet," *Expert Systems with Applications* 237 (2024): 121461.

15. G. Cheng, J. Huang, Y. Wang, et al., "Conditional Privacy-Preserving Multi-Domain Authentication and Pseudonym Management for 6g-Enabled IOV," *IEEE Transactions on Information Forensics and Security* (2023): 1–17.

16. T. Qi, H. Ding, T. Jiang, and X. Cui, "B-Dspa: A Blockchain-Based Dynamically Scalable Privacy-Preserving Authentication Scheme in Vehicular Ad Hoc Networks," *IEEE Internet of Things Journal* 11, no. 1 (2023): 1385–1397.

17. H. Liu, Y. Sun, Y. Xu, R. Xu, and Z. Wei, "A Secure Lattice-Based Anonymous Authentication Scheme for Vanets," *Journal of the Chinese Institute of Engineers* 42, no. 1 (2019): 66–73.

18. A. Gulati, G. S. Aujla, R. Chaudhary, N. Kumar, M. S. Obaidat, and A. Benslimane, "Dilse: Lattice-Based Secure and Dependable Data Dissemination Scheme for Social Internet of Vehicles," *IEEE Transactions on Dependable and Secure Computing* 18, no. 6 (2019): 2520–2534.

19. K. Kaur, S. Garg, G. Kaddoum, F. Gagnon, and S. H. Ahmed, "Blockchain-Based Lightweight Authentication Mechanism for Vehicular Fog Infrastructure," in *2019 IEEE International Conference on Communications Workshops (ICC Workshops)* (IEEE, 2019), 1–6.

20. H. Liu, P. Zhang, G. Pu, T. Yang, S. Maharjan, and Y. Zhang, "Blockchain Empowered Cooperative Authentication With Data Traceability in Vehicular Edge Computing," *IEEE Transactions on Vehicular Technology* 69, no. 4 (2020): 4221–4232.

21. C. Li, Y. Tian, X. Chen, and J. Li, "An Efficient Anti-Quantum Lattice-Based Blind Signature for Blockchain-Enabled Systems," *Information Sciences* 546 (2021): 253–264.

22. J. Wen, L. Bai, Z. Yang, H. Zhang, H. Wang, and D. He, "Larrs: Lattice-Based Revocable Ring Signature and Its Application for Vanets," *IEEE Transactions on Vehicular Technology* 1 (2023): 1–14.

23. S. Prajapat, D. Gautam, P. Kumar, et al., "Secure Lattice-Based Aggregate Signature Scheme for Vehicular Ad Hoc Networks," *IEEE Transactions on Vehicular Technology* 73 (2024): 12370–12384.

24. C. Chen, L. Wang, and Q. Shi, "A Lattice-Based Certificateless Secure Data Transmission Scheme for Internet of Vehicles Based-Blockchain," *IEEE Transactions on Vehicular Technology* (2024): 1–16.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70071 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

25. Y. Yao, X. Chang, J. Mišić, V. B. Mišić, and L. Li, "Bla: Blockchain-Assisted Lightweight Anonymous Authentication for Distributed Vehicular Fog Services," *IEEE Internet of Things Journal* 6, no. 2 (2019): 3775–3784.

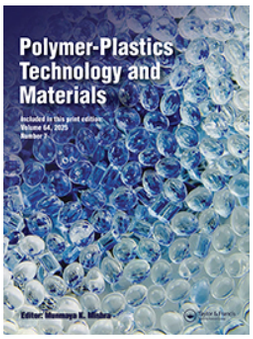
26. E. O. Kiktenko, N. O. Pozhar, M. N. Anufriev, et al., "Quantum-Secured Blockchain," *Quantum Science and Technology* 3, no. 3 (2018): 035004.

27. C.-Y. Li, X.-B. Chen, Y.-L. Chen, Y.-Y. Hou, and J. Li, "A New Lattice-Based Signature Scheme in Post-Quantum Blockchain Network," *IEEE Access* 7 (2018): 2026–2033.

28. R. Saha, G. Kumar, T. Devgun, et al., "A Blockchain Framework in Post-Quantum Decentralization," *IEEE Transactions on Services Computing* 16, no. 1 (2021): 1–12.

29. D. Chaudhary, M. S. P. Durgarao, D. Mishra, and S. Kumari, "A Module Lattice Based Construction of Post Quantum Secure Blockchain Architecture," *Transactions on Emerging Telecommunications Technologies* 35, no. 4 (2024): e4855.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70071 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



Taylor & Francis
Taylor & Francis Group

Polymer-Plastics Technology and Materials

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/lpte21

Carboxymethyl Guar Gum: Bridging the Gap Between Biopolymers and Advanced Functional Materials

Vratika Verma, Jagram Meena, Teena Saini, Sakshi Saini & Rajdeep Malik

To cite this article: Vratika Verma, Jagram Meena, Teena Saini, Sakshi Saini & Rajdeep Malik (01 Apr 2025): Carboxymethyl Guar Gum: Bridging the Gap Between Biopolymers and Advanced Functional Materials, Polymer-Plastics Technology and Materials, DOI: 10.1080/25740881.2025.2478453

To link to this article: https://doi.org/10.1080/25740881.2025.2478453

Published online: 01 Apr 2025.

Submit your article to this journal

Article views: 48

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=lpte21

POLYMER-PLASTICS TECHNOLOGY AND MATERIALS
https://doi.org/10.1080/25740881.2025.2478453

Taylor & Francis
Taylor & Francis Group

Check for updates

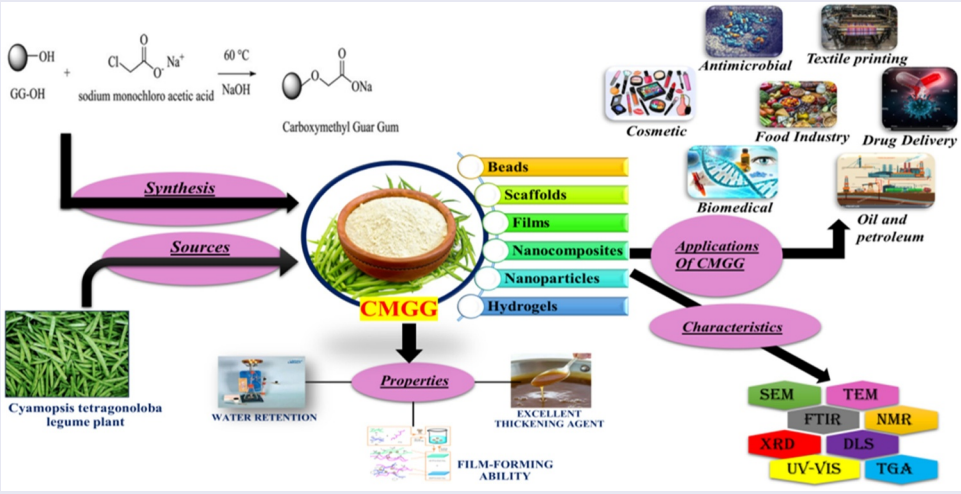
Carboxymethyl Guar Gum: Bridging the Gap Between Biopolymers and Advanced Functional Materials

Vratika Verma, Jagram Meena, Teena Saini, Sakshi Saini, and Rajdeep Malik

Department of Chemistry, Gurukul Kangri (Deemed to be) University, Haridwar, Uttarakhand, India

ABSTRACT

Synthetic polymers pose significant environmental and biomedical challenges, including non-biodegradability, microplastic accumulation, petrochemical reliance, and hazardous additive leaching. These issues drive concerns over ecological contamination and regulatory constraints. These challenges necessitate the development of eco-friendly alternatives, with biodegradable biopolymers emerging as a promising solution. Among them, Carboxymethyl Guar Gum (CMGG) stands out for its enhanced solubility, tunable rheology, and biodegradability, making it a viable candidate for sustainable material applications. This review explores CMGG's synthesis via carboxymethylation, structural modifications, and physicochemical characterizations using FTIR, UV-visible spectroscopy, XRD, TEM, and HRTEM. CMGG demonstrates versatile applications in biomedical engineering (drug delivery, wound healing), sustainable packaging (biodegradable films, stabilizers), and industrial formulations (hydrogels, rheology modifiers). Its potential in next-generation biomaterials highlights its role in advancing sustainable material science while addressing the limitations of conventional polymers.



KEYWORDS

Biodegradable biopolymers; biomedical applications; Carboxymethyl Guar Gum (CMGG); structural characterization etc; sustainable materials

HIGHLIGHTS

- **Shift to Biopolymers**: The paper discusses the limitations of traditional polymers, emphasizing the growing scientific interest in biopolymers as safer, more sustainable alternatives.
- **Biodegradable Polymers**: It highlights the emergence of biodegradable polymers as environmentally friendly materials, focusing on their biodegradability and biocompatibility for various applications.
- **Characterization Techniques**: The review details various characterization techniques for biopolymers, including FTIR, UV-visible microscopy, XRD, TEM, and HRTEM, which are essential for understanding their properties.
- **Focus on Carboxymethyl Guar Gum (CMGG)**: A specific highlight is the in-depth exploration of Carboxymethyl Guar Gum (CMGG), its synthesis through the reaction of Guar Gum with monochloroacetic acid, and its affordability and water solubility.

CONTACT Jagram Meena jagram.meena@gkv.ac.in Department of Chemistry, Gurukul Kangri (Deemed to be) University, Haridwar, Uttarakhand 249404, India

© 2025 Taylor & Francis Group, LLC

- **Applications Across Industries**: Showcases CMGG's diverse applications in sectors such as biomedical, textile, food, agriculture, and construction, highlighting recent research advancements and the material's extensive potential.
- **Recent Research Insights**: The paper provides recent research developments surrounding CMGG, emphasizing its biodegradability and growing potential across various fields.

1. Introduction

The pervasive accumulation of plastic waste has emerged as a pressing global environmental crisis. Synthetic polymers, which have become integral to modern life, exhibit remarkable durability and versatility but pose significant challenges due to their non-biodegradability. The escalating demand for plastics across industries – such as packaging, healthcare, textiles, and construction – has resulted in an alarming increase in plastic waste. Marine ecosystems suffer from plastic debris accumulation, while improper disposal through landfilling and incineration releases hazardous pollutants. With plastic consumption continuing to rise and recycling methods remaining inefficient, the urgent need for sustainable materials has become more critical than ever.

Polymers, derived from the Greek words *poly* (many) and *meros* (parts), are macromolecules composed of repeating monomeric units (Illustration 1). Their diverse mechanical, thermal, and chemical properties have revolutionized industries, ranging from consumer goods to high-performance materials. Historically, natural polymers such as cellulose, starch, and proteins have been widely utilized in textiles, food preservation, and biomedical applications. However, the 20th-century advent of synthetic polymers, including polyethylene, polypropylene, and polyvinyl chloride (PVC), enabled mass production of cost-effective, lightweight, and durable materials. Despite their widespread applicability, synthetic polymers contribute significantly to environmental pollution due to their resistance to biodegradation. Unlike natural polymers, which decompose through biological processes, synthetic plastics persist in ecosystems for centuries, exacerbating land and marine pollution.

The process termed polymerization is the formation of a polymer or large molecule from single units or monomers. In the process, there should be at least two reactive or bonding sites, generally known as functionality in a substance that has to act as a monomer in the polymer formation.[1] Table 1 shows some examples of polymers and their monomeric units, showcasing their applications and structure.

The reliance on petroleum-based polymers further accelerates resource depletion and greenhouse gas emissions, intensifying the environmental burden. Addressing these challenges necessitates a shift toward biopolymers, which offer eco-friendly degradation pathways and reduced dependence on fossil-based feedstocks. The utilization of biopolymers dates back to antiquity, with early civilizations employing cellulose, starch, and proteins for structural and functional applications. Ancient Egyptians used protein-based adhesives and plant gums, while Mesoamerican cultures utilized natural rubber for toolmaking and waterproofing. The scientific characterization of biopolymers

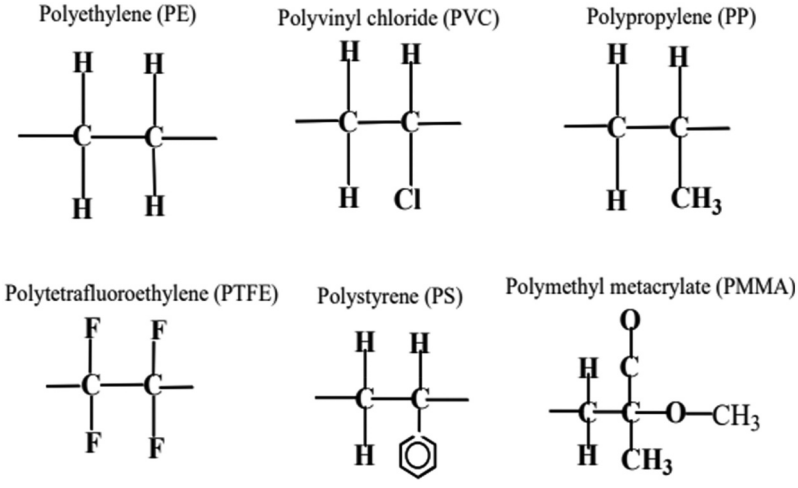
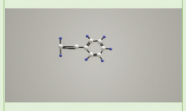
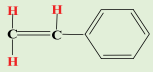
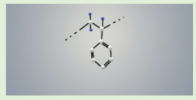
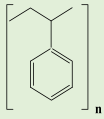
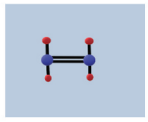
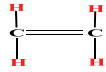
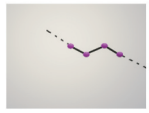
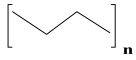
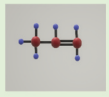
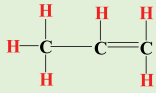
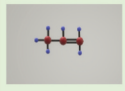
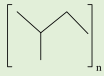
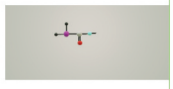
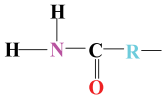
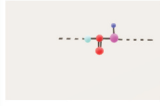
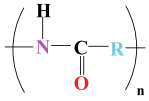
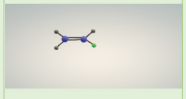
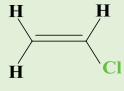
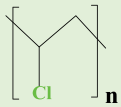


Illustration 1. Structures of the monomers of some polymers.

Table 1. Various examples of polymers along with their monomeric units are shown in the above table.

MONOMER	MONOMERIC UNIT	STRUCTURE	NAME	APPLICATIONS	REFERENCE
Styrene	 	 	Polystyrene	Pharmaceuticals, Building applications, Food and water applications	[2]
Ethylene	 	 	Polyethylene	Biomedical, pharmaceuticals, commercial, biotechnological	[3]
Propylene	 	 	Polypropylene	Packaging, housewares, consumer products	[4]
Amide	 	 	Polyamide	Textile and fiber, automotive industry, engineering plastics	[5]
Vinyl chloride	 	 	Polyvinyl chloride	Packaging, construction, plumbing, wire insulation	[6]

began in the 19th century, but large-scale industrial modifications gained traction in the mid-20th century due to the demand for sustainable alternatives to synthetic polymers. The derivatization of polysaccharides, such as carboxymethyl cellulose (CMC) and hydroxyethyl starch, expanded their functional versatility. Guar gum, a galactomannan from Cyamopsis tetragonoloba, became significant for its rheological properties, leading to derivatives like Carboxymethyl Guar Gum (CMGG), valued for its solubility and biomedical applications. Recent advancements in polymer chemistry, nanotechnology, and green synthesis have driven the development of high-performance biopolymers like polylactic acid (PLA) and polyhydroxyalkanoates (PHA), reinforcing the shift toward biodegradable and eco-friendly material solutions. **Carboxymethyl Guar Gum (CMGG)** has gained attention as a promising biodegradable alternative due to its **enhanced solubility, biocompatibility, and tunable physicochemical properties**. CMGG, a modified guar gum derivative, exhibits significant potential in diverse applications, from food and pharmaceuticals to sustainable material development. By integrating **CMGG into polymer research**, advancements in biopolymer-based materials align with green chemistry principles and environmentally conscious innovations, paving the way for sustainable alternatives to synthetic polymers. The polymers can be classified into various categories on a variety of bases, but we can say they are mainly of two types: natural polymers and synthetic Polymers (Illustration (2)).

Natural polymers, such as cellulose, starch, and proteins, have been utilized for centuries in textiles, food, and construction.[7] For instance, cellulose,[8] the most abundant natural polymer on Earth, forms the structural component of plant cell walls and has long been used in products like paper and cotton fabrics. Proteins, another class of natural polymers, play a crucial role in biological processes and are the building blocks of life, forming everything from hair and skin to enzymes that catalyze vital reactions.[9] In contrast, synthetic polymers, first developed in the early 20th century, have had an even more profound impact on industrial applications. These polymers are typically derived from petrochemical sources, such as crude oil and natural gas, and include materials like polyethylene, polypropylene, polystyrene, and polyvinyl chloride (PVC).[10,11] The development of synthetic polymers has enabled the mass production of durable, flexible, and lightweight materials, which have been instrumental in shaping the modern world. The arduous nature of these materials toward biodegradation leads to their buildup in the environment over long years.[12,13]

Every year, the approximate scale of manufacturing synthetic polymer is 140 million tons globally. Global Plastic production surged from 348 million tons to 359 million tons from 2017 to 2018 alone, as represented in Table 2.[14–16]

While synthetic polymers have undeniably transformed modern life, they also pose significant environmental challenges. Most traditional polymers, such as polyethylene, polypropylene, and PVC, are derived from nonrenewable fossil fuels and are not biodegradable.[17] As a result, plastic waste has accumulated in landfills, oceans, and natural ecosystems, leading to widespread pollution and environmental degradation. In 2015, approximately 6,300 million metric tons of plastic were produced, with 79%

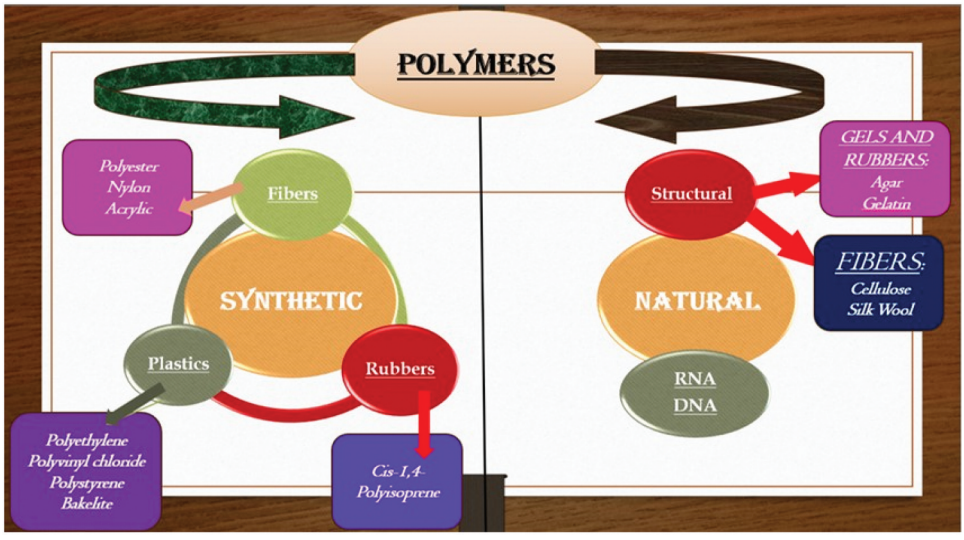
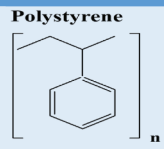
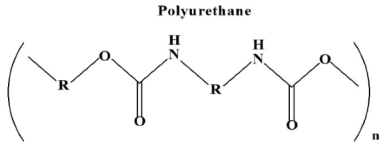
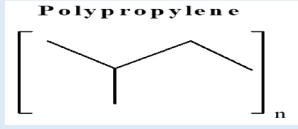
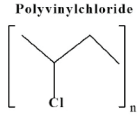
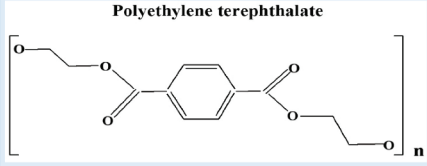
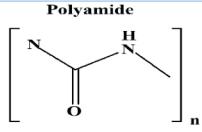
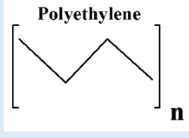


Illustration 2. Schematic representation of existence of polymers.

Table 2. Represents some major synthetic polymers produced worldwide in the year 2017 to 2018.

SYNTHETIC POLYMERS NAMES (GLOBALLY PRODUCED IN 2016)	STRUCTURE	PRODUCTION (IN MILLIONS OF TONNES)
Polystyrene		13
Polyurethane		29
Polypropylene		66
Polyvinylchloride		41
Polyethylene terephthalate		35
Polyamide		07
Polyethylene		122

accumulating in landfills, potentially reaching 12 billion tons by 2050.[18] The durability that makes polymers so valuable in products also makes them resistant to decomposition, with some plastics taking hundreds of years to break down.[20] In addressing the escalating intricacies of global environmental challenges, sustainability paradigms, advanced production methodologies, green chemistry innovations, environmentally conscious engineering, and ecosystem fortification have emerged as pivotal objectives in the design and development of future products, materials, and processes. The shift toward sustainable polymer production focuses on biodegradable starch-based composites from renewable sources. However, their weak mechanical strength and high hydrophilicity limit applications. Incorporating agro-waste cellulosic fibers enhances durability, water resistance, and biodegradability, while advanced techniques like solution casting and ultrasound-assisted extraction optimize fiber dispersion. Persistent challenges, including moisture-induced degradation, drive innovations such as hydrophobic modifications and nanoparticle integration. Future advancements aim to develop sustainable alternatives that replace conventional plastics while minimizing environmental impact.[20] Conventional petroleum-based plastics, especially single-use variants, drive environmental degradation and waste accumulation. The shift to eco-friendly alternatives meets

consumer demand while mitigating the ecological impact of traditional polymer production.[21] Noteworthy strides have been made toward a more sustainable and economically viable world, with the advancement of bio-based products, particularly biopolymers, epitomizing this progressive framework.[22]

2. Biopolymers

Biopolymers are increasingly gaining attention as sustainable alternatives to conventional synthetic polymers, primarily due to their biodegradable nature, renewability, and the reduced environmental impact they offer. As global environmental concerns such as plastic pollution, fossil fuel depletion, and climate change intensify, the need for more sustainable material solutions becomes paramount. Biopolymers, derived from natural sources like plants, animals, and microorganisms, are positioned as a key component in the transition toward greener industrial practices.[23] These materials are eco-friendly and offer diverse applications across multiple industries, including packaging, textiles, biomedical fields, and agriculture. Biopolymers are polymers produced by living organisms or synthesized from renewable biological sources.[24] They consist of long chains of repeating monomer units, similar to synthetic polymers. Still, the critical distinction is that their monomers are derived from renewable biological sources such as carbohydrates, proteins, and lipids. Biopolymers stand out from conventional polymers because they are biodegradable, meaning they can be broken down by microorganisms in natural environments, reducing the accumulation of waste. The environmental advantages of biopolymers make them increasingly vital in combating the global issue of plastic pollution, especially in landfills and oceans. Biopolymers are a class of polymers in which the prefix "bio" alludes to living. Thus, we can say "bio" suggests that the polymers are congenitally formed or synthesized from life forms such as microorganisms, plants, etc. But materials synthesized via synthetic chemistry from biological sources can also be included in this definition.[25] E.g., proteins, fats, oils, sugars, etc. Hence, in strict terms, "*biodegradation of a polymer*" can be characterized as the degradation of a polymer's chemical and physical characteristics, as well as a decrease in its molecular mass caused by microorganisms under both aerobic and anaerobic conditions, aided by abiotic variables and their chemical reactions such as photodegradation, oxidation, and hydrolysis.[26,27] Biopolymers possess diverse properties such as biodegradability, renewability, and biocompatibility. They can vary in mechanical strength, flexibility, thermal stability, and water solubility, depending on their chemical structure and origin. These properties make biopolymers ideal for applications in packaging, medicine, agriculture, and environmental sustainability, providing eco-friendly alternatives to traditional, non-degradable polymers.[28] Table 3 below gives a few examples of the applications and properties of various biopolymers along with their properties.

2.1. Classification of biopolymers

Biopolymers are classified into three major categories based on their origin and production method: natural biopolymers, synthetic biopolymers from renewable resources, and microbial biopolymers. Each category has distinct features and potential applications.[47] Natural biopolymers include cellulose, starch, and silk; synthetic biopolymers derived from renewable resources include polylactic acid and polyhydroxyalkanoates, while microbial biopolymers such as xanthan gum and alginate are produced by microorganisms.[48] Each type offers unique properties and applications that promote sustainability and eco-friendliness (Illustration 3).

2.1.1. Natural biopolymers

Natural biopolymers are derived from biological sources such as plants, animals, and microorganisms. They are typically obtained through direct extraction and modification processes.

- **Polysaccharides**: Examples include cellulose, starch, and chitin, which serve structural and storage functions in plants and arthropods, respectively.
- **Proteins**: Biopolymers such as collagen, keratin, and silk provide mechanical strength and biological functions.
- **Nucleic Acids: DNA** and **RNA** store and transmit genetic information.

2.1.2. Synthetic biopolymers from renewable resources

These biopolymers are synthesized using monomers derived from biological sources through chemical polymerization.

- **Polylactic Acid (PLA)**: Derived from corn starch or sugarcane, PLA is widely used in biodegradable packaging and medical implants.
- **Polyhydroxyalkanoates (PHAs)**: These are microbial polyesters used in biomedical applications and biodegradable plastics.

- **Polycaprolactone (PCL)**: A synthetic biodegradable polyester applied in drug delivery and tissue engineering.

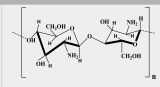
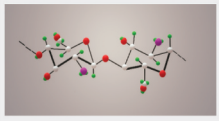
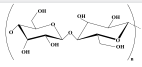
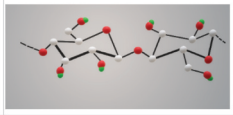
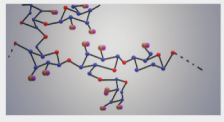
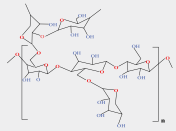
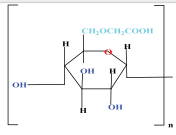
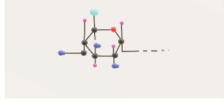
2.1.3. Microbial biopolymers

Microbial biopolymers are produced by microorganisms through fermentation. These polymers have unique properties, including high viscosity and stability.

- **Xanthan Gum**: A polysaccharide produced by *Xanthomonas campestris*, widely used in food thickening and oil recovery.
- **Alginate**: Extracted from brown algae and bacteria, it is utilized in **wound dressing**, **drug delivery**, and **bioprinting**.
- **Bacterial Cellulose**: Produced by *Acetobacter xylinum*, used in **biomedical applications** and **cosmetics**.

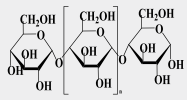
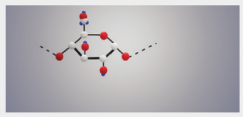
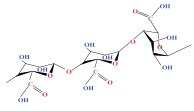
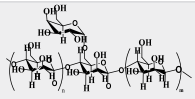
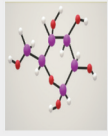
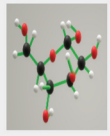
Biopolymers are crucial to the functioning of nature. They are incredibly helpful for carrying out various tasks, including energy storage, transmitting genetic information, and building cells. All biopolymers are produced through enzymatic processes, whether in

Table 3. Outlined examples of biopolymers, distinct applications and properties.

BIOPOLYMER NAME	BIOPOLYMER STRUCTURE	PRODUCTION METHOD	MONOMERIC UNIT	PROPERTIES	APPLICATIONS	REFERENCE
Chitosan	 	Deacetylation of chitin in industries	D-glucosamine and N-acetylglucosamine	High anti-microbial, non-toxicity, and anti-tumor.	Anti-microbial, biomedical, pharmaceutical	[29,30]
Cellulose	 	Altered natural polymer	Beta-glucose	Thermal and acoustic.	Coatings, biomedical	[31,32]
TKP	 		Glucose, Galactose and Xylose	Thickening agent, emulsifying agent, adhesive, hydration	Food industry, textile industry	[33,34]
CMTKG	 		D-xylose, D-glucose, D-galactose	Gelling agents, low cold-water solubility	Pharmaceuticals, food industry, Paper industry, agriculture, tissue engineering	[35-37]

(*Continued*)

Table 3. (Continued).

Starch	 	Plant-based	Glucose	Thickening agent, crystallizing, granules formation, high viscosity	The food industry, textile, paper, petroleum engineering, and agriculture.	[38-40]
Alginate			Mannuronic acid and Guluronic acid	High viscosity, hydrogel formation, intrinsic viscosity	Biomedical	[41-43]
Guar-Gum	  	From Guar plant	Mannose and Galactose	Non-ionic, thickening agents, resin, dispersants, drilling fluid formulations, green corrosion inhibitors, high viscosity, temperature thinning effect	Food industry, petroleum industry, biological applications, waste management, pharmaceutical	[44-46]

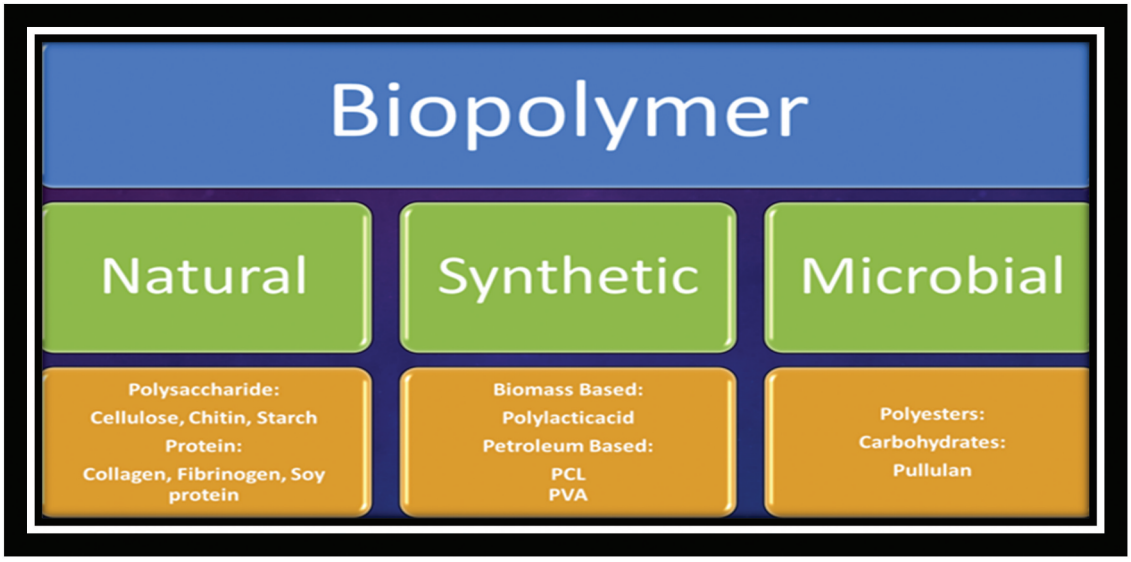


Illustration 3. Classification of biopolymers according to their origin.

the cytoplasm, different organelles, or compartments of the cells, at the membranes of the cytoplasm or cell wall constituents, on the surface of cells, or even extracellularly. The fabrication of a biopolymer may begin in one area of a cell and continue in another area as it takes place.[49–51] Life forms are capable of creating a wide variety of distinct polymers and in most species, these biopolymers are responsible for the majority of the dry matter in the cells. Biopolymers perform several critical and diverse tasks for cells as varied as their

Table 4. Represents several biopolymers and their nanofibers along with their characteristics, applications, and properties.

Biopolymer	Metal	Method	Applications	Characterization	Properties	References
Chitosan	Fe_2O_3	The Co-precipitation Method, Thermal decomposition method	Pharmaceuticals, biomedical,	XRD, FTIR, TEM	Surface modifications, low toxicity, super magnetic,	[55–58]
Chitosan	ZnO NPs	Sol-gel method	Medical, Food packaging	XRD, FTIR, UV, TEM, X-ray, SEM	Antibacterial, antifungal, non-toxic	[57,59,60]
Starch	Silver NPs	Reduction Method	Detection of metal ions,	UV/Vis, TEM, HRTEM	High stability, facile strength, high water-solubility	[61]
Starch	Cationized starch-Pectin NPs	Ionotropic-gelation method	Pharmaceuticals, encapsulation	XRD	Anti-bacterial	[62]
Carboxylate-polyacrylamide	Fe_3O_4	Co-precipitation method				[63]
Alginate-chitosan	Ag-nps	Ionotropic pre-gelation followed by complexation	Encapsulation, drug delivery, therapeutical, Insulin administration, DNA carrier	SEM, FTIR, UV, XRD	Antimicrobial, Anti cancerogenic, Anti sense	[64–66]
Guar Gum	Fe_3O_4	Co-precipitation and cross-linking method	Biological, Biomedical,	TEM, XRD, FTIR, VSM	Cancer treatment, hypothermia, drug delivery, catalytic	[64,66,67]

structural makeup. The functions of biopolymers are as diverse and intricate as their molecular structures. They find applications across numerous sectors, including healthcare, textiles, commercial products, cosmetics, food processing, packaging, and waste management.[52–54] Biopolymers also play critical roles in genetic information storage, catalysis of enzymatic and non-enzymatic reactions, detection of biotic and abiotic stimuli, and cellular defense mechanisms. Table 4 presents literature data on various biopolymers and their nanofibers, detailing their characteristics, properties, and applications.

2.2. Synthesis of biopolymers

The synthesis of biopolymers varies depending on the type of polymer and the source material. Below is an overview of common synthesis methods:

(1) **Enzymatic Polymerization**: Natural biopolymers such as proteins, polysaccharides, and nucleic acids are synthesized in biological systems using enzyme-mediated polymerization.[68]
(2) **Chemical Polymerization**: Synthetic biopolymers such as PLA and PHA are produced through chemical processes such as condensation polymerization and ring-opening polymerization.[69]
(3) **Microbial Fermentation**: Certain biopolymers, like xanthan gum and PHAs, are synthesized by bacteria using fermentation techniques under controlled conditions.[70]
(4) **Co-precipitation and Cross-linking Methods**: These are widely used for biopolymer-based nanocomposites, where metal nanoparticles or other reinforcing agents are incorporated to enhance their properties.[71]
(5) **Electrospinning**: A technique used for fabricating biopolymer nanofibers by applying a high-voltage electric field to polymer solutions, leading to their application in tissue engineering and drug delivery.[72]
(6) **Hydrothermal Synthesis**: This method utilizes high-temperature and high-pressure aqueous environments to synthesize biopolymer-based nanocomposites with improved thermal and mechanical properties.[73]
(7) **Sol-Gel Processing**: A wet-chemical technique for producing biopolymer-based hybrid materials, involving the transition of a system from a liquid solution into a solid gel phase, commonly used in biomaterials and coatings.[74]

2.3. Comparative analysis of biopolymers

The comparative analysis highlights the diverse origins, properties, and applications of biopolymers. Cellulose, derived from plants, is known for its high strength and biodegradability, making it ideal for paper, textiles, and pharmaceuticals.[75] Chitin, obtained from crustaceans and fungi, exhibits biocompatibility and antibacterial properties, supporting applications in wound healing and agriculture.[29] PLA, sourced from corn and sugarcane, is a thermoplastic and compostable polymer

widely used in packaging and medical implants.[21] PHAs, produced by bacteria, are highly biodegradable and versatile, finding applications in biomedical fields and bioplastics. Alginate, derived from algae and bacteria, is hydrophilic with excellent gelling properties, making it suitable for drug delivery and food stabilization.[66]

While biopolymers offer numerous advantages, including biodegradability, biocompatibility, and renewable sourcing, their widespread application faces several challenges. The primary limitations include their mechanical weakness compared to synthetic polymers, susceptibility to microbial degradation in specific environments, and often higher production costs. Additionally, many biopolymers exhibit poor thermal stability, limiting their use in high-temperature applications. Another significant concern is their inconsistent quality due to variations in natural sources, which can impact processing and end-use properties. Addressing these challenges through advanced material modifications, cross-linking techniques, and nanocomposite integration is crucial for expanding their practical applications.

Biopolymers can be classified using various criteria without strict limitations. These classifications can be based on factors such as their origin, biodegradability, thermal responsiveness, and composition, among others. Multiple factors contribute to categorizing biopolymers into distinct classes and sub-classes.[76] Illustration (4) depicts their classification based on biological degradation.

2.4. Applications of biopolymers

Biopolymers provide environmentally friendly alternatives to synthetic materials, finding use in various industries due to their versatility and sustainable properties. Their applications span biomedical devices, drug delivery systems, tissue engineering, and wound dressings, as well as food packaging, agriculture, environmental remediation, biosensing, and energy storage, including lithium-ion batteries.[14,77] These materials also contribute significantly to reducing plastic waste by replacing conventional polymers with biodegradable and renewable options. They have diverse applications across multiple industries, including pharmaceuticals, biomedical, food, agriculture, textiles, and cosmetics.[78,79] Illustration (5) highlights key applications of biopolymers across multiple industries.

Biopolymer-based nanofibers have gained attention due to their biodegradability, biocompatibility, and high surface-area-to-volume ratio. Their unique properties

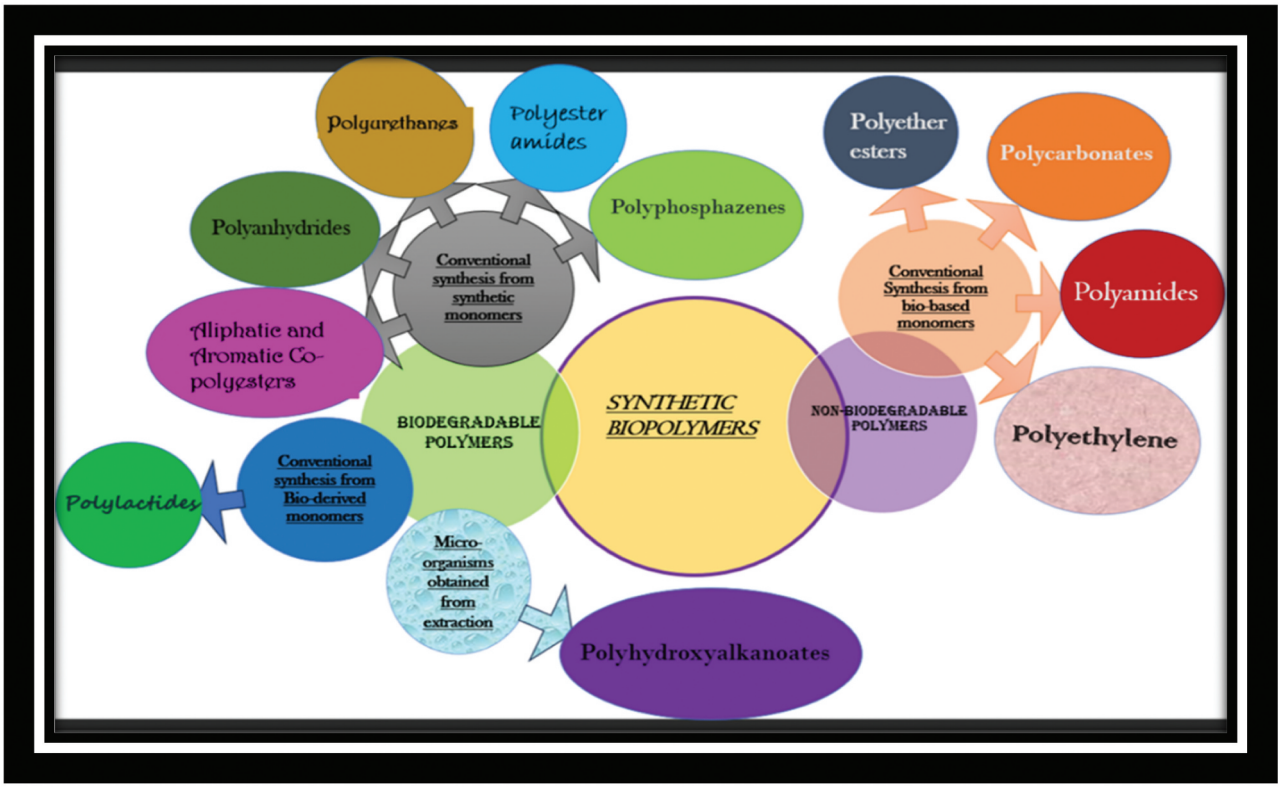


Illustration 4. classification of biopolymers based on their biodegradability.

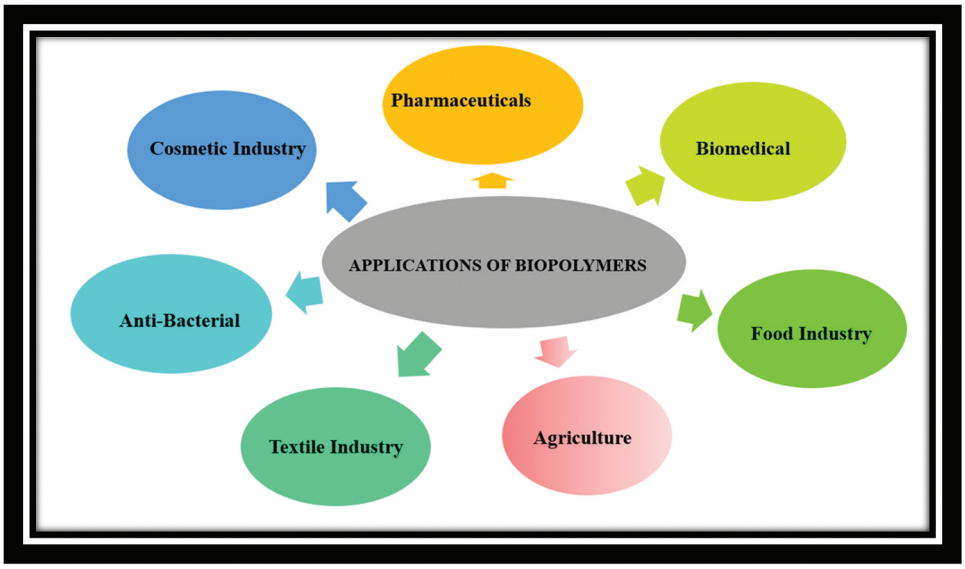


Illustration 5. Key applications of biopolymers across various industries.

make them ideal for drug delivery, wound healing, and environmental remediation. Nanofibers from cellulose, chitosan, and silk fibroin exhibit superior mechanical strength, thermal stability, and antimicrobial properties.[80] Additionally, their functionalization enhances their effectiveness in biosensing, food packaging, and filtration systems, positioning biopolymer nanofibers as promising sustainable materials.[81,82]

Guar Gum, a natural biopolymer, exemplifies these advantages with its exceptional thickening and gelling abilities. Widely used in food, pharmaceuticals, and environmental applications, Guar Gum-based nanocomposites offer enhanced mechanical strength, antibacterial properties, and stability. Table 5 presents various Guar Gum-metal nanostructures, showcasing their characteristics, properties, and applications in biomedical engineering, food packaging, and environmental remediation.

3. Guar Gum

Guar Gum is a natural biopolymer obtained from the seeds of the guar plant *(Cyamopsis tetragonoloba)*. Its molecular structure consists of a backbone of linear mannose units (β-D-mannopyranose) linked by (1→4) glycosidic bonds, with galactose units (α-D-galactopyranose) attached as side branches (Illustration (6)).[90] Typically, the ratio of mannose to galactose in Guar Gum is about 2:1, which contributes to its unique thickening and gelling properties. This structure allows Guar Gum to absorb water and form viscous solutions, making it a versatile ingredient in industries. This unique configuration contributes to its excellent thickening and gelling properties, making it a valuable additive in various industries, including food, pharmaceuticals, and cosmetics.[91] Guar Gum's ability to retain water and form viscous solutions enhances its functionality in applications such as stabilizers, emulsifiers, and binders.

Galactose and mannose (Illustration (7)) are both simple sugars (monosaccharides) that play crucial roles in various biological processes. While galactose is primarily found in lactose and is essential for cellular metabolism, mannose is involved in glycoprotein synthesis and is significant for cell signaling and recognition.[92]

Guar Gum exhibits several notable physical properties that contribute to its versatility in various applications. One of its most significant characteristics is its ability to form highly viscous solutions when hydrated, which allows it to function effectively as a thickening and stabilizing agent.[93,94] Guar gum has excellent water-absorbing and thickening properties, making it highly effective as a stabilizer and emulsifier in various industrial and food applications. Additionally, its solubility in both cold and hot water makes it particularly valuable in the food industry. The viscosity of Guar Gum solutions can be influenced by factors such as concentration, temperature, and pH, providing flexibility in formulation.[95] Furthermore, its nonionic nature allows it to interact well with other ingredients without altering their properties, making it an ideal

Table 5. Represents several nanostructures of Guar Gum along with their characteristics, applications, and properties.

Biopolymer	Metal	Method	ApplicationS	Characterization	Matrix formed	References
Guar Gum	Ag	Green Synthesis	Biomedical, Environmental	XRD, FTIR, TEM	Hydrogels	[83]
Guar Gum	ZnO NPs	Cross-linking method	Food packaging	XRD, FTIR, UV, TEM, X-ray, SEM	Hydrogel Nanocomposite film	[84]
Guar Gum	Silver NPs	In-situ Method	Food packaging, Anti-bacterial	UV/Vis, TEM, HRTEM	Nanocomposite-film	[85]
Guar Gum/ Graphene Oxide	CuO	Ex-Situ method	Waste water purification	XRD, TEM, UV, NMR	Nanocomposite powder	[86]
Guar Gum	TiO_2/ Fe_3O_4	Sol-Gel method	Photocatalytic antimicrobial activity	SEM, FTIR, UV, TEM, XRD	Nanocomposite powder	[87]
Guar Gum	Fe_3O_4	Co-precipitation and cross-linking method	Biological, Biomedical,	TEM, XRD, FTIR, VSM	Cancer treatment, hypothermia, drug delivery, catalytic	[88, 89]

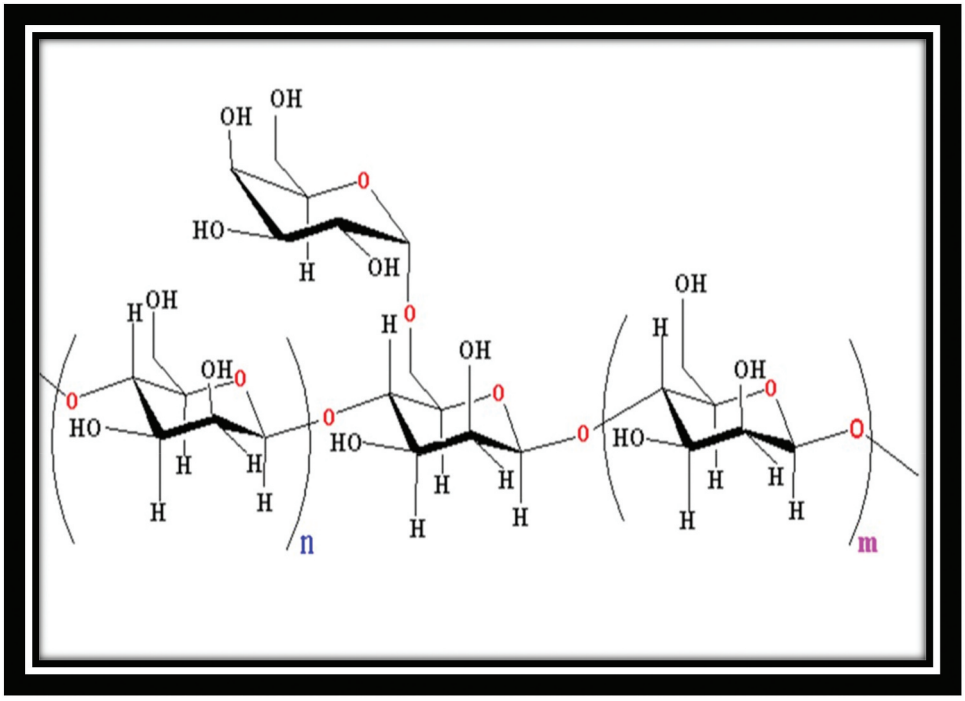


Illustration 6. chemical structure of Guar Gum.

component in pharmaceuticals, cosmetics, and various industrial applications. Guar Gum exhibits several notable physical properties that significantly contribute to its versatility across various applications. One of its most remarkable characteristics is its ability to form highly viscous solutions when hydrated, allowing it to function effectively as a thickening and stabilizing agent.[96] This biopolymer can absorb water up to ten-twenty times its weight, resulting in a gel-like consistency that enhances the texture and mouthfeel of food products. Its high water retention capacity makes it particularly useful in the food industry, where it helps to maintain moisture in baked goods, sauces, and dairy products.[97]

3.1. Properties of Guar Gum

Guar Gum is insoluble in organic solvents and forms a highly viscous solution when dissolved in cold water, especially upon heating. Its solutions exhibit notable buffering capacity and remarkable stability within a pH range of 4.0 to 10.5.[98] The exceptionally high viscosity of Guar Gum is attributed to unique intermolecular interactions, including entanglements, and its substantial hydrodynamic volume in solution.[99] Notably, an increase in temperature causes significant thinning of Guar Gum solutions, as the polymer molecules lose hydration, making it one of the most effective natural polymers.[100] Additionally, as the temperature rises, the pH decreases, and particle size reduces, facilitating faster dissolution and

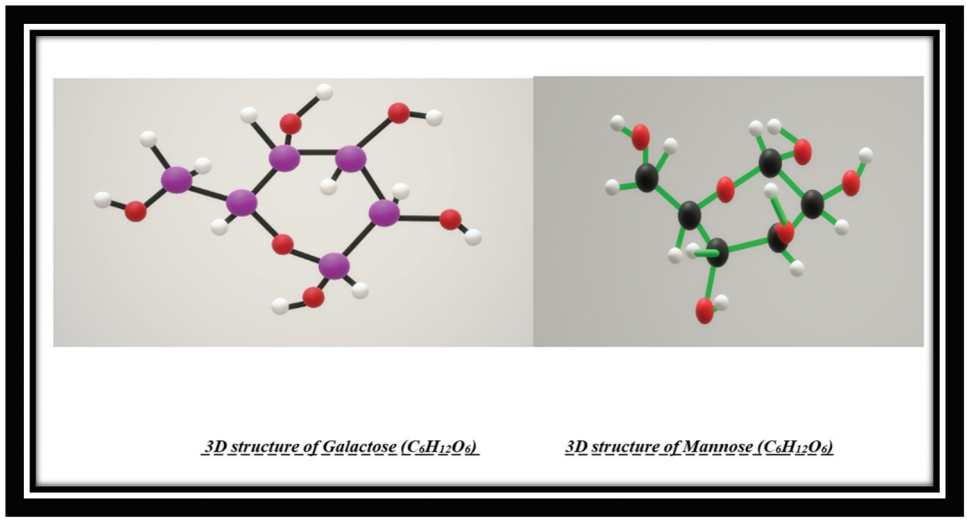


Illustration 7. chemical 3-D structure of galactose and Mannose.

viscosity development.[101] Even at minimal concentrations, Guar Gum rapidly hydrates in hot and cold water to form thick colloidal solutions, demonstrating its gel-forming capacity and excellent water-binding properties through hydrogen bonding in aqueous environments.[102]

This property is particularly valuable in the **food industry**, where Guar Gum is widely used as a thickener and stabilizer in ice creams, dairy products, and sauces, preventing ice crystal formation and improving texture. Its high viscosity also makes it a preferred **fracking agent in the oil and gas industry**, where it is employed in hydraulic fracturing fluids to enhance the efficiency of proppant transport. Additionally, in **pharmaceutical applications**, its ability to form gels and retain moisture makes it an ideal component in controlled drug delivery systems, where it modulates drug release rates. Guar Gum is biodegradable, biocompatible, and nontoxic, further enhancing its appeal for various applications. Additionally, it functions as a dispersant in organic systems and exhibits a high capacity for cross-linking and chemical modification, making it a versatile and valuable biopolymer[103] (Illustration (8)). For instance, in **cosmetic formulations**, Guar Gum is utilized in lotions and creams to improve viscosity and enhance moisture retention, contributing to better skin hydration. Moreover, its use in **wastewater treatment** as a flocculant aids in removing suspended particles, demonstrating its environmental significance.

Given its extensive industrial applications, the production and **processing of Guar Gum** play a crucial role in determining its quality, purity, and functional attributes. Various processing techniques are employed to ensure the efficient extraction and refinement of Guar Gum, optimizing its physicochemical properties for specific applications. The extraction process significantly influences its molecular weight distribution, hydration capacity, and performance in end-use formulations.

3.2. Processing of Guar Gum

India is the world's foremost producer of guar, with various mechanical techniques employed in the industry for extracting Guar Gum from guar seeds. Key methods include polishing, attrition, roasting, and sifting. Due to the varying hardness of the constituent components, the endosperm and its parts can be easily separated. Guar splits are produced following the removal of the hull and germ. After heat treatment, the hull can be effectively separated using attrition milling or various types of impact mills. Sieving facilitates the extraction of the finer fractions of the germ and hull from the endosperm. The subsequent stage involves grinding the Guar Gum into a powder. Finally, the Guar Gum is purified through a process that includes precipitation, dissolution in water, and recovery using ethanol or isopropanol. The result is purified Guar Gum, ready for various applications.[93,104]

3.3. Chemical modification of Guar Gum

Guar Gum, despite its multifaceted utility, exhibits inherent limitations that constrain its applicability across various industrial and scientific domains.[89] One of its principal shortcomings is its restricted solubility in organic solvents, thereby limiting its integration into non-aqueous systems such as pharmaceutical

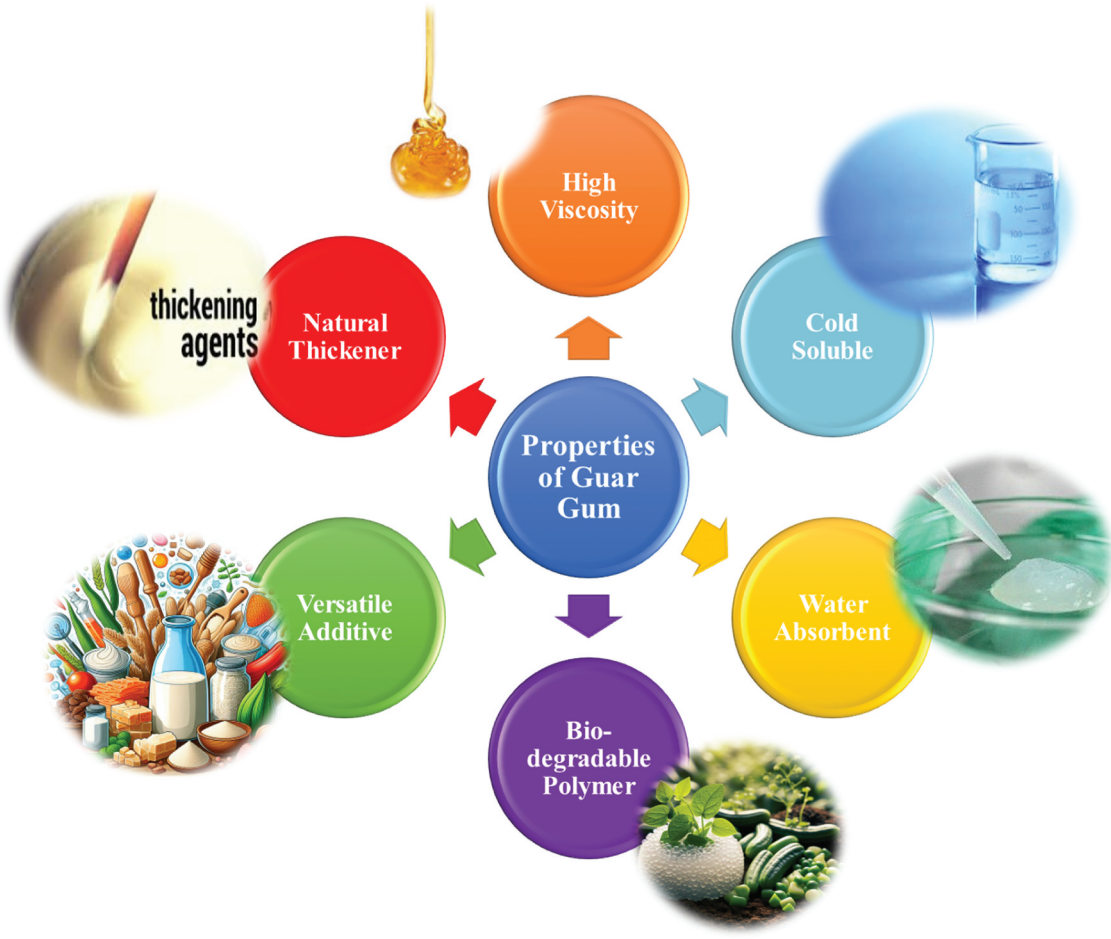


Illustration 8. properties of Guar Gum.

formulations, hydrophobic coatings, and industrial adhesives. This solubility constraint necessitates chemical modification to broaden its compatibility with diverse solvent environments. Furthermore, Guar Gum's physicochemical stability is compromised under extreme pH conditions and elevated temperatures, rendering it unsuitable for applications requiring thermal resilience, such as high-temperature food processing, enhanced oil recovery, and cosmetic formulations demanding extended stability.[105] Another critical drawback is the pronounced variability in its rheological properties, which exhibit significant sensitivity to temperature fluctuations and shear forces. This inherent instability can impede its efficacy as a viscosifier and stabilizing agent in sectors such as food technology, pharmaceuticals, and polymer composites, where precise rheological control is imperative. Additionally, Guar Gum is highly susceptible to microbial degradation, resulting in a reduced shelf life, particularly in aqueous-based formulations. This necessitates the incorporation of antimicrobial agents or molecular modifications to enhance its structural integrity and longevity. Economic volatility in the global market further exacerbates its commercial unpredictability, making cost-effective synthetic or modified biopolymeric alternatives highly desirable for industrial-scale utilization. Moreover, its ionic sensitivity predisposes it to undesirable precipitation in electrolyte-rich environments, thereby restricting its deployment in applications such as wastewater treatment, biomedical formulations, and electrolyte-based industrial processes where colloidal stability is paramount.[106,107]

To circumvent these limitations, targeted chemical modifications have been extensively explored to refine the physicochemical properties of Guar Gum, thereby augmenting its functionality and expanding its industrial viability. Modification strategies such as etherification, esterification, and oxidation facilitate substantial enhancements in solubility, stability, and application-specific performance. For instance, etherification via carboxymethylation introduces carboxymethyl functionalities into the polymer backbone, yielding carboxymethyl Guar Gum (CMGG). This derivatization markedly amplifies hydrophilicity and aqueous



Illustration 9. some examples of the applications of derivatives of Guar Gum.

solubility, thereby broadening its applicability in pharmaceutical excipients, food texturizers, and cosmetic formulations that necessitate superior dispersibility and stabilization.[108–110] Esterification, conversely, engenders the incorporation of ester moieties, conferring enhanced thermal and mechanical robustness to Guar Gum derivatives. These modifications foster the development of polymeric matrices with superior film-forming attributes, rendering them ideal candidates for biodegradable packaging materials, controlled-release systems, and advanced polymeric composites.[88] Oxidative functionalization modulates molecular weight and introduces reactive oxygen-containing groups, thereby tailoring viscosity profiles and reactivity. This structural transformation enables oxidized Guar Gum to exhibit enhanced performance in specialized applications such as drilling fluid rheology optimization, soil conditioning for sustainable agriculture, and high-performance dispersion systems.[94]

Beyond addressing the intrinsic deficiencies of native Guar Gum, these chemical modifications also facilitate its transition into biomedical and high-performance material applications. CMGG, for example, exhibits hydrogel-forming capabilities, making it a promising candidate for controlled drug delivery matrices, wound healing scaffolds, and regenerative medicine frameworks.[95] By leveraging molecular engineering approaches, chemically functionalized Guar Gum derivatives are progressively meeting the stringent performance requirements of industries such as food processing, coatings and pigments, mining, oil and gas, pulp and paper, water purification, cosmetics, and the synthesis of superabsorbent biomaterials (Illustration (9)).

Some of the derivatives of Guar Gum are listed below,

(1) Carboxymethyl Guar Gum (*CMGG*).[108]
(2) Hydroxypropyl Guar Gum (*HPGG*).[111]
(3) O-carboxymethyl- O-hydroxypropyl Guar Gum (*CMHPGG*).[112]
(4) Hydroxymethyl Guar Gum (*HMGG*).[113]
(5) Acryloyloxy Guar Gum.[114]
(6) Methacryloyl Guar Gum.[115]
(7) O-carboxymethyl- O-2 hydroxy-3- (trimethyl ammonia propyl) Guar Gum (CMHTPG).[94]

The most distinctive feature of Guar Gum and its derivatives is the presence of hydroxyl groups, which allows for the modification of their functionalization and structural formula.[116] The diverse properties and applications of Guar Gum lay the groundwork for exploring its derivative, Carboxymethyl Guar, which offers enhanced functionality and versatility in various fields.

4. Carboxymethyl Guar Gum

Carboxymethyl Guar Gum (CMGG) is a chemically modified derivative of Guar Gum, a natural polysaccharide extracted from the seeds of the guar plant (*Cyamopsis tetragonoloba*).[117] Guar Gum is widely recognized for its thickening, gelling, and stabilizing properties, which have made it a valuable ingredient in various industries, including food, pharmaceuticals, and cosmetics. However, the native form of Guar Gum has limitations in terms of solubility and functionality under specific conditions. To overcome these limitations, CMGG is synthesized through the carboxymethylation process,

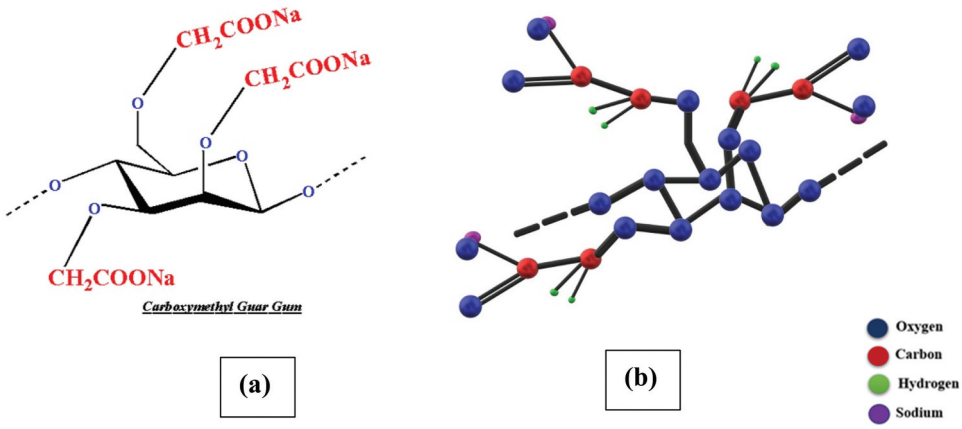


Illustration 10. (a) chemical structure of CMGG; (b) 3-D structure of CMGG.

wherein carboxymethyl groups are introduced into the Guar Gum molecule. This modification enhances the solubility of CMGG in cold water, increases its viscosity, and imparts improved stability across a broader pH range. These properties make CMGG a versatile biopolymer with a wide array of applications.[118] CMGG's unique characteristics, including its ability to form gels and films, have led to its adoption in drug delivery systems, where it acts as a controlled release agent, and in the food industry, where it serves as an effective thickener and stabilizer.[119] Additionally, its biocompatibility and nontoxic nature position CMGG as a promising candidate for biomedical applications, particularly in tissue engineering and pharmaceutical formulations.[120] As research continues to explore the potential of CMGG, its role as a functional additive in various sectors is becoming increasingly significant, driving innovation and expanding the boundaries of biopolymer applications.[121] This section delves into the synthesis, properties, applications, and prospects of Carboxymethyl Guar Gum, highlighting its importance and its contribution to advancements in material science and technology.

4.1. Introspection of chemical structure of carboxymethyl Guar Gum (CMGG)

Carboxymethyl Guar Gum (CMGG) is a chemically modified derivative of Guar Gum, characterized by the introduction of carboxymethyl groups into its polysaccharide backbone. The chemical structure of CMGG consists of a linear chain of mannose and galactose units, which are the primary components of Guar Gum[122] (Illustration 10). The carboxymethylation process occurs through the substitution of hydroxyl groups in the Guar Gum molecule with carboxymethyl groups, resulting in enhanced solubility and reactivity.[123]

This modification introduces anionic charges, allowing CMGG to interact more effectively with water molecules and other ionic species, significantly improving its functional properties.[124] One of the most notable physicochemical properties of CMGG is its solubility in water, particularly at lower temperatures. Unlike native Guar Gum, which can be less soluble in cold water and requires heat for effective hydration, CMGG exhibits a rapid dissolution process due to the presence of carboxymethyl groups.[95,121] This enhanced solubility contributes to its versatility in various applications, particularly in the food and pharmaceutical industries. Furthermore, CMGG demonstrates superior viscosity characteristics compared to native Guar Gum. The viscosity of CMGG solutions is highly dependent on concentration, temperature, and shear rate, making it an excellent thickening agent.[125] Its ability to form gels and films is another critical property that enhances its applicability in areas such as drug delivery systems and food formulation. In comparison to native Guar Gum, CMGG offers distinct advantages due to its modified structure.[100] While native Guar Gum primarily exhibits neutral behavior in aqueous solutions, the anionic nature of CMGG allows it to maintain stability over a broader pH range, making it suitable for diverse formulations. Additionally, the viscosity profile of CMGG is more favorable in terms of achieving desired textural properties in various products.[91] This adaptability makes CMGG a more effective thickening and stabilizing agent, especially in formulations that require consistency and reliability. The structural modifications and resulting physicochemical properties of Carboxymethyl Guar Gum position it as a highly functional and versatile biopolymer, surpassing the limitations of its native counterpart and paving the way for innovative applications across multiple industries.[126]

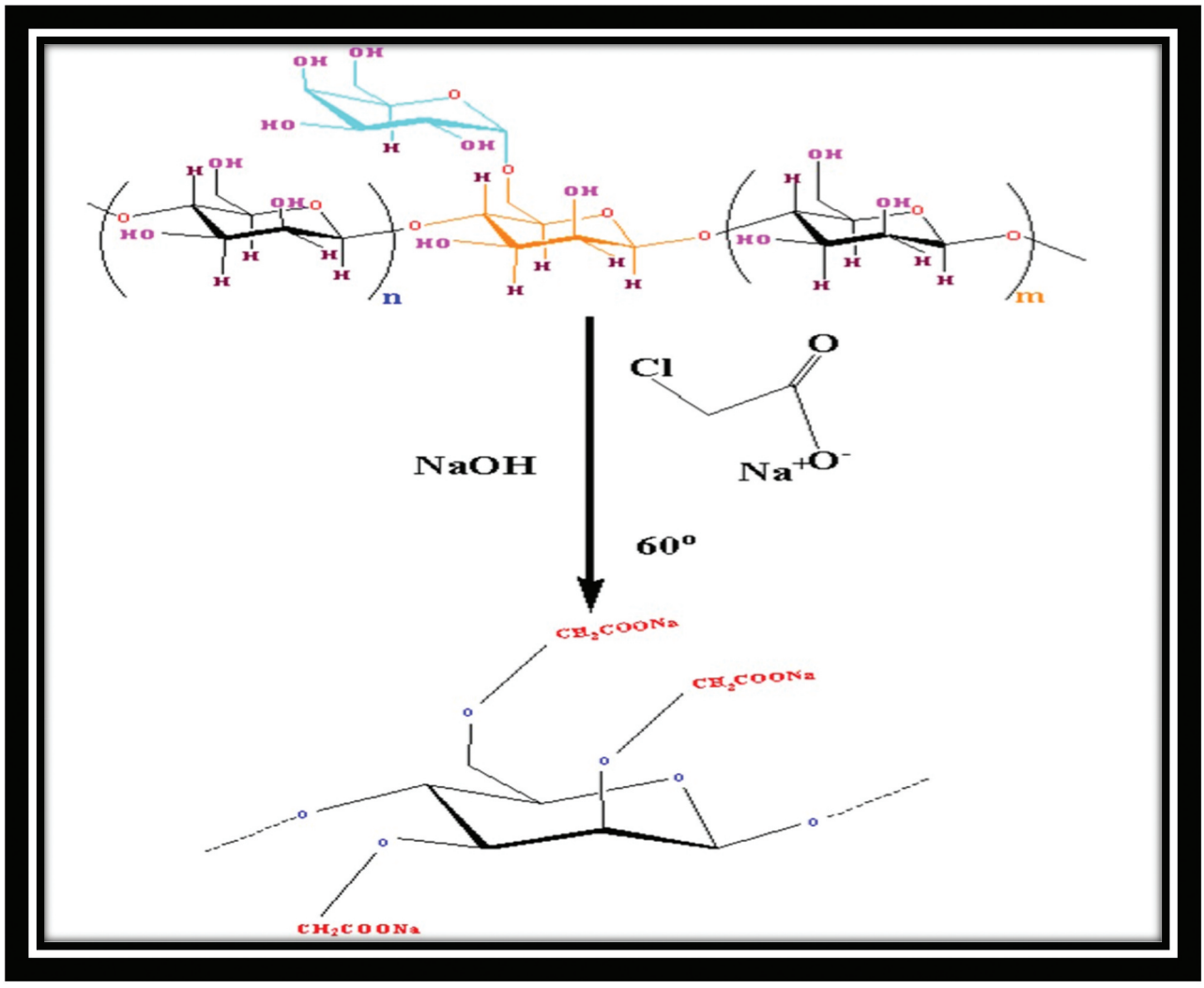


Illustration 11. synthesis of Carboxymethyl Guar Gum.

4.2. Synthesis and applications of carboxymethyl Guar Gum (CMGG)

Carboxymethyl Guar Gum (CMGG) is synthesized through an alkali-catalyzed etherification reaction, wherein Guar Gum undergoes carboxymethylation to introduce carboxymethyl groups into its molecular structure.[108] This process typically employs sodium hydroxide (NaOH) as a catalyst and monochloroacetic acid (MCA) as the carboxymethylating agent. Initially, Guar Gum is dissolved in an aqueous alkaline solution, which activates its hydroxyl groups, making them more reactive. The gradual addition of MCA facilitates the substitution of hydroxyl groups with carboxymethyl groups.[95, 108, 127, 128] The reaction is generally carried out at temperatures ranging from 30°C to 60°C for several hours (typically 4–8 hours) to ensure complete modification. Maintaining an optimal molar ratio of 1:2 (Guar Gum to MCA) and pH control is crucial for efficient etherification and achieving the desired degree of substitution[95] (Illustration (11)). Characterization techniques such as Nuclear Magnetic Resonance (NMR) spectroscopy, Fourier Transform Infrared (FTIR), Scanning Electron Microscopy (SEM), and rheological measurements confirm the successful synthesis of CMGG and its structural modifications.[123]

CMGG finds applications across diverse industries due to its biocompatibility, non-toxicity, and functional properties (Illustration (12)). In pharmaceuticals, it serves as an excipient in drug formulations, particularly for controlled-release systems that enhance therapeutic efficacy.[119] In the food industry, CMGG acts as a thickener, gelling agent, and stabilizer, improving texture and extending shelf life in products like sauces, dressings, and dairy items.[95,98] Its anionic nature ensures compatibility with various ingredients, providing stability across

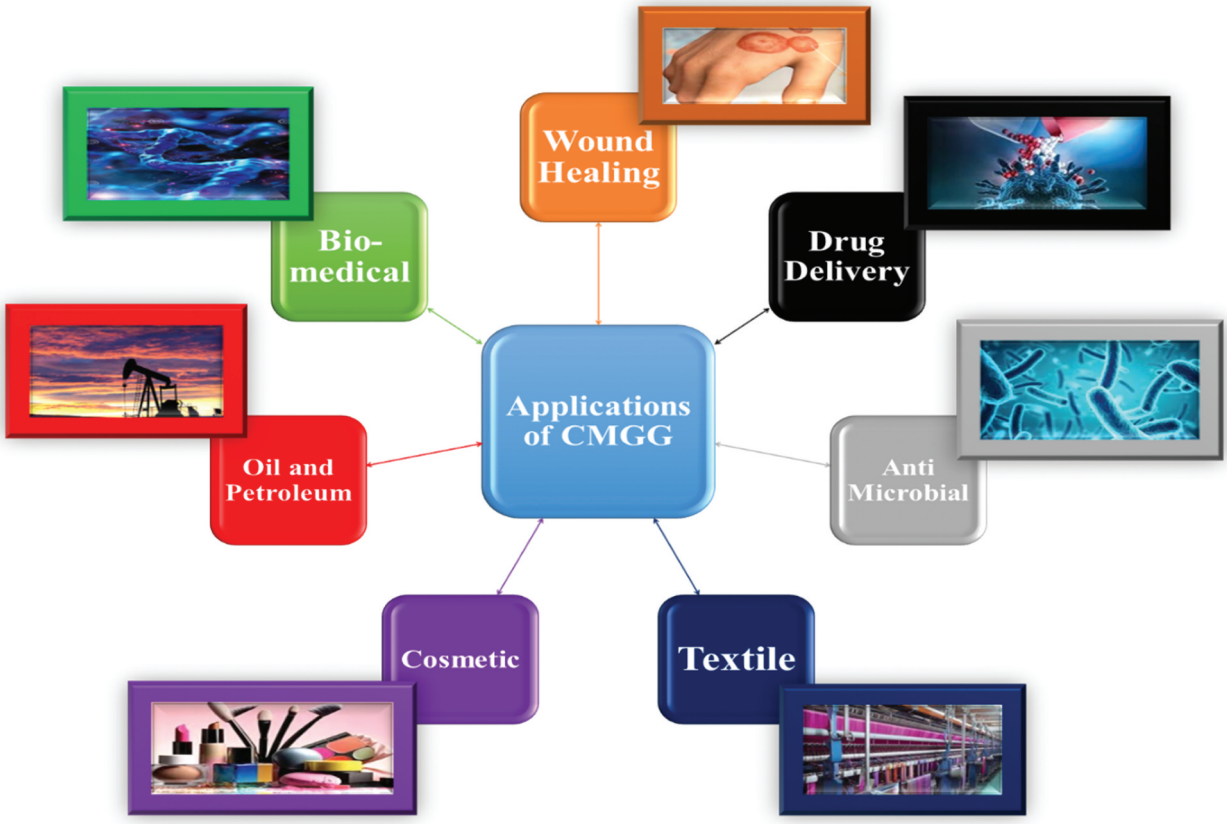


Illustration 12. applications of Carboxymethyl Guar Gum.

different pH levels. In cosmetics and personal care, CMGG contributes to emulsification and thickening in creams, lotions, and shampoos, enhancing product consistency. Industrially, it functions as a fluid loss control agent in oil drilling and as a thickener for textile printing inks, showcasing its versatility.[106]

The multifunctional capabilities of CMGG underscore its significance as an innovative biopolymer, facilitating advancements across multiple fields. Its potential in biomedical applications, including drug delivery and wound healing, highlights its relevance in healthcare. Additionally, its role in nanocomposite development, self-healing materials, and conductive applications demonstrates its versatility in material sciences. By bridging the gap between traditional biopolymers and advanced functional materials, CMGG continues to emerge as a key player in sustainable and high-performance material innovations.

4.2.1. Applications of carboxymethyl Guar Gum (CMGG) with Real-world Examples

CMGG finds applications across diverse industries due to its biocompatibility, non-toxicity, and functional properties.

- **Pharmaceutical Industry**: CMGG is widely utilized in drug formulations, particularly in controlled-release systems that enhance therapeutic efficacy. A recent study demonstrated its role in **nanoparticle-based drug carriers** for targeted cancer therapy, ensuring sustained drug release with minimal side effects.[95]
- **Food Industry**: CMGG acts as a thickener, gelling agent, and stabilizer, improving texture and extending shelf life in products like sauces, dressings, and dairy items. A case study from a leading food corporation highlighted its **use in plant-based dairy alternatives**, where it enhances creaminess and mouthfeel while maintaining stability.[118]
- **Cosmetics and Personal Care**: CMGG contributes to emulsification and thickening in creams, lotions, and shampoos. Recent formulations in **anti-aging serums** have utilized CMGG for enhanced moisture retention and prolonged hydration.[106]
- **Oil Drilling and Textile Printing**: Industrially, CMGG functions as a **fluid loss control agent in oil drilling** and as a **thickener in textile printing inks**, showcasing its versatility. A 2023 study confirmed its efficiency in reducing water loss and

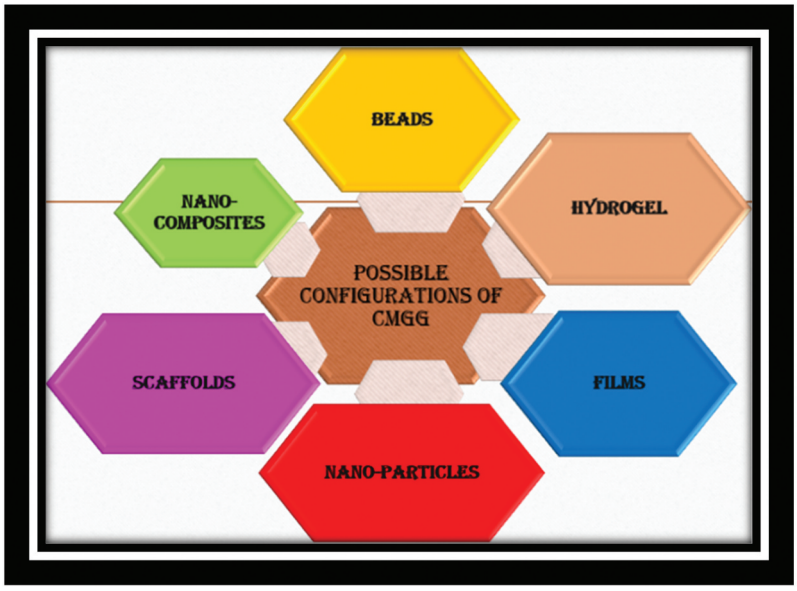


Illustration 13. configurations of Carboxymethyl Guar Gum.

improving drilling fluid viscosity under extreme conditions.

- **Biomedical Applications**: The multifunctional capabilities of CMGG make it an innovative biopolymer in healthcare. **Recent advancements in wound healing** have demonstrated the use of CMGG-based hydrogels in **accelerating tissue regeneration and minimizing infection risks** [129].
- **Nanocomposite Development and Conductive Materials**: CMGG plays a critical role in **self-healing materials and conductive applications**. A study on **CMGG-polyaniline nanocomposites** for wearable electronics showed excellent conductivity and self-repair capabilities.[130]

The ongoing advancements in CMGG research underscore its expanding potential in sustainable and high-performance material innovations. Further studies are investigating its **role in biosensors, smart hydrogels, and biodegradable packaging**, promising groundbreaking applications in various fields.

Table 6 presents various CMGG-based biopolymer-metal matrices, detailing their synthesis methods, characterizations, and applications. These composites exhibit unique properties contributing to developments in biomedical engineering, environmental remediation, and material sciences.

4.3. Characterization techniques for carboxymethyl Guar Gum (CMGG): analysis and insights

Characterization techniques play a crucial role in understanding the structural, thermal, and morphological properties of Carboxymethyl Guar Gum (CMGG). These methods help confirm successful modification, analyze molecular interactions, and assess material performance for various applications. By employing techniques such as FTIR for functional group identification, XRD for crystallinity analysis, and SEM for surface morphology evaluation, researchers can gain comprehensive insights into CMGG's properties. Table 7 summarizes key characterization techniques, detailing their purposes, advantages, and limitations.

The combination of these characterization techniques provides a comprehensive understanding of CMGG's structural, thermal, and morphological properties. By leveraging these analytical methods, researchers can validate chemical modifications, assess material stability, and optimize CMGG for various industrial and biomedical applications. These insights are crucial for tailoring CMGG-based materials to meet the specific requirements of emerging technologies. With a well-characterized profile, CMGG holds significant potential for further advancements in sustainable biopolymer research.

Table 6. Synthesized matrices of CMGG along with their characterizations and preparation methods and applications along with their significance & future impact.

Biopolymer	Matrix	Synthesis methods	Characteristics	Applications	SIZE	Significance & future impact	Reference
CMGG/STMP	Nanoparticles	Ionic Gelation	FTIR, NMR, TGA, DLS, SEM	Drug delivery(nanocarriers)	12-30nm	Enhances drug stability and controlled release, making it a potential carrier for targeted drug delivery.	[131]
CMGG	Nanoparticles	Nanoprecipitation and Sonication	FTIR, SEM, TEM, HRTM	Medical, cosmetic, biosensor, cosmetic, food supplements	12-30nm	Offers biocompatibility and enhanced surface properties, useful in diagnostics and bio-imaging applications.	[128]
CMGG/Citric Acid	Nanocomposite scaffold/ Nanocomposite films	Cross-linking	DSC, ATR-FTIR, DLS, SEM, XRD,	Wound healing		Provides a biodegradable and biocompatible scaffold for tissue regeneration and wound healing applications.	[121]
CMGG/ POLYANILINE	Nanocomposites	Chemical oxidative method	FTIR, XRD, SEM, TEM	Conductive paints	90-280nm	Introduces electrical conductivity, making it suitable for smart coatings and flexible electronics.	[132]
CMGG/PAA-Fe^{3+}	Hydrogels	In-situ radical polymerization	FTIR, UV-Vis	Self-healing in human motion detection		Exhibits self-healing properties, offering potential use in flexible and wearable sensors.	[133]
CMGG/reduced Graphene Oxide	Nanocomposites		SEM, XRD, FTIR	Wound healing		Enhances mechanical strength and antimicrobial properties for biomedical applications.	[134]
CMGG/Ni nps	Ni-polymer composite	Infusion and Reduction	UV, FTIR, XRD, TGA, SEM, TEM,	Catalytic properties and retention of magnetism		Improves catalytic efficiency and magnetic retention, applicable in wastewater treatment and energy storage.	[135]

Table 7. Summary of characterization techniques for CMGG.

Characterization technique	Purpose	Advantages	Disadvantages	References
Fourier Transform Infrared Spectroscopy (FTIR)	Identifies functional groups and confirms chemical modifications	Non-destructive, fast, requires minimal sample	Limited quantitative analysis	[136]
Nuclear Magnetic Resonance (NMR) Spectroscopy	Determines molecular structure and degree of substitution	High accuracy, provides detailed structural information	Expensive, requires specialized expertise	[137]
X-ray Diffraction (XRD)	Analyses crystallinity and phase structure	Provides information on crystal structure	Limited for amorphous materials	[138]
Scanning Electron Microscopy (SEM)	Examines surface morphology and particle size	High-resolution imaging	Requires conductive coating, vacuum environment	[139]
Transmission Electron Microscopy (TEM)	Provides detailed internal structure at the nanoscale	High magnification and resolution	Expensive, complex sample preparation	[140]
Thermogravimetric Analysis (TGA)	Measures thermal stability and decomposition profile	Provides quantitative thermal degradation data	Limited information on chemical composition	[141]
Differential Scanning Calorimetry (DSC)	Determines glass transition and melting points	Useful for studying thermal properties	Requires precise sample preparation	[142]
Dynamic Light Scattering (DLS)	Measures particle size distribution in dispersions	Rapid analysis, non-invasive	Less effective for polydisperse samples	[143]
UV-Vis Spectroscopy	Evaluate optical properties and confirm nanoparticle formation	Simple, cost-effective	Limited structural information	[144]

4.4. Possible configurations of carboxymethyl Guar Gum (CMGG)

Carboxymethyl Guar Gum (CMGG) exhibits a range of possible configurations due to the introduction of carboxymethyl groups into its Guar Gum backbone. These modifications enhance its structural flexibility, allowing for various conformations that depend on environmental conditions such as pH, ionic strength, and temperature.[105] The presence of carboxymethyl groups adds anionic charges to the polymer chain, resulting in an extended configuration in neutral or basic conditions, where electrostatic repulsion between negatively charged groups promotes chain expansion. This leads to higher solubility and increased viscosity in aqueous solutions. In acidic conditions, the anionic charges can be neutralized, allowing the polymer chains to adopt a more compact, coiled structure.[145] Additionally, the degree of substitution of carboxymethyl groups can influence the polymer's behavior, with higher substitution levels leading to more pronounced chain expansion and better interaction with water molecules. CMGG can also form cross-linked networks, particularly in the presence of divalent cations, creating gel-like structures.[146] These configurations, ranging from extended to coiled and from soluble to gelled, enable CMGG to exhibit versatile properties, making it suitable for a wide range of applications, from drug delivery to food thickening and industrial uses. The presence of both carboxyl and hydroxyl groups in the structure of Carboxymethyl Guar Gum (CMGG) attracted significant scientific interest owing to advancements in various properties. CMGG provides a broad spectrum of potential chemical reactions and improved hydrophilicity/water solubility. Carboxymethyl Guar Gum (CMGG) has amassed wide support as a potential functional base for various applications.[44,147] Equipped with increased stability and water solubility, CMGG opens the door to a wide range of additional changes. The domain of the applications of Carboxymethyl Guar Gum can therefore be expanded by blending, grafting, cross-linking, and co-polymerization owing to the presence of carboxyl and hydroxyl groups in its backbone. Fabrication and synthesis of various morphology of CMGG are possible, like scaffolds, films, beads, nanoparticles, and nanocomposites presented in Illustration (13).

Carboxymethyl Guar Gum (CMGG) exhibits remarkable versatility in its configurations, allowing it to serve a variety of functions across multiple industries. It can be processed into beads, which are commonly used in drug delivery systems to provide controlled release of therapeutic agents. Hydrogels made from CMGG offer excellent moisture retention and biocompatibility, making them ideal for applications in wound healing and tissue engineering.[145] Additionally, CMGG forms films that are utilized in biodegradable packaging and as protective layers in wound care products.[121] In the field of nanotechnology, CMGG-based nanoparticles enable targeted drug delivery and enhanced biosensing capabilities [124], [131]. Moreover, scaffolds created from CMGG play a crucial role in tissue engineering by supporting cellular growth and tissue regeneration. Lastly, CMGG can be incorporated into nanocomposites, enhancing their mechanical, thermal, and antimicrobial properties for use in environmental remediation, packaging, and biomedical applications. This adaptability makes CMGG a valuable biopolymer across a wide range of sectors. Carboxymethyl Guar Gum (CMGG) stands out as a highly adaptable biopolymer, offering enhanced properties compared to native Guar Gum through its chemical modification. Its versatility in forming beads, hydrogels, films, nanoparticles, scaffolds, and nanocomposites has expanded its

applicability across diverse fields such as pharmaceuticals, food, environmental remediation, and biomedicine.[149,150] The improved solubility, viscosity, and gelling capacity of CMGG make it an excellent choice for formulations requiring controlled release, biocompatibility, and structural integrity.[44, 95, 132] As industries increasingly seek sustainable and functional materials, CMGG's ability to enhance product performance while remaining biodegradable and nontoxic positions it as a key player in advancing innovative technologies.

5. Conclusion

Carboxymethyl Guar Gum (CMGG) has established itself as a versatile biopolymeric material with significant potential across biomedical, pharmaceutical, packaging, and environmental applications. This review has elucidated its synthesis, structural modifications, and physicochemical properties, highlighting its superiority over unmodified Guar Gum. The functionalization with carboxymethyl groups enhances aqueous solubility, rheological modifiability, and interfacial interactions, positioning CMGG as a high-performance candidate for next-generation biomaterials. Given the global imperative to transition from synthetic polymers to biodegradable alternatives, CMGG emerges as a strategic solution that aligns with contemporary sustainability paradigms. Its intrinsic biodegradability and biocompatibility mitigate environmental concerns associated with conventional polymers, thereby contributing to a reduced ecological footprint. Furthermore, its integration into advanced functional materials – such as hydrogels, nanocomposites, and bioactive films – underscores its potential in sustainable material science. However, despite its promising attributes, several challenges must be addressed to facilitate its large-scale industrial adoption. Enhancing CMGG's mechanical robustness, optimizing crosslinking strategies, and improving its processability in scalable manufacturing systems remain critical areas of research. Additionally, elucidating its long-term stability, degradation kinetics, and interactions in complex biological and environmental matrices will be pivotal for regulatory approvals and commercialization. The development of hybrid CMGG-based materials with synergistic functionalities – such as stimuli-responsiveness, self-healing capabilities, and multimodal bioactivity – could further elevate its applicability in high-performance domains. Future research should also explore eco-friendly synthesis routes and green crosslinking agents to reinforce its role in circular economy models, minimizing waste generation and microplastic pollution.

Beyond CMGG, the field of biopolymers as a whole faces similar challenges in scalability, cost-effectiveness, and functional adaptability. Despite their environmental advantages, many biopolymers exhibit limitations such as lower mechanical strength, shorter lifespan, and sensitivity to processing conditions. The development of **multifunctional biopolymeric composites**, incorporating nanoscale reinforcements and bio-based crosslinking strategies, could offer a pathway to overcome these limitations. Additionally, research efforts must focus on enhancing **biopolymer processability** using advanced fabrication techniques such as **3D printing, electrospinning, and molecular self-assembly** to create structurally refined and application-specific biomaterials.

Another crucial area of future research is **the development of fully biodegradable and recyclable biopolymeric systems** to mitigate polymeric waste accumulation. The incorporation of **bioinspired design principles**, leveraging molecular interactions seen in natural biopolymers, could lead to the next generation of high-performance sustainable materials. Furthermore, interdisciplinary collaboration between material science, nanotechnology, and biotechnology will be essential for unlocking novel functionalities in CMGG and other biopolymers, particularly in drug delivery, regenerative medicine, and environmental remediation. In essence, CMGG epitomizes the paradigm shift toward bio-based polymers that not only exhibit superior functional efficacy but also align with ecological sustainability. As advancements in polymer engineering, nanotechnology, and biomaterial science continue to accelerate, CMGG holds immense potential to redefine the landscape of sustainable materials. Bridging the gap between renewable biopolymers and technologically sophisticated functional materials, its evolution will be instrumental in fostering the next generation of environmentally benign, high-performance polymeric innovations. Likewise, continued advancements in biopolymer research will play a transformative role in **reducing dependency on petrochemical-derived materials, promoting circular economy models, and revolutionizing material applications across diverse scientific and industrial domains**

Acknowledgments

We gratefully acknowledge the invaluable contribution of our colleagues and mentors, who provided guidance and support throughout the writing of this review paper. Their insights and encouragement were instrumental in shaping our understanding and refining the content. We are also thankful to all researchers and authors whose work has been referenced in this review paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study did not receive financial support from any public, commercial, or non-profit funding organizations.

Notes on contributors

Vratika Verma is a dedicated researcher currently pursuing her Ph.D. at the Department of Chemistry, Gurukula Kangri University, Haridwar, Uttarakhand, under the guidance of Dr. Jagram Meena, having commenced her doctoral journey in September 2023. She completed her Master's degree from Hemvati Nandan Bahuguna Garhwal University, Srinagar (Garhwal), Uttarakhand, in 2021, which ignited her passion for research and academia. Her scholarly contributions are primarily focused on biopolymers and their integration into advanced functional materials. She has authored multiple high-impact publications, including the review paper "Carboxymethyl Guar Gum: Bridging the Gap Between Biopolymers and Advanced Functional Materials," which was recently accepted in Polymer-Plastics Technology and Materials. Additionally, her review paper "Studies on Biopolymer-Based Nanocomposites Reinforced with Metallic Nanoparticles" was published in Acta Biology Forum on July 8, 2024. Vratika has also contributed to esteemed book chapters in Industrial Chemistry and Biopolymers. Furthermore, she has co-authored two notable review papers: "A Comprehensive Scrutinization on Tamarind Kernel Powder-Based Derivatives and Nanomaterials in Modern Research" (as third author) published in Nano-Structures and Nano-Objects, and "Polyvinyl Alcohol: Recent Advances and Applications in Sustainable Materials" (as second author), published in Polymer-Plastics Technology and Materials. With a steadfast commitment to scientific innovation, her research endeavors continue to explore the frontiers of biopolymers and nanocomposites, bridging the gap between sustainable materials and technological advancements.

Dr. Jagram Meena is a Ph.D. in Biopolymer Chemistry from Delhi Technological University, Dr. Jagram Meena possesses over a decade of teaching and research experience. Currently an Assistant Professor at Gurukula Kangri University, his expertise lies in synthesizing and characterizing biopolymers, nanoparticles, and their applications in various fields. Dr. Meena's contributions extend beyond academia, as evidenced by his administrative roles and memberships in prestigious organizations. With numerous publications, conference presentations, and participation in faculty development programs, he remains dedicated to advancing knowledge in nanotechnology and biopolymer chemistry. Proficient in handling major instrumentation, Dr. Meena exemplifies a commitment to interdisciplinary research in sustainable materials.

Teena Saini is currently enrolled as a research scholar in the Department of Chemistry at Gurukula Kangri University, Haridwar (Uttarakhand). She has been working towards her Ph.D. under the supervision of Dr. Jagram Meena, since September 2023. She completed her Master's degree in organic chemistry from Hemvati Nandan Bahuguna Garhwal University, Srinagar (Garhwal), Uttarakhand in 2021. Before her doctoral studies, she gained practical experience at Patanjali Organic Research Institute Pvt. Ltd. Her current research endeavors are centered around Biopolymers.

Sakshi Saini is a research scholar in the Department of Chemistry at Gurukula Kangri University, Haridwar. She is pursuing her Ph.D. under the supervision of Dr. Jagram Meena since Sep. 2023. She has completed her Master's degree in Industrial Chemistry from Gurukula Kangri University, Haridwar,2022. During her tenure in the oil and soap industry, she honed her skills in quality assurance methodologies. She is working in Biochemistry and focuses on Natural Biopolymers. Her research interests include Nanomaterials of Biopolymers with a particular focus on Tamarind Kernel powder.

Dr. Rajdeep Malik is an accomplished assistant professor in the Department of Chemistry at Gurukula Kangri (Deemed a university student), Haridwar. With a Ph.D. in Chemistry completed in 2008, he has been dedicated to academia and research for over a decade. His expertise spans Physical Chemistry, Polymer Chemistry, Inorganic Chemistry, Electrochemistry, and Nanomaterials. Dr. Malik possesses extensive experience in handling advanced physioanalytical techniques, including UV-visible spectrophotometry, GC-MS, potentiometry, and IR spectroscopy. He has contributed significantly to scientific literature, with several research papers and conference presentations to his credit. Passionate about advancing chemical sciences, he continues to mentor students and contribute to innovative research in his field.

Data availability statement

All the data supporting this review are present with the authors and will be provided in case a request is made.

Ethical approval

All authors have agreed with the manuscript's content, and there is no financial interest in the report. We certify that the submission is original work and is not under review at any other publication.

References

[1] Botvin, V.; Karaseva, S.; Salikova, D.; Dusselier, M. Syntheses and Chemical Transformations of Glycolide and Lactide as Monomers for Biodegradable Polymers. *Polym. Degrad. Stab.* 2021, *183*, 109427. DOI: 10.1016/j.polymdegradstab.2020.109427.

[2] Zhao, M.; Zhao, C.; Han, C.; Zhao, X.; Wu, Y. Ozone Resistance of Three Natural Antioxidants in Solution Polymerized Styrene-Butadiene Rubber—Molecular Simulation and Experimental Study. *Polym. Degrad. Stab.* 2024, *222*, 110711. DOI: 10.1016/j.polymdegradstab.2024.110711.

[3] Hartman, S.; Sasidharan, R.; Voesenek, L. A. C. J. The Role of Ethylene in Metabolic Acclimations to Low

Oxygen. *New Phytol.* 2020, *229*(1), 64–70. Portico. 10.1111/nph.16378.

[4] Hossain, M. T.; Shahid, M. A.; Mahmud, N.; Habib, A.; Rana, M. M.; Khan, S. A.; Hossain, M. D. Research and Application of Polypropylene: A Review. *Discover Nano* 2024, *19*(1). DOI: 10.1186/s11671-023-03952-z.

[5] Zhang, J.; Fan, H.; Bai, J.; Liu, X.; Ding, Y.; Yang, M. A Novel Polyhedral Oligomeric Silsesquioxane Antioxidant Based on Amide-Linked Hindered Phenols and Its Anti-Oxidative Behavior in Polyamide 6,6. *Polym. Degrad. Stab.* 2024, *229*, 110939. DOI: 10.1016/j.polymdegradstab.2024.110939.

[6] Rijavec, T.; Strlič, M.; Kralj Cigić, I. Damage Function for Poly(vinyl Chloride) in Heritage Collections. *Polym. Degrad. Stab.* 2023, *211*, 110329. DOI: 10.1016/j.polymdegradstab.2023.110329.

[7] Benalaya, I.; Alves, G.; Lopes, J.; Silva, L. R. A Review of Natural Polysaccharides: Sources, Characteristics, Properties, Food, and Pharmaceutical Applications. *Int. J. Mol. Sci.* 2024, *25*(2), 1322. DOI: 10.3390/ijms25021322.

[8] Marcioni, M.; Maddalena, L.; Avolio, R.; Castaldo, R.; Gentile, G.; Carosio, F. Layer-By-Layer Coated Cellulose Reduces the Fire Risk of Polyurethane Foam Biocomposites. *Polym. Degrad. Stab.* 2024, *228*, 110935. DOI: 10.1016/j.polymdegradstab.2024.110935.

[9] Wolf, P.; Reimer, M.; Maier, M.; Zollfrank, C. Biodegradation of Polysaccharides, Polyesters, and Proteins in Soil Based on the Determination of Produced Carbon Dioxide. *Polym. Degrad. Stab.* 2023, *217*, 110538. DOI: 10.1016/j.polymdegradstab.2023.110538.

[10] Thbayh, D. K.; Fiser, B. Computational Study of Synthetic and Natural Polymer Additives — the Antioxidant Potential of BHA, TBHQ, BHT, and Curcumin. *Polym. Degrad. Stab.* 2022, *201*, 109979. DOI: 10.1016/j.polymdegradstab.2022.109979.

[11] AL-Oqla, F. M.; Hayajneh, M. T.; Nawafleh, N. Advanced Synthetic and Biobased Composite Materials in Sustainable Applications: A Comprehensive Review. *Emergent Mater.* 2023, *6*(3), 809–826. DOI: 10.1007/s42247-023-00478-z.

[12] Gu, J.-D.; Wu, E. K. W. Biodegradability of Synthetic Plastics and Polymeric Materials: An Illusion or Reality in Waste Managements? *Appl. Env. Biotechnol.* 2021, *5*(2), 9–27. DOI: 10.26789/aeb.2020.02.003.

[13] Gowthaman, N. S. K.; Lim, H. N.; Sreeraj, T. R.; Amalraj, A.; Gopi, S. Advantages of Biopolymers Over Synthetic Polymers. *Biopolymers Their Ind. Appl.* 2021, 351–372. DOI: 10.1016/b978-0-12-819240-5.00015-8.

[14] Abedsoltan, H. Applications of Plastics in the Automotive Industry: Current Trends and Future Perspectives. *Polym Eng. Sci.* 2023, *64*(3), 929–950. Portico. 10.1002/pen.26604.

[15] Liu, P.; Shi, Y.; Wu, X.; Wang, H.; Huang, H.; Guo, X.; Gao, S. Review of the Artificially-Accelerated Aging Technology and Ecological Risk of Microplastics. *Sci. Total Environ.* 2021, *768*, 144969. DOI: 10.1016/j.scitotenv.2021.144969.

[16] Yang, Z.; Lü, F.; Zhang, H.; Wang, W.; Shao, L.; Ye, J.; He, P. Is Incineration the Terminator of Plastics and Microplastics? *J. Hazard. Mater.* 2021, *401*, 123429. DOI: 10.1016/j.jhazmat.2020.123429.

[17] Singh, N.; Ogunseitan, O. A.; Wong, M. H.; Tang, Y. Sustainable Materials Alternative to Petrochemical Plastics Pollution: A Review Analysis. *Sustain. Horiz.* 2022, *2*, 100016. DOI: 10.1016/j.horiz.2022.100016.

[18] Dejene, B. K.; Gudayu, A. D. Eco-Friendly Packaging Innovations: Integrating Natural Fibers and ZnO Nanofillers in Polylactic Acid (PLA) Based Green Composites - a Review. *Polym. Plast. Technol. Mater.* 2024, *63*(12), 1645–1681. DOI: 10.1080/25740881.2024.2350686.

[19] Dey, S.; Veerendra, G. T. N.; Babu, P. S. S. A.; Manoj, A. V. P.; Nagarjuna, K. Degradation of Plastics Waste and Its Effects on Biological Ecosystems: A Scientific Analysis and Comprehensive Review. *Biomed. Mater. Devices* 2023, *2*(1), 70–112. DOI: 10.1007/s44174-023-00085-w.

[20] Dejene, B. K.; Geletaw, T. M. Development of Fully Green Composites Utilizing Thermoplastic Starch and Cellulosic Fibers from Agro-Waste: A Critical Review. *Polym. Plast. Technol. Mater.* 2024, *63*(5), 540–569. DOI: 10.1080/25740881.2023.2296649.

[21] Dejene, B. K.; Gudayu, A. D. Achieving Performance and Functionality in PLA-Based Packaging: Insights from Hybrid Composites with False Banana (Enset) Fiber and ZnO Nanoparticles. *J. Thermoplast. Composite Mater.* 2024. DOI: 10.1177/08927057241288125.

[22] Priya, A. K.; Alagumalai, A.; Balaji, D.; Song, H. Bio-Based Agricultural Products: A Sustainable Alternative to Agrochemicals for Promoting a Circular Economy. *RSC Sustainability* 2023, *1*(4), 746–762. DOI: 10.1039/d3su00075c.

[23] Singh, S.; Chunglok, W. Biopolymers Towards Green and Sustainable Development. 2022. DOI: 10.2174/9789815079302122010l.

[24] Kumaran, S. K.; Chopra, M.; Oh, E.; Choi, H.-J. Biopolymers and Natural Polymers. *Polym Sci. Nanotechnol.* 2020, 245–256. DOI: 10.1016/b978-0-12-816806-6.00011-x.

[25] Tang, T.-C.; An, B.; Huang, Y.; Vasikaran, S.; Wang, Y.; Jiang, X.; Lu, T. K.; Zhong, C. Materials Design by Synthetic Biology. *Nat Rev Mater* 2020, *6*(4), 332–350. DOI: 10.1038/s41578-020-00265-w.

[26] Lim, B. K. H.; Thian, E. S. Biodegradation of Polymers in Managing Plastic Waste — a Review. *Sci. Total Environ.* 2022, *813*, 151880. DOI: 10.1016/j.scitotenv.2021.151880.

[27] Šašinková, D.; Serbruyns, L.; Julinová, M.; FayyazBakhsh, A.; De Wilde, B.; Koutný, M. Evaluation of the Biodegradation of Polymeric Materials in the Freshwater Environment—An Attempt to Prolong and Accelerate the Biodegradation Experiment. *Polym. Degrad. Stab.* 2022, *203*, 110085. DOI: 10.1016/j.polymdegradstab.2022.110085.

[28] Kumar, A.; Mishra, R. K.; Verma, K.; Aldosari, S. M.; Maity, C. K.; Verma, S.; Patel, R.; Thakur, V. K. A Comprehensive Review of Various Biopolymer Composites and Their Applications: From Biocompatibility to Self-Healing. *Mater. Today*

Sustainability 2023, *23*, 100431. DOI: 10.1016/j.mtsust.2023.100431.

[29] Román-Doval, R.; Torres-Arellanes, S. P.; Tenorio-Barajas, A. Y.; Gómez-Sánchez, A.; Valencia-Lazcano, A. A. Chitosan: Properties and Its Application in Agriculture in Context of Molecular Weight. *Polymers* 2023, *15*(13), 2867. DOI: 10.3390/polym15132867.

[30] Verma, D. K.; Malik, R.; Meena, J.; Rameshwari, R. Synthesis, Characterization, and Applications of Chitosan Based Metallic Nanoparticles: A Review. *J. Appl. Nat. Sci.* 2021, *13*(2), 544–551. DOI: 10.31018/jans.v13i2.2635.

[31] Lopez Hurtado, P.; Rouilly, A.; Vandenbossche, V.; Raynaud, C. A Review on the Properties of Cellulose Fiber Insulation. *Building Environ.* 2016, *96*, 170–177. DOI: 10.1016/j.buildenv.2015.09.031.

[32] Chandrashekhar Upadhye, M.; Chetan Kuchekar, M.; Revansiddhappa Pujari, R.; Uttam Sable, N. Biopolymers: A Comprehensive Review. *Open Access Res. J. Sci. Technol.* 2022, *4*(1), 013–018. DOI: 10.53022/oarjst.2022.4.1.0070.

[33] Rani, I.; Warkar, S. G.; Kumar, A. Nano ZnO Embedded Poly (Ethylene Glycol) Diacrylate Cross-Linked Carboxymethyl Tamarind Kernel Gum (CMTKG)/Poly (Sodium Acrylate) Composite Hydrogels for Oral Delivery of Ciprofloxacin Drug and Their Antibacterial Properties. *Mater. Today Commun.* 2023, *35*, 105635. DOI: 10.1016/j.mtcomm.2023.105635.

[34] Meena, J.; Warkar, S. G.; Verma, D. K. Carboxymethyl Tamarind Kernel Gum Nanoparticles; as an Antioxidant Activity. *J. New Mater. Electrochem. Syst.* 2023, *26*(3), 145–150. DOI: 10.14447/jnmes.v26i3.a01.

[35] Khushbu, K.; Warkar, S. G. Carboxymethyl Tamarind Kernel Gum Based Controlled Drug Delivery Excipients: A Review. *J. Eng. Res.* 2022. DOI: 10.36909/jer.icapie.15061.

[36] Khushbu, K.; Warkar, S. G. Potential Applications and Various Aspects of Polyfunctional Macromolecule-Carboxymethyl Tamarind Kernel Gum. *Eur. Polym. J.* 2020, *140*, 110042. DOI: 10.1016/j.eurpolymj.2020.110042.

[37] Meena, J.; Chandra, H.; Warkar, S. G. Carboxymethyl Tamarind Kernel Gum/znO- Biocomposite: As an Antifungal and Hazardous Metal Removal Agent. *J. New Mater. Electrochem. Syst.* 2022, *25*(3), 206–213. DOI: 10.14447/jnmes.v25i3.a08.

[38] Compart, J.; Singh, A.; Fettke, J.; Apriyanto, A. Customizing Starch Properties: A Review of Starch Modifications and Their Applications. *Polymers* 2023, *15*(16), 3491. DOI: 10.3390/polym15163491.

[39] Omoregie Egharevba, H. Chemical Properties of Starch and Its Application in the Food Industry; *Chemical Properties of Starch*, 2020. DOI: 10.5772/intechopen.87777.

[40] Dejene, B. K.; Birilie, A. A.; Yizengaw, M. A.; Getahun, S. A. Thermoplastic Starch-ZnO Nanocomposites: A Comprehensive Review of Their Applications in Functional Food Packaging. *Int. J. Biol. Macromol.* 2024, *282*, 137099. DOI: 10.1016/j.ijbiomac.2024.137099.

[41] Kumari, B.; Sit, N. Comprehensive Review on Single and Dual Modification of Starch: Methods, Properties and Applications. *Int. J. Biol. Macromol.* 2023, *253*, 126952. DOI: 10.1016/j.ijbiomac.2023.126952.

[42] Milivojević, M.; Popović, A.; Pajić-Lijaković, I.; Šoštarić, I.; Kolašinac, S.; Stevanović, Z. D. Alginate Gel-Based Carriers for Encapsulation of Carotenoids: On Challenges and Applications. *Gels* 2023, *9*(8), 620. DOI: 10.3390/gels9080620.

[43] Deng, Y.; Yang, N.; Valentine Okoro, O.; Shavandi, A.; Nie, L. Alginate-Based Composite and Its Biomedical Applications. *Prop. Appl. Alginates* 2022. DOI: 10.5772/intechopen.99494.

[44] Jiang, H.; Zhang, W.; Chen, L.; Liu, J.; Cao, J.; Jiang, W. Recent Advances in Guar Gum-Based Films or Coatings: Diverse Property Enhancement Strategies and Applications in Foods. *Food Hydrocolloids* 2023, *136*, 108278. DOI: 10.1016/j.foodhyd.2022.108278.

[45] Prabaharan, M. Prospective of Guar Gum and Its Derivatives as Controlled Drug Delivery Systems. *Int. J. Biol. Macromol.* 2011, *49*(2), 117–124. DOI: 10.1016/j.ijbiomac.2011.04.022.

[46] Wang, S.; He, L.; Guo, J.; Zhao, J.; Tang, H. Intrinsic Viscosity and Rheological Properties of Natural and Substituted Guar Gums in Seawater. *Int. J. Biol. Macromol.* 2015, *76*, 262–268. DOI: 10.1016/j.ijbiomac.2015.03.002.

[47] Singh, R.; Gautam, S.; Sharma, B.; Jain, P.; Chauhan, K. D. Biopolymers and Their Classifications. *Biopolymers Their Ind. Appl.* 2021, 21–44. DOI: 10.1016/b978-0-12-819240-5.00002-x.

[48] Balart, R.; Garcia-Garcia, D.; Fombuena, V.; Quiles-Carrillo, L.; Arrieta, M. P. Biopolymers from Natural Resources. *Polymers* 2021, *13*(15), 2532. DOI: 10.3390/polym13152532.

[49] Pei, Y.; Wang, L.; Tang, K.; Kaplan, D. L. Biopolymer Nanoscale Assemblies as Building Blocks for New Materials: A Review. *Adv. Funct. Mater.* 2021, *31*(15). Portico. 10.1002/adfm.202008552.

[50] Meena, J.; Verma, S. K.; Rameshwari, R.; Verma, D. K. Polyaniline/Carboxymethyl Guar Gum Nanocomposites: As Biodegradable, Conductive Film. *Rasayan J. Chem.* 2022, *15*(2), 1021–1027. DOI: 10.31788/rjc.2022.1526820.

[51] Meena, J.; Jassal, P. S. Phenol Organic Impurity Remove from Pollutants Water by Batch Adsorption Studies Using Magneto Chitosan Nanoparticle. *AIJREAS* 2017, *2*, 2455–6300.

[52] Dziuba, R.; Kucharska, M.; Madej-Kiełbik, L.; Sulak, K.; Wiśniewska-Wrona, M. Biopolymers and Biomaterials for Special Applications within the Context of the Circular Economy. *Materials* 2021, *14*(24), 7704. DOI: 10.3390/ma14247704.

[53] Meena, J.; Ram, J. Fabrication of Biopolymers and Their Use with Metal Zinc Oxide Nanoparticles: A Review. *Acta Biol. Forum* 2023, *V02i01*, 23–32. DOI: 10.5281/zenodo.8056710.

[54] Verma, D. K.; Meena, J.; Verma, S. K. Synthesis of Nickel Nanorods and Their Conversion into Nanoparticles. *Res. J. Chem. Environ.* 2022, *26*(12), 7–12. DOI: 10.25303/2612rjce007012.

[55] Jadon, N.; Bhat, G. A.; Sharma, M. V.; Sharma, H. K. Photocatalytic Degradation of Methyl Orange Dye with Synthesized Chitosan/Fe2O3Nanocomposite and Its Isotherm Studies. *Curr. Nanosci.* 2022, *18*(1), 78–85. DOI: 10.2174/1573413716666201217123318.

[56] Bojorges, H.; López-Rubio, A.; Martínez-Abad, A.; Fabra, M. J. Overview of Alginate Extraction Processes: Impact on Alginate Molecular Structure and Techno-Functional Properties. *Trends Food Sci. Technol.* 2023, *140*, 104142. DOI: 10.1016/j.tifs.2023.104142.

[57] Warkar, S. G. Synthesis and Applications of Biopolymer/FeO Nanocomposites: A Review. *J. New Mater. Electrochem. Syst.* 2022, *25*(1), 7–16. DOI: 10.14447/jnmes.v25i1.a02.

[58] Harugade, A.; Sherje, A. P.; Pethe, A. Chitosan: A Review on Properties, Biological Activities and Recent Progress in Biomedical Applications. *Reactive Funct. Polym.* 2023, *191*, 105634. DOI: 10.1016/j.reactfunctpolym.2023.105634.

[59] Aouadi, A.; Hamada Saud, D.; Rebiai, A.; Achouri, A.; Benabdesselam, S.; Mohamed Abd El-Mordy, F.; Pohl, P.; Ahmad, S. F.; Attia, S. M.; Abulkhair, H. S., et al. Introducing the Antibacterial and Photocatalytic Degradation Potentials of Biosynthesized Chitosan, Chitosan–ZnO, and Chitosan–ZnO/pvp Nanoparticles. *Sci. Rep.* 2024, *14*(1). DOI: 10.1038/s41598-024-65579-z.

[60] Abdelfattah, E. M.; Elzanaty, H.; Elsharkawy, W. B.; Azzam, M. A.; Elqahtani, Z. M.; Alotibi, S.; Alyami, M.; Fahmy, T. Enhancement of the Structure, Thermal, Linear/Nonlinear Optical Properties, and Antibacterial Activity of Poly (Vinyl Alcohol)/Chitosan/ZnO Nanocomposites for Eco-Friendly Applications. *Polymers* 2023, *15*(21), 4282. DOI: 10.3390/polym15214282.

[61] Azad Poshtmakhi, S.; Rad-Moghadam, K.; Hosseinzadeh, H.; Toorchi-Roudsari, S. Finely Dispersed Fe3O4 and Ag Nanoparticles Adhered by Starch Nano-Layers: An Efficient Catalyst for the Synthesis of Pyrano[2,3-D]pyrimidines. *ChemistrySelect* 2023, *8*(45). Portico. 10.1002/slct.202300533.

[62] dos Santos, A. M.; Meneguin, A. B.; Akhter, D. T.; Fletcher, N.; Houston, Z. H.; Bell, C.; Thurecht, K. J.; Gremião, M. P. D. Understanding the Role of Colon-Specific Microparticles Based on Retrograded Starch/Pectin in the Delivery of Chitosan Nanoparticles Along the Gastrointestinal Tract. *Eur. J. Pharm. Biopharmaceutics* 2021, *158*, 371–378. DOI: 10.1016/j.ejpb.2020.12.004.

[63] Gautam, S.; Sharma, B.; Jain, P. Structural Applications of Graphene-Based Biopolymer Nanocomposites. *Compos. Sci. Technol.* 2020, 61–81. DOI: 10.1007/978-981-15-9180-8_4.

[64] Taghiloo, S.; Ghajari, G.; Zand, Z.; Kabiri-Samani, S.; Kabiri, H.; Rajaei, N.; Piri-Gharaghie, T. Designing Alginate/Chitosan Nanoparticles Containing Echinacea angustifolia: A Novel Candidate for Combating Multidrug-Resistant Staphylococcus Aureus. *Chem. Biodivers.* 2023, *20*(7). Portico. 10.1002/cbdv.202201008.

[65] Jafarimanesh, M. A.; Ai, J.; Shojaei, S.; Khonakdar, H. A.; Darbemamieh, G.; Shirian, S. Sustained Release of Valproic Acid Loaded on Chitosan Nanoparticles within Hybrid of Alginate/Chitosan Hydrogel With/Without Stem Cells in Regeneration of Spinal Cord Injury. *Prog Biomater* 2023, *12*(2), 75–86. DOI: 10.1007/s40204-022-00209-3.

[66] Carrasco-Sandoval, J.; Aranda, M.; Henríquez-Aedo, K.; Fernández, M.; López-Rubio, A.; Fabra, M. J. Impact of Molecular Weight and Deacetylation Degree of Chitosan on the Bioaccessibility of Quercetin Encapsulated in Alginate/chitosan-Coated Zein Nanoparticles. *Int. J. Biol. Macromol.* 2023, *242*, 124876. DOI: 10.1016/j.ijbiomac.2023.124876.

[67] Selvaraj, C.; Dinesh, D. C.; Rajaram, K.; Sundaresan, S.; Singh, S. K. Macromolecular Chemistry: An Introduction. *Silico Approaches Macromol. Chem.* 2023, 71–128. DOI: 10.1016/b978-0-323-90995-2.00007-2.

[68] Selvaraj, C.; Dinesh, D. C.; Rajaram, K.; Sundaresan, S.; Singh, S. K. Macromolecular Chemistry: An Introduction. *Silico Approaches Macromol. Chem.* 2023, 71–128. DOI: 10.1016/b978-0-323-90995-2.00007-2.

[69] Cherian, R. M.; Varghese, R. T.; Joy, J.; Jose, C.; Thomas, S. Synthetic Biopolymers. *Appl. Biopolymers Sci. Biotechnol. Eng.* 2024, 31–56. Portico. 10.1002/9781119783473.ch2.

[70] Saharan, B. S.; Kamal, N.; Badoni, P.; Kumar, R.; Saini, M.; Kumar, D.; Sharma, D.; Tyagi, S.; Ranga, P.; Parshad, J., et al. Biopolymer and Polymer Precursor Production by Microorganisms: Applications and Prospects. *J. Chem. Technol. Biotechnol.* 2023, *99*(1), 17–30. Portico. 10.1002/jctb.7524.

[71] Saini, S.; Saini, T.; Verma, V.; Meena, J. Studies on Biopolymer -Based Nanocomposites Reinforced with Metallic Nanoparticles. *Acta Biol. Forum* 2024, *4*(2), 6–15. DOI: 10.51470/abf.2024.4.2.06.

[72] Jeevanandam, J.; Pan, S.; Rodrigues, J.; Elkodous, M. A.; Danquah, M. K. Medical Applications of Biopolymer Nanofibers. *Biomater. Sci.* 2022, *10*(15), 4107–4118. DOI: 10.1039/d2bm00701k.

[73] Hoque, M.; McDonagh, C.; Tiwari, B. K.; Kerry, J. P.; Pathania, S. Effect of High-Pressure Processing on the Packaging Properties of Biopolymer-Based Films: A Review. *Polymers* 2022, *14*(15), 3009. DOI: 10.3390/polym14153009.

[74] Scala-Benuzzi, M. L.; Piguillem Palacios, S. V.; Takara, E. A.; Fernández-Baldo, M. A. Biomaterials and Biopolymers for the Development of Biosensors. *Biomater.-Based Sensors* 2023, 3–24. DOI: 10.1007/978-981-19-8501-0_1.

[75] Malektaj, H.; Drozdov, A. D.; deClaville Christiansen, J. Mechanical Properties of Alginate Hydrogels Cross-Linked with Multivalent Cations. *Polymers* 2023, *15*(14), 3012. DOI: 10.3390/polym15143012.

[76] Mtibe, A.; Motloung, M. P.; Bandyopadhyay, J.; Ray, S. S. Synthetic Biopolymers and Their Composites: Advantages and Limitations—An Overview. *Macromol. Rapid Commun.* 2021, *42*(15). Portico. 10.1002/marc.202100130.

[77] Meena, J.; Jassal, P. S. Cresol and Its Derivative Organic Pollutant Removal from Wastewater by Adsorption of the Magneto Chitosan Nanoparticle. *Int. J. Chem. Stud.* 2017, *5*, 850–854.
[78] Udayakumar, G. P.; Muthusamy, S.; Selvaganesh, B.; Sivarajasekar, N.; Rambabu, K.; Banat, F.; Sivamani, S.; Sivakumar, N.; Hosseini-Bandegharaei, A.; Show, P. L. Biopolymers and Composites: Properties, Characterization and Their Applications in Food, Medical and Pharmaceutical Industries. *J. Environ. Chem. Eng.* 2021, *9*(4), 105322. DOI: 10.1016/j.jece.2021.105322.
[79] Meena, J.; Kumar, M.; Rasool, A.; Krismastuti, F. S. H. Optimizing Antimicrobial Efficacy and Ammonia Sensing in a Novel Carboxymethyl Tamarind Kernel Gum/Fe Nanocomposite. *Sustain. Chem. One World* 2024, *3*, 100010. DOI: 10.1016/j.scowo.2024.100010.
[80] Deshmukh, S.; Kathiresan, M.; Kulandainathan, M. A. A Review on Biopolymer-Derived Electrospun Nanofibers for Biomedical and Antiviral Applications. *Biomater. Sci.* 2022, *10*(16), 4424–4442. DOI: 10.1039/d2bm00820c.
[81] Srivastava, P.; Kumari, B. M.; Rajeswari, S. U.; Meena, J.; Gangopadhyay, S.; Packirisamy, S.; Gupta, S.; Packirisamy, S. A Comprehensive Review of Advancements in Pharmacology and Drug Discovery. *J. Exp. Zool. India* 2024, *27*(2). DOI: 10.51470/JEZ.2024.27.2.2005.
[82] Padil, V. V. T.; Ramesh, M. V.; Cheong, J. Y. A Comprehensive Review on Gum-Based Electrolytes for Energy Applications: Current Status and Future Projection. *Mater. Today Chem.* 2023, *28*, 101373. DOI: 10.1016/j.mtchem.2022.101373.
[83] Palem, R. R.; Shimoga, G.; Kang, T. J.; Lee, S.-H. Fabrication of Multifunctional Guar Gum-Silver Nanocomposite Hydrogels for Biomedical and Environmental Applications. *Int. J. Biol. Macromol.* 2020, *159*, 474–486. DOI: 10.1016/j.ijbiomac.2020.05.041.
[84] Parmar, I.; Tanwar, M.; Gupta, R. K. Curcumin and ZnO-NPs Loaded Xanthan Gum and Hydroxypropyl Guar Gum-Based Hydrogel Film for Food Packaging. *Indian J. Chem. Technol.* 2024. DOI: 10.56042/ijct.v31i4.9102.
[85] Gasti, T.; Hiremani, V. D.; Kesti, S. S.; Vanjeri, V. N.; Goudar, N.; Masti, S. P.; Thimmappa, S. C.; Chougale, R. B. Physicochemical and Antibacterial Evaluation of Poly (Vinyl Alcohol)/Guar Gum/Silver Nanocomposite Films for Food Packaging Applications. *J. Polym. Environ.* 2021, *29*(10), 3347–3363. DOI: 10.1007/s10924-021-02123-4.
[86] Akbarzadeh, A. R.; Esmaeili, M. S.; Parvaz, S.; Nazarizadeh, P.; Malekshah, R. E. Wastewater Purification from Drug (BEN) and Poison (MAL) by Graphene Oxide/Guar Gum/CuO (GOGC): A Novel MD Simulation of Carbon Nanocomposite Architecting for Experimental Contaminants Removal. *Diamond Relat. Mater.* 2023, *136*, 110039. DOI: 10.1016/j.diamond.2023.110039.
[87] Michael, M. P.; Singh, S. K.; Sahini, M. G. Facile Conglomeration of Guar Gum/TiO2/Fe3O4 Composite Materials for Photocatalytic Antimicrobial Activities. *J. Indian Chem. Soc.* 2022, *99*(10), 100688. DOI: 10.1016/j.jics.2022.100688.
[88] Bag, J.; Mukherjee, S.; Ghosh, S. K.; Das, A.; Mukherjee, A.; Sahoo, J. K.; Tung, K. S.; Sahoo, H.; Mishra, M. Fe3O4 Coated Guar Gum Nanoparticles as Non-Genotoxic Materials for Biological Application. *Int. J. Biol. Macromol.* 2020, *165*, 333–345. DOI: 10.1016/j.ijbiomac.2020.09.144.
[89] Krishna, D. V.; Sankar, M. R. Synthesis and Characterization of SiO2 Nanoparticles Reinforced 3D Printable Gelatin/PVA/Guar Gum/Hydroxypropyl Methylcellulose-Based Biocomposite Hydrogel. *Ind. Crops Products* 2024, *218*, 118977. DOI: 10.1016/j.indcrop.2024.118977.
[90] Mandal, S.; Chi, H.; Moss, R. E.; Dhital, P.; Babatunde, E. O.; Gurav, R.; Hwang, S. Seed Gum-Based Polysaccharides Hydrogels for Sustainable Agriculture: A Review. *Int. J. Biol. Macromol.* 2024, *263*, 130339. DOI: 10.1016/j.ijbiomac.2024.130339.
[91] Tahmouzi, S.; Meftahizadeh, H.; Eyshi, S.; Mahmoudzadeh, A.; Alizadeh, B.; Mollakhalili-Meybodi, N.; Hatami, M. Application of Guar (Cyamopsis Tetragonoloba L.) Gum in Food Technologies: A Review of Properties and Mechanisms of Action. *Food Sci. nutr.* 2023, *11*(9), 4869–4897. Portico. 10.1002/fsn3.3383.
[92] Conte, F.; van Buuringen, N.; Voermans, N. C.; Lefeber, D. J. Galactose in Human Metabolism, Glycosylation, and Congenital Metabolic Diseases: Time for a Closer Look. *Biochem. Et Biophys. Acta (BBA) - Gener. Subj.* 2021, *1865*(8), 129898. DOI: 10.1016/j.bbagen.2021.129898.
[93] Sharma, G.; Sharma, S.; Kumar, A.; Al-Muhtaseb, A. H.; Naushad, M.; Ghfar, A. A.; Mola, G. T.; Stadler, F. J. Guar Gum and Its Composites as Potential Materials for Diverse Applications: A Review. *Carbohydr. Polym.* 2018, *199*, 534–545. DOI: 10.1016/j.carbpol.2018.07.053.
[94] Sharahi, M.; Bahrami, S. H.; Karimi, A. A Comprehensive Review on Guar Gum and Its Modified Biopolymers: Their Potential Applications in Tissue Engineering. *Carbohydr. Polym.* 2025, *347*, 122739. DOI: 10.1016/j.carbpol.2024.122739.
[95] Dalei, G.; Das, S. Carboxymethyl Guar Gum: A Review of Synthesis, Properties, and Versatile Applications. *Eur. Polym. J.* 2022, *176*, 111433. DOI: 10.1016/j.eurpolymj.2022.111433.
[96] Chauhan, D.; Kumar, A.; Warkar, S. G. Synthesis, Characterization, and Metal Ions Sensing Applications of Meta-Benziporphodimethene-Embedded Polyacrylamide/Carboxymethyl Guar Gum Polymeric Hydrogels in Water. *Environ. Technol.* 2020, *43*(7), 991–1002. DOI: 10.1080/09593330.2020.1812730.
[97] Singh, D.; Viswakarma, P.; Kumar, B. A Review of Guar Gum Processing, Properties, and Its Food Applications. *Asian J. Multidimension. Res.* 2021, *10*(10), 965–969. DOI: 10.5958/2278-4853.2021.00962.9.
[98] Garg, S. S.; Gupta, J. Guar Gum-Based Nanoformulations: Implications for Improving Drug Delivery. *Int. J. Biol. Macromol.* 2023, *229*, 476–485. DOI: 10.1016/j.ijbiomac.2022.12.271.

[99] Mandal, S.; Hwang, S.; Shi, S. Q. Guar Gum, a Low-Cost Sustainable Biopolymer, for Wastewater Treatment: A Review. *Int. J. Biol. Macromol.* 2023, *226*, 368–382. DOI: 10.1016/j.ijbiomac.2022.12.039.

[100] Verma, D.; Sharma, S. K. Recent Advances in Guar Gum-Based Drug Delivery Systems and Their Administrative Routes. *Int. J. Biol. Macromol.* 2021, *181*, 653–671. DOI: 10.1016/j.ijbiomac.2021.03.087.

[101] Ullah, R.; Shah, L. A.; Khan, M.; Ara, L. Guar Gum Reinforced Conductive Hydrogel for Strain Sensing and Electronic Devices. *Int. J. Biol. Macromol.* 2023, *246*, 125666. DOI: 10.1016/j.ijbiomac.2023.125666.

[102] Jayakody, M. M.; Kaushani, K. G.; Vanniarachchy, M. P. G.; Wijesekara, I. Hydrocolloid and Water-Soluble Polymers Used in the Food Industry and Their Functional Properties: A Review. *Polym. Bull.* 2022, *80*(4), 3585–3610. DOI: 10.1007/s00289-022-04264-5.

[103] Grządka, E.; Matusiak, J.; Godek, E.; Maciołek, U. Mixtures of Cationic Guar Gum and Anionic Surfactants as Stabilizers of Zirconia Suspensions. *J. Mol. Liq.* 2021, *343*, 117677. DOI: 10.1016/j.molliq.2021.117677.

[104] Herald, C. T. Guar Gum. *Food Hydrocolloids* 2020, 171–184. DOI: 10.1201/9780429290459-10.

[105] Saraf, P.; Montazer, M. Extraction, Structural Properties, and Applications of Guar Gum. *Nat. Gums* 2023, 185–260. DOI: 10.1016/b978-0-323-99468-2.00007-3.

[106] Madni, A.; Khalid, A.; Wahid, F.; Ayub, H.; Khan, R.; Kousar, R. Preparation and Applications of Guar Gum Composites in Biomedical, Pharmaceutical, Food, and Cosmetics Industries. *Curr. Nanosci.* 2021, *17*(3), 365–379. DOI: 10.2174/1573413716999201110142551.

[107] Kaur, S.; Santra, S. Recent Progress in Chemical Modification of the Natural Polysaccharide Guar Gum. *Curr. Org. Synth.* 2022, *19*(2), 197–219. DOI: 10.2174/1570179418666211109105416.

[108] Gong, H.; Liu, M.; Chen, J.; Han, F.; Gao, C.; Zhang, B. Synthesis and Characterization of Carboxymethyl Guar Gum and Rheological Properties of Its Solutions. *Carbohydr. Polym.* 2012, *88*(3), 1015–1022. DOI: 10.1016/j.carbpol.2012.01.057.

[109] Hasan, A. M. A.; Abdel-Raouf, M. E. Applications of Guar Gum and Its Derivatives in the Petroleum Industry: A Review. *Egypt. J. Petroleum* 2018, *27*(4), 1043–1050. DOI: 10.1016/j.ejpe.2018.03.005.

[110] Le, T.-A.; Huynh, T.-P. Current Advances in the Chemical Functionalization and Potential Applications of Guar Gum and Its Derivatives. *Eur. Polym. J.* 2023, *184*, 111852. DOI: 10.1016/j.eurpolymj.2023.111852.

[111] Nsengiyumva, E. M.; Heitz, M. P.; Alexandridis, P. Carboxymethyl Hydroxypropyl Guar Gum Physicochemical Properties in Dilute Aqueous Media. *Int. J. Biol. Macromol.* 2024, *262*, 129775. DOI: 10.1016/j.ijbiomac.2024.129775.

[112] Shi, H.-Y.; Zhang, L.-M. New Grafted Polysaccharides Based on O-Carboxymethyl-O-Hydroxypropyl Guar Gum and N-Isopropyl Acrylamide: Synthesis and Phase Transition Behavior in Aqueous Media. *Carbohydr. Polym.* 2007, *67*(3), 337–342. DOI: 10.1016/j.carbpol.2006.06.005.

[113] Sasi, S.; Anjum, N.; Tripathi, Y. C. Ethnomedicinal, Phytochemical and Pharmacological Aspects of Flacourtia Jangomas: A Review. *Int. J. Pharm. Pharm. Sci.* 2018, *10*(3), 9. DOI: 10.22159/ijpps.2018v10i3.23998.

[114] Shenoy, M. A.; D'Melo, D. J. Synthesis and Characterization of Acryloyloxy Guar Gum. *J. Appl. Polym. Sci.* 2010, *117*(1), 148–154. Portico. 10.1002/app.31872.

[115] Xiao, W.; Dong, L. Novel Excellent Property Film Prepared from Methacryloyl Chloride-Graft-Guar Gum Matrixes. 2011 International Conference on Consumer Electronics, Communications and Networks (CECNet), 2011. DOI: 10.1109/cecnet.2011.5768424.

[116] Singh, M.; Raorane, C. J.; Alka; Shastri, D.; Raj, V.; Kim, S.-C.; Tuteja, M. Recent Progress on Modified Gum Katira Polysaccharides and Their Various Potential Applications. *Polymers* 2022, *14*(17), 3648. DOI: 10.3390/polym14173648.

[117] Abdel-Raouf, M. E.-S.; Sayed, A.; Mostafa, M. Application of Guar Gum and Its Derivatives in Agriculture. *Gums, Resins, Latexes Plant. Origin* 2021, 1–17. DOI: 10.1007/978-3-030-76523-1_5-1.

[118] Hritcu, D.; Dodi, G.; Silion, M.; Popa, N.; Popa, M. I. Composite Magnetic Chitosan Microspheres: In situ Preparation and Characterization. *Polym. Bull.* 2011, *67*(1), 177–186. DOI: 10.1007/s00289-010-0439-y.

[119] Badwaik, H. R.; Kumari, L.; Maiti, S.; Sakure, K.; Ajazuddin; Nakhate, K. T.; Tiwari, V.; Giri, T. K. A Review on Challenges and Issues with Carboxymethylation of Natural Gums: The Widely Used Excipients for Conventional and Novel Dosage Forms. *Int. J. Biol. Macromol.* 2022, *209*, 2197–2212. DOI: 10.1016/j.ijbiomac.2022.04.201.

[120] George, A.; Shah, P. A.; Shrivastav, P. S. Guar Gum: Versatile Natural Polymer for Drug Delivery Applications. *Eur. Polym. J.* 2019, *112*, 722–735. DOI: 10.1016/j.eurpolymj.2018.10.042.

[121] Orsu, P.; Matta, S. Fabrication and Characterization of Carboxymethyl Guar Gum Nanocomposite for Application of Wound Healing. *Int. J. Biol. Macromol.* 2020, *164*, 2267–2276. DOI: 10.1016/j.ijbiomac.2020.07.322.

[122] Mukherjee, B.; Kumari, L.; Ehsan, I.; Ghosh, P.; Banerjee, S.; Chakraborty, S.; Ray, M.; Al Hoque, A.; Sahoo, R. Guar Gum-Based Nanomaterials in Drug Delivery and Biomedical Applications. *Biopolymer-Based Nanomater. Drug. Delivery Biomed. Appl.* 2021, 143–164. DOI: 10.1016/b978-0-12-820874-8.00016-6.

[123] Iqbal, D. N.; Nazir, A.; Iqbal, M.; Yameen, M. Green Synthesis and Characterization of Carboxymethyl Guar Gum: Application in Textile Printing Technology. *Green Process. Synth.* 2020, *9*(1), 212–218. DOI: 10.1515/gps-2020-0022.

[124] Rahmatpour, A.; Alijani, N. An All-Biopolymer Self-Assembling Hydrogel Film Consisting of Chitosan and Carboxymethyl Guar Gum: A Novel

Bio-Based Composite Adsorbent for Cu2+ Adsorption from Aqueous Solution. *Int. J. Biol. Macromol.* 2023, *242*, 124878. DOI: 10.1016/j.ijbiomac.2023.124878.

[125] Khan, M. Chemical and Physical Architecture of Macromolecular Gels for Fracturing Fluid Applications in the Oil and Gas Industry. *Curr. Status, Challenges, Prospects. Gels* 2024, *10*(5), 338. DOI: 10.3390/gels10050338.

[126] Rana, V.; Kamboj, S.; Sharma, R.; Singh, K. Modification of Gums: Synthesis Techniques and Pharmaceutical Benefits. *Handb. Polym. Pharm. Technol.* 2015, 299–364. Portico. 10.1002/9781119041450.ch10.

[127] Adhikary, P.; Krishnamoorthi, S.; Singh, R. P. Synthesis and Characterization of Grafted Carboxymethyl Guar Gum. *J. Appl. Polym. Sci.* 2010, *120*(5), 2621–2626. Portico. 10.1002/app.33471.

[128] Gupta, A. P.; Verma, D. K. Preparation and Characterization of Carboxymethyl Guar Gum Nanoparticles. *Int. J. Biol. Macromol.* 2014, *68*, 247–250. DOI: 10.1016/j.ijbiomac.2014.05.012.

[129] Chen, H.; Kuru, E.; Hu, K. Dynamic Filtration Loss Control Through Optimization of Drilling Fluid Rheological Properties: A Comparative Study of the Fluid Viscoelasticity versus Shear Viscosity Effects. 2024. DOI: 10.2139/ssrn.4765326.

[130] Li, S.; Zhou, X.; Dong, Y.; Li, J. Flexible Self-Repairing Materials for Wearable Sensing Applications: Elastomers and Hydrogels. *Macromol. Rapid Commun.* 2020, *41*(23). Portico. 10.1002/marc.202000444.

[131] Dodi, G.; Pala, A.; Barbu, E.; Peptanariu, D.; Hritcu, D.; Popa, M. I.; Tamba, B. I. Carboxymethyl Guar Gum Nanoparticles for Drug Delivery Applications: Preparation and Preliminary in-Vitro Investigations. *Mater. Sci. Eng.: C* 2016, *63*, 628–636. DOI: 10.1016/j.msec.2016.03.032.

[132] Gupta, A. P.; Verma, D. K. Synthesis and Characterization of Carboxymethyl Guar Gum Nanoparticles Stabilized Polyaniline/Carboxymethyl Guar Gum Nanocomposites. *J Nanostruct Chem.* 2015, *5*(4), 405–412. DOI: 10.1007/s40097-015-0172-z.

[133] Chen, W.; Bu, Y.; Li, D.; Liu, Y.; Chen, G.; Wan, X.; Li, N. Development of High-Strength, Tough, and Self-Healing Carboxymethyl Guar Gum-Based Hydrogels for Human Motion Detection. *J. Mater. Chem. C* 2020, *8*(3), 900–908. DOI: 10.1039/c9tc05797h.

[134] Orsu, P.; Haider, H. Y.; Koyyada, A. Bioengineering for Curcumin-Loaded Carboxymethyl Guargum/Reduced Graphene Oxide Nanocomposites for Chronic Wound Healing Applications. *Int. J. Pharm.* 2021, *606*, 120928. DOI: 10.1016/j.ijpharm.2021.120928.

[135] Sardar, D.; Sengupta, M.; Bordoloi, A.; Ahmed, M. A.; Neogi, S. K.; Bandyopadhyay, S.; Jain, R.; Gopinath, C. S.; Bala, T. Multiple Functionalities of Ni Nanoparticles Embedded in Carboxymethyl Guar Gum Polymer: Catalytic Activity and Superparamagnetism. *Appl. Surf. Sci.* 2017, *405*, 231–239. DOI: 10.1016/j.apsusc.2017.01.229.

[136] Montoya-Escobar, N.; Ospina-Acero, D.; Velásquez-Cock, J. A.; Gómez-Hoyos, C.; Serpa Guerra, A.; Gañan Rojo, P. F.; Vélez Acosta, L. M.; Escobar, J. P.; Correa-Hincapié, N.; Triana-Chávez, O., et al. Use of Fourier Series in X-Ray Diffraction (XRD) Analysis and Fourier-Transform Infrared Spectroscopy (FTIR) for Estimation of Crystallinity in Cellulose from Different Sources. *Polymers* 2022, *14*(23), 5199. DOI: 10.3390/polym14235199.

[137] Gupta, P.; Das, S. S.; Singh, N. B. *Nuclear Magnetic Resonance Spectroscopy*; Spectroscopy, 2023. DOI: 10.1201/9781003412588-2.

[138] Tomaszewski, P. E. Comments on the Paper "How to Calculate Crystallite Size from X-Ray Diffraction (XRD) Using Scherrer Method" by Siti Fatimah Et Al. Published in ASEAN Journal of Science and Engineering 2 (2022) 65. *ASEAN J. Sci. Eng.* 2022, *3* (3), 301–304. DOI: 10.17509/ajse.v3i3.56905.

[139] Mtanis, T.; Biadsee, A.; Ormianer, Z. Assessing the Cleanliness of Dental Implants Using Scanning Electron Microscopy and Energy-Dispersive X-Ray Spectroscopy Analysis—A SEM and EDS in vitro Study. *J. Funct. Biomater.* 2023, *14*(3), 172. DOI: 10.3390/jfb14030172.

[140] Alcorn, F. M.; Jain, P. K.; van der Veen, R. M. Time-Resolved Transmission Electron Microscopy for Nanoscale Chemical Dynamics. *Nat Rev Chem.* 2023, *7*(4), 256–272. DOI: 10.1038/s41570-023-00469-y.

[141] Natali, M.; Torre, L.; Puri, I.; Rallini, M. Thermal Degradation of Phenolics and Their Carbon Fiber Derived Composites: A Feasible Protocol to Assess the Heat Capacity as a Function of Temperature Through the Use of Common DSC and TGA Analysis. *Polym. Degrad. Stab.* 2022, *195*, 109793. DOI: 10.1016/j.polymdegradstab.2021.109793.

[142] Janovszky, D.; Sveda, M.; Sycheva, A.; Kristaly, F.; Zámborszky, F.; Koziel, T.; Bala, P.; Czel, G.; Kaptay, G. Amorphous Alloys and Differential Scanning Calorimetry (DSC). *J. Therm. Anal. Calorim.* 2021, *147*(13), 7141–7157. DOI: 10.1007/s10973-021-11054-0.

[143] Jia, Z.; Li, J.; Gao, L.; Yang, D.; Kanaev, A. Dynamic Light Scattering: A Powerful Tool for in situ Nanoparticle Sizing. *Colloids Interface* 2023, *7*(1), 15. DOI: 10.3390/colloids7010015.

[144] Chirumamilla, P.; Dharavath, S. B.; Taduri, S. Eco-Friendly Green Synthesis of Silver Nanoparticles from Leaf Extract of Solanum Khasianum: Optical Properties and Biological Applications. *Appl. Biochem. Biotechnol.* 2022, *195*(1), 353–368. DOI: 10.1007/s12010-022-04156-4.

[145] Hasan, A. M. A.; Keshawy, M.; Abdel-Raouf, M. E.-S. Atomic Force Microscopy Investigation of Smart Superabsorbent Hydrogels Based on Carboxymethyl Guar Gum: Surface Topography and Swelling Properties. *Mater. Chem. Phys.* 2022, *278*, 125521. DOI: 10.1016/j.matchemphys.2021.125521.

[146] Askari, V. R.; Fadaei, M. S.; Dabbaghi, M. M.; Fadaei, M. R.; Baradaran Rahimi, V. Ionotropical Cross-Linked Carboxymethylated Gums-Based Systems in Drug Delivery. *Ionotropic Cross-Link. Biopolymers* 2024, 245–274. DOI: 10.1016/b978-0-323-96116-5.00022-3.

[147] Raeisi Estabragh, M. A.; Sajadi Bami, M.; Dehghannoudeh, G.; Noudeh, Y. D.; Moghimipour, E. Cellulose Derivatives and Natural Gums as Gelling Agents for Preparation of Emulgel-Based Dosage Forms: A Brief Review. *Int. J. Biol. Macromol.* 2023, *241*, 124538. DOI: 10.1016/j.ijbiomac.2023.124538.

[148] Soubhagya, A. S.; Balagangadharan, K.; Selvamurugan, N.; Prabaharan, M. Porous Wound Dressings Based on Chitosan/Carboxymethyl Guar Gum/TiO2 Nanoparticles. The Vii International Young Researchers' Conference – Physics, Technology, Innovations (Pti-2020), 2020; Vol. 2313, pp 020029. DOI: 10.1063/5.0030433.

[149] Thakur, S. Gum-Based Green Nanocomposites and Their Applications. *Green-Based Nanocomposite Mater. Appl.* 2023, 295–315. DOI: 10.1007/978-3-031-18428-4_15.

[150] Meena, J.; Saini, S. A Novel Acryloyloxy Tamarind Kernel Powder Biocomposite and Enhanced Antibacterial Activity. 2024. DOI: 10.2139/ssrn.4871679.

Comparative analysis of in silico tools for Drug Side effect prediction

Harshita Thakur
Department of Biotechnology
Delhi Technological University
Shahbad Daulatpur, Main Bawana Road,
New Delhi, India
Email: harshitathakur_23bio06@dtu.ac.in

Asmita Das
Department of Biotechnology
Delhi Technological University
Shahbad Daulatpur, Main Bawana Road,
New Delhi, India
Email: asmitadas1710@dce.ac.in

2025 International Conference on Intelligent Control, Computing and Communications (IC3) | 979-8-3315-2749-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/IC363308.2025.10956741

Abstract— In silico tools are being developed and used these days to make predictions in the field of medicine. They are being used in scientific research where they are helping to ease out the process of developing new techniques and products. Cancer research is also being benefitted from these tools. There are special algorithms developed which form basis of these in silico tools. Molecular Docking is one such tool that can be based on principles like Artificial Intelligence, Machine Learning etc. It is used to understand interactions between drugs and how do they impact the body, what makes the difference in the biologic system and how do these compounds actually cause side effects. Here is a review of Molecular Docking tools which can predict the interactions between two compounds on basis of their biding energy and other parameters. These tools will be analyzed on basis of predictions they make and on basis of this analysis, best tool will be used chosen for making adverse reaction analysis of Dabrafenib, which is an anti-cancer drug administered to melanoma patients.

Keywords — molecular docking, autoDock, swissDock, autoDock vina

I. INTRODUCTION

Bioinformatics has contributed significantly in the healthcare and disease management sectors. From identifying genes to deciphering proteins involved in any event, these in silico tools have helped scientists predict unforeseen outcomes, make generous choices and prevent unwanted impacts. Big data has reached the diagnostics as well, where the genomic sequences are analyzed for presence of mutations and other alterations leading to some diseased phenotypes. Cancer research has also been benefitted from advances in these in silico approaches. Multiple studies are being published every year which make use of these techniques and have noteworthy findings have been elucidated from these approaches. AI, ML, ANN etc have been a building block of in silico tools. They help in analyzing data, making relevant predictions and performing complex tasks involving problem solving. In context of immunology, computational vaccines and drugs are a major product of bioinformatics. With this plethora of information and easy accessibility of data and tools, it is possible for young scientists and researchers to design new beneficial compounds and medicines which are comparatively safe and effective. Use of bioinformatics tools significantly reduces the time of development and also helps in easing the process of designing and preparing a drug which otherwise takes nearly two decades just because of the time utilized in the discovery phase. Another possible way to bring bioinformatics in use is to predict the possible adverse effects of the drugs which are already in use. It has been observed that some compounds generate different response in different individuals. Such responses may make the body susceptible to unwanted reactions and may show unexpected symptoms which are often called side effects. One approach to address this issue is use of pharmacogenomic analysis to determine the what actually is causing the body of patient to respond differently. Such reactions are associated with the metabolism and how the drug is broken down in the body. The altered response is associated with different genetic makeup of certain populations which give them the altered phenotype of varying drug metabolism. These "side-effects" are dealt with by altering the dosage of the medications and such warnings are usually labelled on the packaging. Another approach is to find alternatives to those drugs which are causing side effects rendering patients more prone to illness. Cancer therapy often takes a toll on patient, both on physical and psychological aspects. There are multiple drugs and therapies which have been approved for general public use which have shown to cause the unwanted effects in the body. These effects appear due to unsupervised interactions of the administered molecules with various non target components. To address such issues, a different approach has to be used. Identifying such altered responses is possible through databases and various resources available online. Identifying the cause of those unanticipated effects is a challenging task. Computational approach to discover these non-target interactions include pathway- based methods and the chemical structure-based methods. Pathway based methods identify the various interacting molecules for a particular component. Network based tools using pictorial representations of interconnected molecules are applied for such analysis. These interactions can be used to predict the how the binding to multiple sites can lead to activation of some other pathway by influencing a completely different cascade of molecules. Structure based methods are focused on determining the actual binding energies in order to figure out how strongly do the molecules interact. These results help in validating the predicted interactions and hence proposing reasons behind the side effects. Molecular Docking [1] and QSAR [2] are major tools for making such structure-based predictions. Here are some tools and databases which can help in analyzing what are the alternate pathway, a drug is following leading to expression of a varied response in addition to the desired one. It is helpful in deriving safe and effective drugs.

A. Pathway Based Approaches

In order to obtain data pertaining to the side effect of the drug, first understand how does it actually work. For this, thorough information about the pathogenesis of disease and mechanism of action of the drug is a prerequisite. This gives a crucial understanding of the functioning of that molecule and should be retrieved through literature study. Information about anti-cancer drugs is available on National Cancer Institute. A list of approved FDA drugs is available with observed side effects and risks. These drugs can be searched in DrugBank [3] and PubChem where more information regarding the drugs can be obtained. All possible information regarding these

2025 International Conference on Intelligent Control, Computing and Communications (IC3) | 979-8-3315-2749-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/IC363308.2025.10956741

979-8-3315-2749-5/25/$31.00 ©2025 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:43:31 UTC from IEEE Xplore. Restrictions apply.

therapeutic compounds is available on these databases. Next, determine the target sites. These are the sites in the entire body which have shown to interact positively with the ligand alias the drug. Databases like BindingDB [4], Therapeutic Target Database [5], PHAROS [6], DrugBank[3], SwissTarget [7], etc provide information on what targets does a particular compound interact with and what causes off-target binding that is determining the factors possibly responsible for the adverse reaction.

B. Structure Based Approaches

Molecular docking is the most favored computational approach for elucidating the interaction of two molecules. It determines the behavior of small molecules in the target binding site. Conformation and orientations of a ligand are determined and what is the best way possible for a strong binding can be elucidated from docking results. Such orientations are called poses and the best one is opted. The pose with a low energy is selected to be the best pose and it is used for future references. Molecular docking has been used in researches for drug designing and disease management. There are multiple tools and software programs for the implementation of this techniques. AutoDock [8], AMDock [9], MCDock [10], Glide [11], GOLD [12], FRED [13], CHARMM [14], ACID [15], AMBER [16], SwissDock [17,18] etc.

C. Molecular Docking

Molecular Docking has become a crucial component in drug discovery. It has now become comparatively easy to predict the possible interactions that can occur in the body using this method. Also, it has found many applications in the other areas of medicine and research. Repurposing drugs, finding better alternatives, personalized medicine etc are some of the major benefits. Interactions between proteins, protein-ligands, nucleic acids can be analyzed using docking. Various methods like Machine Learning, Deep Learning, Neural Networks have been used to design tools for performing this task. These tools are available on various platforms and can be used as standalone softwares or are subscription based. The main purpose behind using docking is to predict the binding energy of the two molecules. To do this, certain scores are provided based on the bonds provided. All the data regarding structure of the molecule is derived from Protein Data Bank (PDB). The coordinates file of PDB is converted into an actual structure by the software and the binding scores are predicted. Binding energy and RMSD values are used to predict if the participating compound can interact in vivo or not. Molecular docking usually helps in filtering out the undesirable ligands in the process of drug discovery. This is used in virtual screening to reduce the number of ligands in order to screen out the ligand which are less likely to interact with the target. As many tools are available for this process, it is crucial to determine which is handy and more favorable one to use. As molecular docking is performed at the academic level as well, it is important to know what particular tools can be used to make better predictions at a lower cost, in order to make some fruitful research. Here, we present a comparative analysis of some molecular docking tools.

D. In silico molecular docking tools

The major tools for Molecular Docking are AutoDock and AutoDock Vina [19]. They are mainly used for academic purposes. AutoDock4 is available at https://autodock.scripps.edu. It is designed to operate for protein-ligand interaction studies. It is based on Lamarckian Genetic Algorithm. A force field scoring function is used to predict the binding energy. Ligand is optimized for different conformations and each conformation is estimated and interaction energy is predicted. This keeps on improving till a satisfactory score is achieved. AutoDock Vina is an improvement over AutoDock. It is capable of making more accurate and faster predictions as compared to AutoDock. It is based on Local Search Global Optimizer. This tool finds major application in virtual screening. SwissDock (https://www.swissdock.ch/index.php) is a web server- based tool based on Dihedral Space Sampling and EADock. CHARMM energy values are estimated based on grid scores. CB Dock2 [20, 21] is available at https://cadd.labshare.cn/cb-dock2/index.php. It is a protein- ligand docking server which integrates cavity detection, docking and template fitting. It can perform blind docking and has mainly been used for drug discovery. It is based on AutoDock Vina for docking predictions and curvature- based cavity detection for active site prediction.

II. METHODS

To perform such a task, first one should have structures of drug and the receptor in the PDB format. They can be downloaded from Protein Data Bank itself or from PubChem. Along with this, a thorough literature analysis has to be performed to determine the drug targets in the body and possible reasons for the side effects of that particular drug. For performing this, Molecular Docking tools were used to determine the extent of interaction and then they are compared on basis of the output they present. Dabrafenib [22] is a BRAF inhibiting drug which is used for treatment of melanoma patients. It has been observed that it shows side effects like bone marrow depression, hyperglycemia, lymphopenia. To determine how does Dabrafenib cause these side effects, that is, where else does it bind in the body, docking has to be used. Before that, it is important what tool should be used to perform this task so that best predictions are made. Molecular Docking tools Autodock, AutoDock Vina, SwissDock, CB-Dock2 were used to perform docking analysis. Different tools make use of different methods in predicting binding. Docking was performed using these tools and results were analyze to determine a suitable tool to perform this crucial task of understanding the adverse reactions of Dabrafenib. Blind docking was performed using AutoDock to compare its results with the specific docking.

III. RESULT

Molecular Docking was performed using Autodock, AutoDock Vina, SwissDock, CB-Dock2. AutodockVina is a command line- based tool which can be used for precise docking. AutoDock can perform docking to a large extent. It can be used for both precise and blind docking but the blind docking results of AutoDock vary from the precise docking in terms of the RMSD values. CB-Dock2 and SwissDock are web server- based docking tools which predict docking poses and binding energy. They first determine the pockets and then perform docking generating comparable results. Comparative outcomes are predicted by all these tools. Also, other tools like the visualization tools, used in docking were also compared for quality and ease of use. As seen in Fig.1, RMSD values predicted by these tools are presented. It can be seen that specific docking results generated by AutoDock Vina are more precise and accurate. Fig.2 focusses on the binding energy comparison. CB- Doc2 makes one unusual prediction

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:43:31 UTC from IEEE Xplore. Restrictions apply.

making it less reliable to use. Though these tools make use of different parameters, but they interpret the same thing, that is the binding energy.

TABLE I. TABLE Ⅰ AUTODOCK RESULTS

Parameters→ Protein - Ligand ↓	Binding Energy (kcal/mol)	Inhibition Constant	Reference RMSD	STATISTICAL MECHANICAL ANALYSIS			
				Q	A (kcal/mol)	U (kcal/mol)	S (kcal/mol/K)
BRAF-Dabrafenib	-8.26	881.15 nM	29.73	10.13	-1372.02	-7.78	4.58
AKT-Dabrafenib	-8.02	1.32µM	20.74	10.12	-1371.18	-6.94	4.58

TABLE II. VINA RESULTS

Parameters→ Protein-Ligand↓	Binding Affinity (kcal/mol)	Distance from best mode	
		Upper bound RMSD	Lower bound RMSD
BRAF-Dabrafenib	-11.7	0.000	0.000
AKT-Dabrafenib	-13.7	0.000	0.000

TABLE III. CB-DOCK2 RESULTS

Parameters→ Protein-Ligand↓	FitDock score	Cavity volume (Å^3)	Pocket RMSD
BRAF-Dabrafenib	-12.8	1931	0.0
AKT-Dabrafenib	100.4	208	3.15

TABLE IV. SWISSDOCK RESULTS

Parameters → Docking↓	SwissParam	AC Score	RMSD	Polar Energy	Non-Polar Energy
BRAF-Dabrafenib	-8.837	-107.373731	46.4366	24.5571	-67.1287
AKT-Dabrafenib	-9.9326	-121.673935	17.5711	17.5711	-74.4133

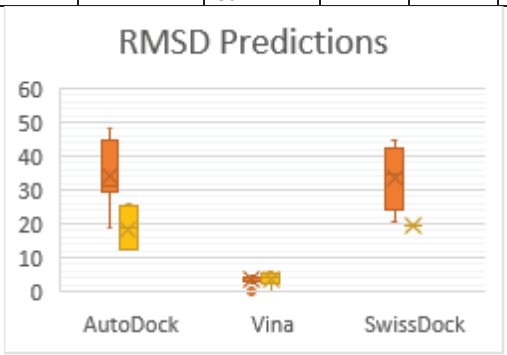


Fig. 1. Comparison of RMSD calculations by various tools

IV. DISCUSSION

To analyze Dabrafenib for its side effects [23], docking has to be performed and for that it is crucial to know which tool will be best suited for this process. For this, Autodock, AutoDock Vina, SwissDock, CB-Dock2 were used to compare and determine the best tool for academic use. Blind docking using AutoDock was compared with targeted docking. It is suggested that the active sites should be known so as to make correct predictions as AutoDock can not make satisfactory predictions in terms of binding site hence generates varying RMSD [24] results. RMSD is a comparative analysis of the poses. It is determined if the given ligand matches the actual ligand in terms of conformation. It was found that targeted docking using AutoDock Vina gave satisfactory results. Also, the binding energy analysis proved that all the four tools were able to generate correct energy relations. AutoDock provides more details in terms of bonds and interactions as compared to other tools. If the user has to visually analyze the interactions, a separate tool has to be used like PyMol, Chimera or Discovery Studio. SwissDock and CB Dock2 on the other hand provide simultaneous visualization. Correct docking sites can be visualized and interacting atoms of the receptor can be determined using both these tools. They make use of different algorithms to function. Major issue with SwissDock is it doesn't let the user decide the binding pocket and hence allows only blind docking in the freely available version. So, the RMSD values vary in the scores predicted by SwissDock. SwissParam scores and AC scores are the output of SwissDock. Attracting cavities and AutoDock Vina are the two methods by which docking was performed. The AC scores [18] are based on CHARMM force fields [25] and FACTS model. SwissParam score is approximation of binding energy and the AC scores are actual energy contributions of the interacting atoms and molecules. CB Dock2 can predict both the binding pockets and the energy functions in one go. The major fact here is that it generates on output only that is it doesn't give multiple poses as output whereas other three tools generate a list of possible poses that can be analyzed for interactions. The FitDock [21] score provided in output is energy parameter which predicts the likelihood of ligand being the match of selected receptor site. The RMSD values generated here are also comparable to that produced by other tools. All these tools when used for blind docking generate different results. PyMol, CHIMERA, Discovery Studio were analyzed to pick the best one out for publication quality images and visualization. Discovery Studio is the most suitable tools for such task as it allows a high-quality image saving, better visualization and molecule preparation. One can determine active sites in this software and this information can be used for targeted docking in any other tool. This package offers all those features making it the best pick.

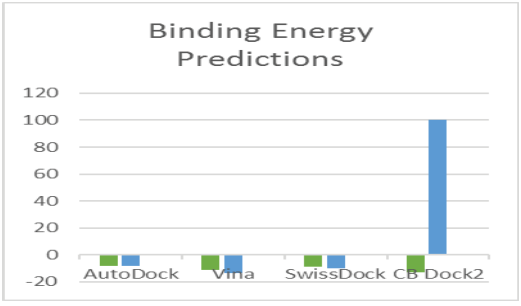


Fig. 2. Comparison of predicted binding energies

V. VCONCLUSION

Melanoma is a very aggressive form of skin cancer and the changing environmental conditions and modern living standards have led to its spread and progression. This study was attempted to understand how do the various tools work and how well are they able to predict the effect of that molecule in the body. It can be determined how does a drug work and what additional impacts does it leave on body at the molecular level. Off-target interactions can be successfully identified and can help in relating the role of Dabrafenib with

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:43:31 UTC from IEEE Xplore. Restrictions apply.

other underlying conditions like Hyperglycemia [26, 27] which can develop in melanoma patients. It has been found that performing blind docking can generate varying outcomes [28]. So, it is suggested to know the binding site of receptor as it generates more useful outcomes. This eliminates the use of SwissDock and CB Dock2 for Side effect analysis. AutoDock can provide sufficient results when used for targeted docking. Taking into consideration, the ease of access and usage, it is suggested that AutoDock Vina is the best choice as it saves the user from generating multiple files and the system failure which AutoDock often undergoes. Also, it can generate separate files for different poses making them easy to interpret. The results generated in Vina can be opened in a visualization softwares which can predict the structural features of binding molecules. Out of available softwares, BIOVIA Discovery Studio can be used as it generates publication quality images in just one click. Receptor and ligand both can be prepared in Discovery Studio and it can also be used for visualization. If suggesting a pipeline for side effect prediction, SIDER [29] can be used to determine side effects caused due to a particular drug, SwissTarget can be used for off-target site analysis, confirm this through literature and then docking can be performed using Vina to compare binding energies of proposed ligand and the actual ligand for that receptor. This study has opened a scope for designing safer drug candidates for melanoma which successfully target only the mutant B-RAF. Similar methods can help in discovering new molecular targets for some other drug as well and can help in deciphering new compounds with potential anti-cancer properties which are safer and easy to use.

REFERENCES

[1] Agu, P.C., Afiukwa, C.A., Orji, O.U. et al., "Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management". Sci Rep 13, 13398 (2023). Available: https://doi.org/10.1038/s41598-023-40160-2.

[2] Tropsha, A., Isayev, O., Varnek, A. et al., "Integrating QSAR modelling and deep learning in drug discovery: the emergence of deep QSAR", Nat Rev Drug Discov 23, 141–155 (2024) Available: https://doi.org/10.1038/s41573-023-00832-0.

[3] David S Wishart, et al., "DrugBank 5.0: a major update to the DrugBank database for 2018", Nucleic Acids Research, Volume 46, Issue D1, 4 January 2018, Pages D1074–D1082, https://doi.org/10.1093/nar/gkx1037.

[4] M.K. Gilson, T. Liu, M. Baitaluk, G. Nicola, L. Hwang, J. Chong, "BindingDB in 2015: A public database for medicinal chemistry, computational chemistry and systems pharmacology", Nucleic Acids Research, Volume 44, Issue D1, 4 January 2016, Pages D1045–D1053, https://doi.org/10.1093/nar/gkv1072.

[5] Y. Zhou, et al., "TTD: Therapeutic Target Database describing target druggability information", Nucleic Acids Research, Volume 52, Issue 5 January 2024, Pages D1465–D1477, Available: https://doi.org/10.1093/nar/gkad751.

[6] Keith J Kelleher et al., "Pharos 2023: an integrated resource for the understudied human proteome", Nucleic Acids Research, Volume 51, Issue D1, 6 January 2023, Pages D1405–D1416, Available: https://doi.org/10.1093/nar/gkac1033.

[7] A.Daina, O. Michielin, V. Zoete, "SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules", Nucleic Acids Research, Volume 47, Issue W1, 02 July 2019, Pages W357–W364,Available: https://doi.org/10.1093/nar/gkz382.

[8] S Forli, R. Huey, M.E.Pique, M.F Sanner, DS Goodsell, AJ Olson, "Computational protein–ligand docking and virtual drug screening with the AutoDock suite", Nat Protoc 11, 905–919 (2016). https://doi.org/10.1038/nprot.2016.051.

[9] Tresanco, M.S., M.E., Valiente, P.A. et al. "AMDock: a versatile graphical tool for assisting molecular docking with Autodock Vina and Autodock4", Biol Direct 15, 12 (2020). Available: https://doi.org/10.1186/s13062-020-00267-2

[10] M Lui, S Wang, "MCDOCK: a Monte Carlo simulation approach to the molecular docking problem", J Comput Aided Mol Des. 1999 Sep;13(5):435-51. doi: 10.1023/a:1008005918983. PMID: 10483527.

[11] Kirkpatrick, P., "Gliding to success", Nat Rev Drug Discov 3, 299 (2004). https://doi.org/10.1038/nrd1364.

[12] J Cole, J. Nissink, R. Taylor, "Protein-Ligand Docking and Virtual Screening with GOLD". 10.1201/9781420028775.ch15 (2005).

[13] M. McGann, "FRED and HYBRID docking performance on standardized datasets", J Comput Aided Mol Des. 2012 Aug;26(8):897-906. doi: 10.1007/s10822-012-9584-8. Epub 2012 Jun 5. PMID: 22669221.

[14] H. Guterres, W. Im, "CHARMM-GUI-Based Induced Fit Docking Workflow to Generate Reliable Protein–Ligand Binding Modes", Journal of Chemical Information and Modeling", J. Chem. Inf. Model. - 4772 - 4779 – 631, American Chemical Society 1549-9596, Available: https://doi.org/10.1021/acs.jcim.3c00416.

[15] F. Wang, et al., "ACID: a free tool for drug repurposing using consensus inverse docking strategy", J Cheminform 11, 73 (2019). https://doi.org/10.1186/s13321-019-0394-z.

[16] G. Heinzelmann, M. Gilson, "Automation of absolute protein-ligand binding free energy calculations for docking refinement and compound evaluation", Sci Rep 11, 1116 (2021). https://doi.org/10.1038/s41598-020-80769-1.

[17] M. Bugnon, et al., "SwissDock 2024: major enhancements for small-molecule docking with Attracting Cavities and AutoDock Vina", Nucleic Acids Research, Volume 52, Issue W1, 5 July 2024, Pages W324–W332, Available: https://doi.org/10.1093/nar/gkae300

[18] Röhrig UF, Goullieux M, Bugnon M, Zoete V., "Attracting Cavities 2.0: improving the flexibility and robustness for small-molecule docking", J. Chem. Inf. Model., 2023.

[19] O. Trott, A. J. Olson, "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading", Journal of Computational Chemistry 31 (2010), 455-461, DOI 10.1002/jcc.21334.

[20] Yang Liu, et al., "CB-Dock2: improved protein-ligand blind docking by integrating cavity detection, docking and homologous template fitting", Nucleic Acids Research, 2022.

[21] Xiaocong Yang, et al., "FitDock: protein-ligand docking by template fitting". Briefings In Bioinformatics, 2022.

[22] N. Garzon-Orjuela, et al., "Efficacy and safety of dabrafenib-trametinib in the treatment of unresectable advanced/metastatic melanoma with BRAF-V600 mutation: A systematic review and network meta-analysis", Dermatol Ther. 2020 Mar;33(2):e13145. doi: 10.1111/dth.13145. Epub 2019 Nov 14. PMID: 31664762.

[23] Ng TSC, et al., "Overcoming differential tumor penetration of BRAF inhibitors using computationally guided combination therapy". Sci Adv. 2022 Apr 29;8(17):eabl6339. doi: 10.1126/sciadv.abl6339. Epub 2022 Apr 29. PMID: 35486732; PMCID: PMC9054019.

[24] A. Castro-Alvarez, AM Costa, J. Vilarrasa, "The Performance of Several Docking Programs at Reproducing Protein-Macrolide-Like Crystal Structures", Molecules. 2017 Jan 17;22(1):136. doi: 10.3390/molecules22010136. PMID: 28106755; PMCID: PMC6155922.

[25] X. Zhu, P.EM Lopes, A.M. MacKerell, "Recent Developments and Applications of the CHARMM force fields. Wiley Interdiscip Rev Comput Mol Sci. 2012 Jan;2(1):167-185. doi: 10.1002/wcms.74. Epub 2011 Jun 28. PMID: 23066428; PMCID: PMC3468154.

[26] Yim C, Mansell K, Hussein N, Arnason T., "Current cancer therapies and their influence on glucose control", World J Diabetes. 2021 Jul 15;12(7):1010-1025. doi: 10.4239/wjd.v12.i7.1010. PMID: 34326951; PMCID: PMC8311484.

[27] J.W. Goldman, M.A. Mendenhall, S.R. Rettinger, "Hyperglycemia Associated With Targeted Oncologic Treatment: Mechanisms and Management", The Oncologist, Volume 21, Issue 11, November 2016, Pages 1326–1336, https://doi.org/10.1634/theoncologist.2015-0519.

[28] D. Ghersi, R. Sanchez, "Improving accuracy and efficiency of blind protein-ligand docking by focusing on predicted binding sites", Proteins. 2009 Feb 1;74(2):417-24. doi: 10.1002/prot.22154. PMID: 18636505; PMCID: PMC2610246.

[29] M. Kuhn, I. Letunic, L.J Jensen, P. Bork, "The SIDER database of drugs and side effects", Nucleic Acids Research, Volume 44, Issue D1, 4 January 2016, Pages D1075–D1079, Available: https://doi.org/10.1093/nar/gkv1075.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:43:31 UTC from IEEE Xplore. Restrictions apply.

2025 First International Conference on Advances in Computer Science, Electrical, Electronics, and Communication Technologies (CE2CT) | 979-8-3315-1857-8/25/$31.00 ©2025 IEEE | DOI: 10.1109/CE2CT64011.2025.10939314

Conventional and Intelligent Control Schemes for Handling A Helicopter's Pitch Axis

Abhishek Chaudhary
Electrical Engineering Department
Delhi Technological University, DTU
New Delhi, India
abhishek@dtu.ac.in

***Abstract* — In this paper, a comparative study of conventional LQR controller, optimized PID controller and intelligent Fuzzy logic controller is demonstrated while achieving the desired pitch axis trajectories of a 2-DOF helicopter system. During this process, TLBO optimized PID controller is used to track down the trajectory angle of pitch axis of helicopter system and this performance is compared with inference rules of fuzzy system and a conventional LQR controller. The membership function parameters, exploration – exploitation scheme of optimization algorithm are discussed while addressing the problem of trajectory tracking. The parameter tuning and stable response of pitch axis of helicopter model is achieved on MATLAB/Simulink platform, on which the comparison is accomplished through time response analysis.**

Keywords— 2-DOF Helicopter; Optimization Control, Unmanned Aerial Vehicle; Fuzzy Logic; Pitch Trajectory: LQR.

I. INTRODUCTION

The issue of trajectory tracking control in unmanned aerial vehicles (UAVs), particularly in helicopter models, has been extensively explored in the literature for many years [1, 2]. A three-degree-of-freedom (3-DOF) helicopter's trajectory tracking can be effectively managed using a combination of two controllers: a feedforward controller for guiding the trajectory and a feedback controller for stabilizing any deviations from the intended path [1]. Experimental studies have shown that these controllers are effective in achieving precise trajectory tracking. Another approach involves dividing the dynamics of a small-scale unmanned helicopter into two subsystems—heading-heave and longitudinal-lateral movements—to focus on heading reference trajectory and inertial position tracking. Feedback loops within these subsystems are used to correct any trajectory errors [2]. Further improvements in attitude and position control have been achieved by integrating the parallel tracking and mapping method with a wireless camera, demonstrating its effectiveness in helicopter control [3]. Despite the range of control strategies available in the literature, the traditional and intelligent controller tuning remains crucial for effective axis control.

In one study, a combination of Linear Quadratic Regulator (LQR) and PID control was employed to reduce tracking errors in pitch and yaw by optimizing a quadratic function with state feedback [4]. Another study introduced a robust adaptive fuzzy controller, developed using a gradient descent algorithm, which further enhanced control accuracy [5]. To address the challenges posed by higher-order flight control actuator failures, a reorganized Fault-Tolerant Control (FTC) system was proposed [6]. Fuzzy logic, since its introduction by Professor Lotfi A. Zadeh in 1965, has become a vital tool in various fields such as pattern recognition, telecommunications, and control systems. By utilizing fuzzy set theory and mathematics, fuzzy logic systems can incorporate human knowledge and intuition through linguistic rules, making them highly effective for trajectory control in helicopter models using Gaussian-type membership functions [7, 8]. Additionally, a dual fuzzy-neural network control system was tested for an intelligent landing system in air transportation, and simulation studies confirmed its effectiveness under varying and unpredictable weather conditions [9].

Continuous monitoring of a helicopter system is essential for real-time optimization of controller parameters. Over the past decade, numerous algorithms have been developed to address this need. Computational intelligence plays a critical role in the real-time tuning of controller parameters for multivariable systems. From classical methods like the Ziegler-Nichols (ZN) method [10] to modern computational algorithms, these approaches are fundamental to optimizing nonlinear systems. Velocity control in a helicopter model is often implemented using a fuzzy-PID controller, with PID auto-tuning achieved through adjustments based on the previous control period [11]. The fuzzy-PID controller has also demonstrated its effectiveness in trajectory control, providing robust and stable responses [12]. An auto-tuned fuzzy-PID controller is designed using a twin-loop strategy [13]. The Teaching-Learning-Based Optimization (TLBO) algorithm, which mimics the learning process in a classroom, has proven to be a versatile tool in various engineering applications for finding both local and global optima [14].

This study is focused on the comparative analysis of three different controllers, used to control the pitch axis of helicopter model. These three controllers are TLBO optimized PID controller, LQR controller and intelligent fuzzy controller. Based on this, the paper is structured as follows; Section I (this section) is focused on the literature survey, Section II is about the demonstration of helicopter model. In section III, the control strategies are discussed at length. In section IV, simulation analysis of results are demonstrated. The work will be concluded with section V.

II. PITCH AXIS IN HELICOPTER

Dynamics of nonlinear equations for helicopter model, as shown in figure 1, This figure shows the physical setup available in the laboratory, which comprises a yoke, two distinct DC motors, and a helicopter-shaped body, which are acquired by using Lagrangian mechanics, are shown for the pitch axis as follows [15]:

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:15:07 UTC from IEEE Xplore. Restrictions apply.

$$(J_{eq,p} + m_{heli}\, l_{cm}^2)\ddot{\theta} = K_{pp}\, V_{m,p} + K_{py}\, V_{m,y} - m_{heli} g l_{cm} \cos\theta - B_p\, \dot{\theta} - m_{heli}\, l_{cm}^2 sin\theta cos\theta \dot{\Psi}^2$$

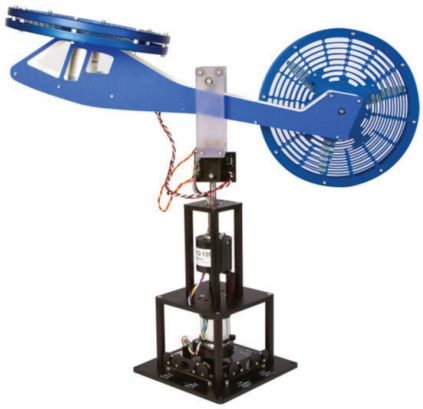


Fig. 1. Laboratory setup of 2-DOF helicopter [15]

The laboratory model showing the horizontal and vertical propellors of helicopter model, as provided by Quanser [15], is depicted in Fig. 2. The vertical motion is the responsive product of vertical propellor itself. One of the motors is designated for controlling the pitch movement, while the other is responsible for yaw movement. To measure these movements, two separate encoders are used—one for pitch and one for yaw. The helicopter system is mounted on a stationary fixed base platform.



Fig. 2. Horizontal and vertical propellors of 2-DOF helicopter [15]

The first DC motor, which operates the front propeller, adjusts the altitude of the helicopter's nose along its pitch axis, thereby controlling the pitch angle, denoted by θ. The pitch encoders are used to measure the corresponding pitch angles. Through a slip-ring mechanism, the measured angle is transmitted as signals to achieve the desired pitch and yaw angles. The nonlinear helicopter system model is subject to the following constraints:

- A pitch angle of zero indicates that the helicopter is in a horizontal position.

- When the helicopter's nose moves vertically upward, the pitch angle increases, leading to a positive pitch thrust force ($F_p > 0$) and a positive pitch angle ($\theta > 0$).

Helicopter model works with the following state space representation:

$$\dot{x} = Ax + Bu \quad (1)$$

Where, state vector x represents:

$$x = [\theta, \Psi, \dot{\theta}, \dot{\Psi}, \dot{\alpha}, \dot{\gamma}]^T \quad (2)$$

and output y is given by:

$$y = Cx + Du \quad (3)$$

III. Control Strategies

A. Fuzzy Logic Controller (FLC)

Fuzzy logic controllers encapsulate human expertise in system management by utilizing membership functions and IF-THEN fuzzy rules. This technique, allowing machines to simulate human reasoning using fuzzy logic, is referred to as a fuzzy logic controller. Fuzzy logic is particularly effective in addressing challenges related to uncertainty, imprecision, and qualitative decision-making. In fuzzy control applications, it is important to note that the control process often involves converting a range of values into a specific point or vice versa. By defining the appropriate membership functions, both inputs and outputs are fuzzified, leading to the derivation of the desired outputs.

The implementation of a fuzzy logic controller involves three fundamental steps: fuzzification, the fuzzy inference process, and defuzzification. Initially, the measured input and its derivative are processed by the fuzzy logic controller, which is connected to the 2-DOF helicopter. The output is then fed back into the input through a closed-loop feedback mechanism within the same unmanned vehicle. For system monitoring, this approach utilizes human-like reasoning to effectively manipulate, represent, and implement the fuzzy logic controller. In the context of helicopter control, a Takagi-Sugeno (T-S) fuzzy model is used, featuring two inputs and one output to manage the pitch angle error. The inputs consist of the pitch error and its derivative, while Gaussian-type membership functions are employed to ensure a smooth, continuously differentiable curve. Seven distinct linguistic variables are used for each input state variable, labeled as NS, ZE, PS, PM, NB, NM, and PB, with an associated rule base.

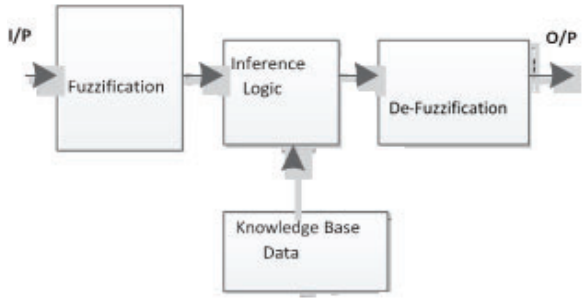


Fig. 3. FLC working Architecture

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:15:07 UTC from IEEE Xplore. Restrictions apply.

B. Proportional Integral Derivative (PID) Controller

To implement a proportional-integral-derivative (PID) controller as a cross-coupled controller on a helicopter model, the MATLAB SIMULINK platform is used once again. The error signal, e(s), is expressed as:

$$e(s) = X_d(s) - X(s) \quad (4)$$

where, X_d (s) represents the desired set points for the pitch control axisband X(s) is the actual process value. The PID controller is applied to this error signal. The set points, error signals e(s) and process values are calculated separately with PID controllers for pitch axis. TLBO algorithm is employed to optimize the parameters of PID controller. While implementing the PID controller in the frequency domain on the helicopter system, the gains $k_{d,heli}$, $k_{i,heli}$ and $k_{p,heli}$, which are the derivative, integral, and proportional gains respectively, are used to control the helicopter system, with the initial operating points given as:

$$\beta_d(s) = \left(k_{d,heli}.s + \frac{k_{i,heli}}{s} + k_{p,heli}\right)e(s) \quad (5)$$

Initial parameters of PID are calculated using ZN method [10]. Then gain parameters are updated through optimization algorithms after each iteration using describing functions.

C. Teaching Learning Based Optimization (TLBO)

TLBO algorithm works in two different phases, with 2nd phase is initiated after the end of 1st. Both phases combines up to complete one iteration in algorithm. These phase are – teaching phase and learning phase.

Teaching phase demonstrates how students in a class learn through their teacher. In this phase, the teacher's goal is to improve the average performance of the students in a specific subject based on the teacher's abilities. Mean marks of any subject (X_{mean}) and the highest marks of class ($X_{j,best,i}$) are used to calculate difference mean as:

$$Dmean_{j,k,i} = r*(X_{best} - T_f * X_{mean}) \quad (6)$$

Where i stand for iteration, k for population size, j is any subject and T_f stand for teaching factor, which is given by:-

T_f = round (1+r), where r is a random number. New solution ($X_{j,k,i}(new)$) is calculated as:

$$X_{j,k,i}(new) = X_{j,k,i}(old) + Dmean_{j,k,i} \quad (7)$$

With this, the teaching phase comes to end and learning phase starts. In learning phase, students are supposed to improve their knowledge through sharing and learning. In here, one students selects a random students from class and improves his/her knowledge through sharing. For a population size of 'n' learners/students, A and B are two randomly selected learners such that $X_{A,i}(new)$ *and* $X_{B,i}(new)$ are the updated solutions for $X_{A,i}(old)$ *and* $X_{B,i}(old)$, which are defined as:

$$X_{A,i}(new) = X_{A,i}(old) + r\,[X_{A,i}(old) - X_{B,i}(old)] \quad (8)$$

$$X_{A,i}(new) = X_{A,i}(old) + r\,[X_{B,i}(old) - X_{A,i}(old)] \quad (9)$$

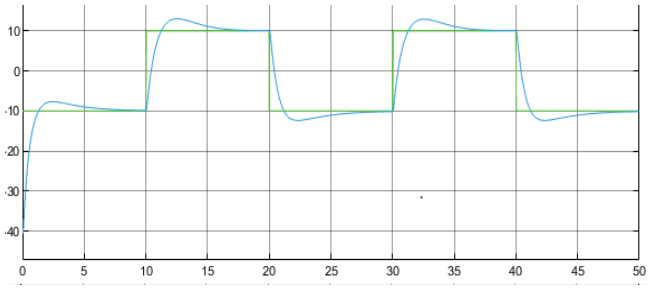
Execution of above equations particularly depends upon the greedy solution conditions. Greedy solution is a criteria of deciding whether the problem is maximization type or minimization type. In pitch axis trajectory problem, aim is to reduce the error to minimum. This makes current problem as minimization problem. Hence, the greedy solution is followed considering the aim as minimization of error problem.

Time response parameters are taken as describing function for the pitch axis problem. After the completion of learning phase, first iteration comes to an end. This process continues until the minimization condition satisfies and the final gain parameters after the completion of all iteration, gives optimized results for pitch axis problem.

IV. Simulation and Results

In this study, the 2-DOF helicopter system, which was initially controlled using the default Linear Quadratic Regulator (LQR) provided by Quanser [15], is subsequently managed using two alternative strategies. The performance of the default LQR controller is graphically depicted in Figure 4. Following this, a PID controller, optimized through the Teaching-Learning-Based Optimization (TLBO) algorithm, and a fuzzy logic controller are implemented to track a predefined reference trajectory for the pitch axis. The system, commonly used as an experimental setup in laboratory research, consists of a fixed base with two separate propellers. The base facilitates movement along its center, with each propeller driven by a DC motor mounted at opposite ends of the system. The primary function of these propellers is to regulate the pitch angle's elevation trajectory. Specifically, the rear propeller controls the yaw angle, while the front propeller manages the elevation angle, thereby ensuring stable lateral motion. A key control challenge is mitigating the cross-coupling effects between the pitch and yaw movements. The real-time dynamics of this system are simulated using MATLAB software, and the results are thoroughly analyzed in this section.

The 2-DOF helicopter represents a standard MIMO (Multiple Input Multiple Output) nonlinear system, with individual controllers assigned to manage the pitch and yaw axes. For the implementation of the Fuzzy Logic Controller (FLC), the pitch axis controller is provided with two inputs: the error (e(s)), which is the difference between the desired and actual pitch angles (measured in radians), and the derivative of this error. Each controller in this setup operates with these two inputs, which are fuzzified using seven distinct Gaussian membership functions. These membership functions are distributed over closed intervals for the fuzzy logic controllers.



Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:15:07 UTC from IEEE Xplore. Restrictions apply.

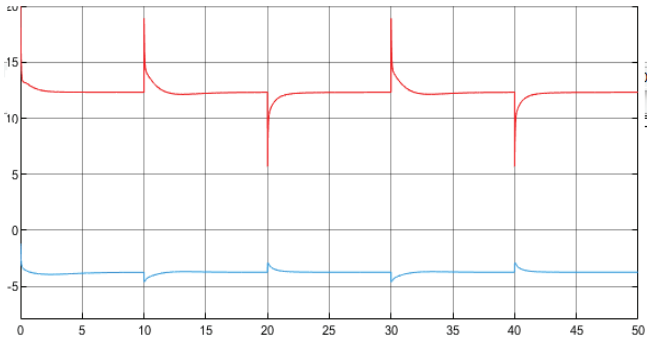


Fig. 4. Pitch angle variations (amplitude vs time) and voltage variations (volt. vs time) on 2DOF helicopter model after using LQR controller

The input and output fuzzy subsets are categorized into linguistic values: Positive Big (PB), Negative Big (NB), Positive Medium (PM), Negative Medium (NM), Positive Low (PL), Negative Low (NL), and Zero (ZE). These categories are utilized in a set of 49 fuzzy rules based on IF-THEN logic, forming a knowledge base that mimics human reasoning, which is essential for the effective operation of the FLC. The degree of membership for these fuzzy sets ranges from 0 to 1, with 0 indicating complete inconsistency and 1 representing full consistency with the fuzzy levels. The input signals for the fuzzy logic controller are the errors in pitch angle, which are the differences between the desired and actual angles. "Negative Big (NB)" represents a large negative velocity, while "Positive Big (PB)" indicates a large positive velocity. The rules function as guidelines to return to the desired position, with the proportional coefficients managed according to the membership function logic. The linguistic control strategy progressively transforms into an automated control strategy. This involves applying experimental knowledge to implement fuzzy logic rules with different membership functions for controlling the yaw and pitch angles. Variations observed in pitch angle with square reference trajectory and voltage variations observed after this control strategy are shown in figure 5. The green signal is reference trajectory and the red signal is followed output trajectory. The improved time response is visibly observed after using FLC controller as compared to default LQR controller. This nonlinear system is once again tested using a third controller for the tracing of pitch reference trajectory, further in this section.

The PID controller is further employed to manage pitch angle variations in the helicopter model, utilizing the Teaching-Learning-Based Optimization (TLBO) algorithm. Figure 6 presents the block diagram used to execute this optimization strategy. The gain parameters depicted in this figure are tuned using the TLBO algorithm and are refined after each iteration of the teaching and learning phases, as discussed in the previous section of this work. The iteration by iteration improvement in tuned gain parameters are observed and the final tuned parameter after the completion of all the iterations are taken as optimized result.

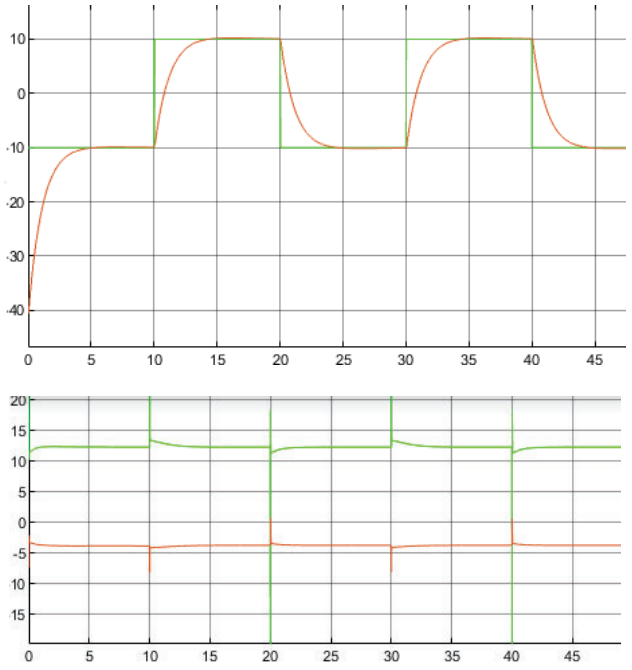


Fig. 5. Pitch angle variations (amplitude vs time) and voltage variations (volt. vs time) on 2DOF helicopter model after using FLC controller

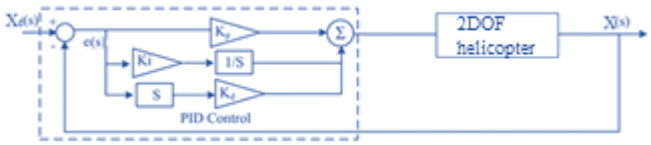


Fig. 6. Block diagram representing the implementation of PID controller on helicopter model.

The control model for the helicopter system's cross-coupled pitch axis, as shown in Figure 6, is specifically designed to ensure stable trajectory tracking. The controller's parameters are optimized using the Teaching-Learning-Based Optimization (TLBO) algorithm. In scenarios with multiple constraints, techniques like TLBO tend to provide near-optimal solutions; however, small adjustments in algorithm parameters can significantly influence their effectiveness. This consideration is particularly relevant in the present study, where TLBO is preferred over other metaheuristic algorithms due to its robustness. A key advantage of this algorithm is that its performance on the helicopter model remains independent of specific parameter settings, contributing to a more stable control environment for the 2-DOF helicopter system's cross-coupled design. The optimized performance of the helicopter model is depicted in Figure 7, where the pink line indicates the square reference trajectory, and the blue line shows the actual tracking achieved. A visual comparison highlights the improved performance compared to the default Linear Quadratic Regulator (LQR) control strategy.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:15:07 UTC from IEEE Xplore. Restrictions apply.

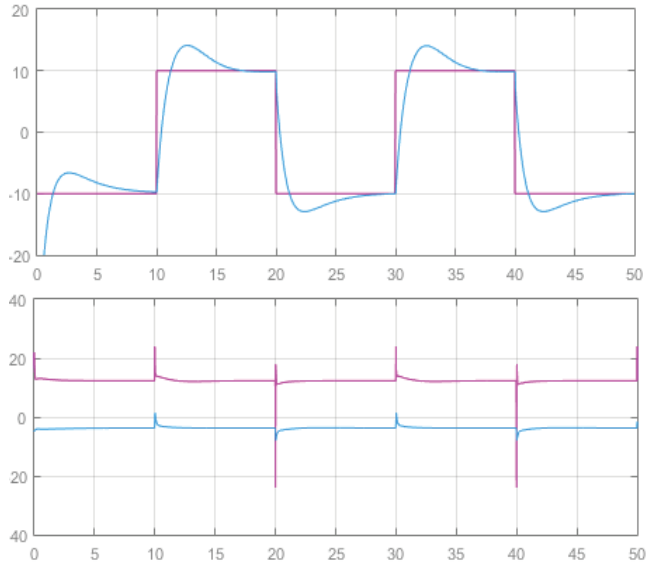


Fig. 7. Pitch angle variations (amplitude vs time) and voltage variations (volt. vs time) on 2DOF helicopter model after using TLBO-PID controller

The graphical results obtained using the three different control strategies are further analyzed to demonstrate a mathematical comparison between these techniques, as shown in Table I. Demonstration of helicopter model is conducted by implementation of conventional LQR controller, the TLBO-optimized PID controller and the fuzzy logic controller on the Simulink/MATLAB platform. Step reference trajectories are used to analyze the time response results, and the variations in pitch angle are compared across the different control schemes.

TABLE I. TIME RESPONSE ANALYSIS OF PITCH AXIS

	Settling Time (milli seconds)	**Peak Over-shoot (in %)**
LQR	4.72	8.99
TLBO-PID	3.99	4.18
FUZZY	0.95	0.5

V. CONCLUSION

Study of LQR, TLBO-optimized P-I-D and intelligent Fuzzy controllers for achieving predefined pitch axis trajectory of a helicopter model is performed. The TLBO-optimized PID controller was used to track the pitch axis trajectory and compared with the fuzzy system's inference rules and the conventional LQR controller. The membership function parameters and the optimization algorithm's exploration-exploitation scheme were discussed in the context of trajectory tracking. The parameter tuning and stable response of the pitch axis were achieved using Simulink/MATLAB and the comparison was conducted through time response analysis. Graphical and mathematical results are explaining the superiority of optimized PID control strategy and intelligent FLC control strategy over the conventional LQR control strategy and the same is explained with the help of different figures and a separate table in the different sections of this paper.

REFERENCES

[1] T. Kiefer, K. Graichen, and A. Kugi, "Trajectory Tracking of a 3DOF Laboratory Helicopter Under Input and State Constraints," *IEEE Trans. Control Syst. Technol.*, vol. 18, no. 4, pp. 944-952, 2010.

[2] I. A. Raptis, K. P. Valavanis, and G. J. Vachtsevanos, "Linear Tracking Control for Small-Scale Unmanned Helicopters," *IEEE Trans. Control Syst. Technol.*, vol. 20, no. 4, pp. 995-1010, 2012.

[3] K. Tanaka, H. Ohtake, M. Tanaka, and H. O. Wang, "Wireless Vision-Based Stabilization of Indoor Microhelicopter," *IEEE/ASME Trans. Mechatronics*, vol. 17, no. 3, pp. 519-524, 2012.

[4] C. Liu, J. Pan, and Y. Chang, "PID and LQR Trajectory Tracking Control for an Unmanned Quadrotor Helicopter: Experimental Studies," in *Proc. 35th Chin. Control Conf. (CCC)*, Chengdu, China, Jul. 2016, pp. 10845–10850.

[5] M. Jahed and M. Farrokhi, "Robust Adaptive Fuzzy Control of Twin Rotor MIMO System," *Soft Comput.*, vol. 17, no. 10, pp. 1847–1860, Oct. 2013.

[6] J. D. Boskovic and R. K. Mehra, "A Decentralized Fault-Tolerant Control System for Accommodation of Failures in Higher-Order Flight Control Actuators," *IEEE Trans. Control Syst. Technol.*, vol. 18, no. 5, pp. 1103–1115, Sep. 2010.

[7] A. Chaudhary, "Innovative Control Framework for Nonlinear Vibrations of a Balancer System Through Hybrid FPID Controller," *J. Vib. Eng. Technol.*, vol. 13, p. 115, 2025.

[8] A. Chaudhary, "Axis Control of a Nonlinear Helicopter Model Using Intelligent Controller," in *Proc. 2023 Int. Conf. Adv. Technol. (ICONAT)*, Goa, India, 2023, pp. 1-6.

[9] K. Xu, G. Zhang, and Y. Xu, "Dual Fuzzy Neural Network Control in Civil Aviation Intelligent Landing System," in *Proc. AICI 2011, Part II*, LNAI 7003, Springer-Verlag, 2011, pp. 166-174.

[10] J. G. Ziegler and N. B. Nichols, "Optimum Settings for Automatic Controllers," *Trans. ASME*, vol. 64, pp. 759–765, 1942.

[11] X. X. Chen, Y. Liu, and X. W. Hu, "Velocity Control of Unmanned Helicopter Based on Fuzzy-PID," *Adv. Mater. Res.*, vol. 756, pp. 327–331, 2013.

[12] X. Hou, P. Li, and Z. Fang, "An Application of Fuzzy PID Algorithm on Unmanned Helicopter Attitude Control," in *Proc. 6th World Congr. Intell. Control Autom. (WCICA)*, 2006.

[13] G. Xing, W. Lifeng, and M. Zhijun, "Fuzzy Self-Tuning PID Controller for an Unmanned Helicopter," in *Proc. 34th Chin. Control Conf. (CCC)*, 2015.

[14] R. V. Rao, V. J. Savsani, and D. P. Vakharia, "Teaching–Learning-Based Optimization: A Novel Method for Constrained Mechanical Design Optimization Problems," *Comput.-Aided Des.*, vol. 43, no. 3, pp. 303-315, 2011.

[15] Quanser Inc., *Quanser 2-DOF Helicopter Manual*, Tech. Rep., Quanser, 2010.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:15:07 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the International Conference on Machine Learning and Autonomous Systems (ICMLAS-2025)
IEEE Xplore Part Number: CFP258A9-ART; ISBN: 979-8-3315-0574-5

2025 International Conference on Machine Learning and Autonomous Systems (ICMLAS) | 979-8-3315-0574-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICMLAS64557.2025.10968795

Coordination Frameworks for IoT-Driven Swarm Robotics Communication

Ayush Kumar Agrawal
Department of Electronics and Communication
National Institute of Technology Jamshedpur
Jamshedpur, India
2020rsec001@nitjsr.ac.in

Jayendra Kumar
Department of Electronics and Communication
National Institute of Technology Jamshedpur
Jamshedpur, India
jkumar.ece@nitjsr.ac.in

Avanish Kumar
Department of Electrical Engineering
Delhi Technological University
Delhi, India
avanish180803@gmail.com

Anumeha
Department of Electronics and Communication
Government Women's Polytechnic Jamshedpur
Jamshedpur, India
amehanitjsr@gmail.com

Arvind R. Yadav
Department of Electronics and Instrumentation Engineering
Institute of Technology, Nirma University
Ahmedabad, India
arvindkumar.yadav@nirmauni.ac.in

***Abstract*—The Internet of Things (IoT) has revolutionized global device interconnectivity, fostering seamless data exchange through expansive sensing networks. Similarly, swarm robotics emulates the collective behaviors observed in nature, such as bird flocks and ant colonies, aiming to adapt these principles for practical use. This paper explores the dynamics of robotic swarms operating under a centralized network model using master-slave communication. Two implementation approaches are examined: direct, small-scale interactions utilizing device MAC addresses and large-scale connectivity facilitated by an IoT-based cloud server. Our experiments employ the NodeMCU ESP8266 platform, which utilizes the IEEE 802.11 Wi-Fi standard for data handling and transfer. For cloud-based communication, we utilize the ThingSpeak server to manage data flow. This study evaluates the strengths and limitations of these approaches in fostering swarm behavior, aiming to derive strategies for enhancing the performance and coordination of swarm robotic systems.**

***Index Terms*—Centralized Communication, Collaborative Robotics, Internet of Things (IoT), Swarm Robotics, Device Interconnectivity**

I. Introduction

Effective communication is at the heart of swarm robotics, much like how a team must communicate to work toward a common goal. Think of a swarm of bees: one bee might find food, and then shares that information with the rest of the swarm so they can all act in coordination. Similarly, robots in a swarm need to exchange information to perform tasks as a unified group. Various communication methods help make this possible, including master-slave setups, mesh networks, and ring topologies. In a master-slave configuration, one robot takes the lead and directs the others, like a coach guiding a team. Mesh and ring networks, however, allow a more distributed approach, where robots communicate directly with each other. In a mesh network, messages are passed along between robots like notes in a chain, while in a ring topology, information circulates in a continuous loop, much like a game where players pass a ball around. These frameworks play a crucial role in making swarm robotics work smoothly [1]–[3].

In this study, we explore these communication frameworks on the NodeMCU ESP8266 platform, which uses Wi-Fi for both online and offline communication. In the online mode, robots connect via a router, while in offline mode, they use MAC addresses for direct interactions without relying on the internet. We specifically examine master-slave communication in both scenarios and assess mesh and ring topologies in offline settings to observe how the robots interact [4]. Additionally, we look into IoT-enabled communication through cloud servers, where robots can exchange data over the internet, allowing them to collaborate from different locations. By analyzing these frameworks, we aim to discover ways to improve the efficiency and robustness of swarm robotic systems [1]–[3].

II. Related Work

Existing models in swarm robotics, while foundational, have several limitations, particularly when it comes to scalability and communication reliability. Master-slave configurations, although effective for smaller systems, face challenges in larger swarms due to the reliance on a single master to manage all communications. As the number of robots increases, this can create bottlenecks, reducing the overall system efficiency. Similarly, mesh and ring topologies, which offer more dis-

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:38:24 UTC from IEEE Xplore. Restrictions apply.

tributed communication, encounter issues such as network congestion and message loss, especially when operating offline. In offline scenarios, the lack of a centralized server can result in disconnected or delayed communication, undermining the system's reliability. These limitations hinder the ability of swarm robotic systems to scale effectively in complex environments [1], [2].

Another major challenge is the dependency on centralized infrastructure, such as routers or cloud servers, to facilitate communication. This creates a vulnerability where a failure in the central server or network can disrupt the entire swarm's operation. Furthermore, many models lack adaptive mechanisms that can adjust to dynamic or changing environments, limiting the system's flexibility and resilience. As swarm robotics moves toward more autonomous and decentralized systems, addressing these issues is crucial to improving both their efficiency and robustness in real-world applications [3], [4].

III. Swarm Robotics

Swarm robotics focuses on the coordinated operation of numerous simple robots that adhere to basic interaction rules to achieve complex objectives. Inspired by social species such as ants, birds, and fish, this approach leverages collective behaviors observed in nature to address various challenges. Swarm intelligence, a paradigm rooted in algorithms mimicking self-organizing behaviors of social insects, underpins the foundation of swarm robotics. In these systems, robots engage with each other and their environment to exhibit collective behaviors. While individual robots perform simple tasks, their collective actions enable the achievement of complex goals [2], [5]–[7]. Swarm robots differ from traditional multi-robot systems through their decentralized and collective operational nature. Key characteristics distinguishing swarm robotics include scalability, fault tolerance, and adaptability, all of which make them suitable for diverse applications, which are:

1) Autonomous robots that can sense their environment and act accordingly.
2) The swarm is composed of a large number of robots working together.
3) The robots within the swarm should be uniform in design and functionality.
4) Every robot must have the ability to sense its surroundings and communicate with others.
5) Collaboration between robots is essential to accomplish tasks efficiently.

Swarm robotics is capable of performing various tasks, often tested through experimental setups. Some common tasks are listed below:

1) Aggregation: This task involves gathering all robots in a designated area within the environment.
2) Dispersion: Focuses on spreading the robots evenly across a defined space to maximize coverage.
3) Pattern Formation: The robots adjust their positions to form a specific pattern or shape.
4) Collective Movement: Ensures the synchronized movement of multiple robots as a unified group.
5) Source Search: Used for operations such as locating sounds or odors in a search area.

Other behaviors observed in swarm robotics include task allocation, collaborative object transport, and joint mapping efforts. To implement swarm robotics in real-world applications, several algorithmic approaches are applied, such as Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), and Artificial Bee Colony (ABC). These algorithms are utilized in various fields, including healthcare, smart city development, and transportation. Additionally, swarm robotics is beneficial for applications such as disaster relief, surveillance in hazardous areas, and smart management of urban systems [8]–[11].

IV. NodeMCU in Swarm Robotics

Swarm robots are typically composed of many simple units that individually perform basic tasks, but collectively carry out more complex operations. A key feature of swarm robotics is leveraging a large number of low-cost, simple robots to achieve complex behaviors. Effective communication is essential in a swarm, so each robot must be able to interact with others within the group. The NodeMCU ESP8266 is a suitable choice for this purpose, serving two primary functions: acting as the robot's microprocessor to process sensor data and control robot behavior, and enabling communication through TCP/IP protocols.

The NodeMCU ESP8266 features a Tensilica Xtensa 32-bit LX106 RISC microprocessor, capable of operating between 80 MHz and 160 MHz clock speeds. It includes 128KB of RAM and 4MB of flash memory for storing data and programs [7], [12], [13]. The ESP8266 chip, known for its low cost, supports the IEEE 802.11 Wi-Fi standard and includes a complete TCP/IP stack, enabling wireless communication between robots. Its affordability, ease of use, and communication capabilities make the NodeMCU ESP8266 an ideal choice for decentralized swarm robotics communication, even without an internet connection.

In this study, different network topologies are employed to analyze the swarm behavior of robots. Some topologies represent centralized systems, while others exhibit decentralized swarm behavior. Three network topologies are considered: a) master-slave, b) mesh, and c) ring.

The master-slave topology involves one robot acting as the master, overseeing and controlling the actions of the other robots, which function as slaves. In this configuration, the master robot monitors data from the slave robots and issues commands if any issues arise. Slave robots can only receive data from the master and cannot send instructions back. Therefore, data transfer is bidirectional, but command transfer is unidirectional [14]. If a slave robot fails, the master can replace it with another functioning robot to continue the task. Figure 1 illustrates master-slave communication between NodeMCU ESP8266 units using their MAC addresses.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:38:24 UTC from IEEE Xplore. Restrictions apply.

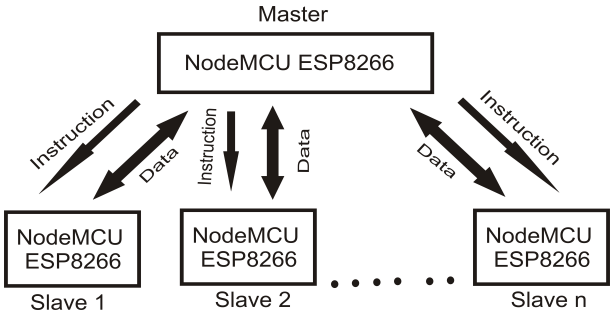


Fig. 1. Master-Slave Communication between Multiple NodeMCU ESP8266 Devices

In contrast, the mesh topology does not rely on a single robot to control the others. Each robot in a mesh network is connected to multiple other robots, allowing messages to be relayed through various paths. This architecture enhances fault tolerance, as communication can continue even if one path fails. The density of information flow depends on the number of connections between robots in the swarm. For n robots, the total number of connections in the mesh network is calculated as:

$$\text{Total connections} = \frac{n(n-1)}{2}$$

where n is the number of robots. This setup allows for more robust communication due to the multiple pathways available.

As Figure 2 shows the example of mesh communication between five NodeMCU ESP8266 and the total number connection is calculated using above formula is 10.

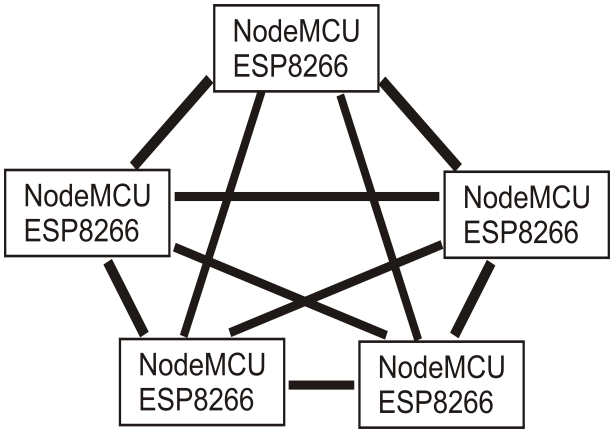


Fig. 2. Mesh Communication Between Multiple NodeMCU ESP8266 Devices

The third network configuration is the ring topology, which, like the mesh topology, does not assign control to any single robot in the swarm. In this configuration, the robots are arranged in a circular structure, where information circulates unidirectionally from one robot to the next until it reaches its destination. This setup limits the flow of information, as each robot is connected to only two others, creating a longer communication path and causing delays in transmission [15], [16]. As the number of robots increases in the swarm, the transmission time grows due to the extended route between the source and the destination. Therefore, the reliability of communication in a ring topology decreases as the swarm size increases, meaning the probability of successful message delivery diminishes. However, one advantage of using the ring topology is the ease with which a faulty robot can be replaced within the loop. Figure 3 illustrates a setup where five NodeMCU ESP8266 devices are connected in a ring configuration for communication.

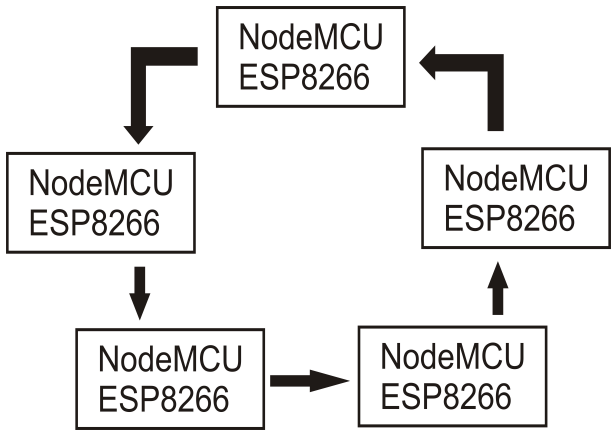


Fig. 3. Ring Communication Among Multiple NodeMCU ESP8266 Devices

As demonstrated, all three topologies—master-slave, mesh, and ring—are viable for implementing swarm behavior in robotic swarms, with the choice depending on specific needs. Both the mesh and ring configurations support decentralized communication, while the master-slave topology provides a more centralized communication structure. In each of these configurations, communication can occur without requiring internet access, utilizing the MAC addresses of the robots to facilitate message passing within the swarm.

V. Swarm Robotics with the Internet (IoT)

Over the past two decades, the Internet of Things (IoT) has seen remarkable progress, enabling widespread connectivity and computing capabilities. This transformation has positioned IoT as a crucial platform for linking devices to the internet, allowing for Machine-to-Machine (M2M) communication and data transfer through standard networking protocols such as TCP/IP. Swarm robotics is one of the key applications that benefits from M2M communication and data transfer. In many swarm robotics systems, cloud servers are utilized to provide centralized communication within the network [2], [5], [17]. The cloud infrastructure also supports computation and storage, facilitating the use of swarm intelligence algorithms that enhance communication and exhibit collective swarm behaviors.

Swarm intelligence, a subfield of artificial intelligence, focuses on creating intelligent systems that mimic the collaborative behaviors of social animals and insects. Several

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:38:24 UTC from IEEE Xplore. Restrictions apply.

swarm intelligence algorithms can be applied to IoT-based applications, including:

1) Ant Colony (AC) Algorithm,
2) Ant Colony Optimization (ACO) Algorithm,
3) Particle Swarm Optimization (PSO) Algorithm,
4) Artificial Bee Colony (ABC) Algorithm,
5) Bacterial Foraging Optimization Algorithm (BFOA), and others.

Swarm robotics can utilize various techniques for enabling communication between devices, one of which is the master-slave configuration. In some cases, internet connectivity is employed to establish communication. When internet-based communication and data storage are necessary, a cloud server is often used [18]. In this study, the ThingSpeak cloud server was selected to store data from both the master and slave devices. As shown in Figure 4, the slave devices can only upload data to the cloud and communicate bidirectionally with the master. In contrast, the master device has the capability to both read from and write to the cloud server, allowing it to monitor the data from the slave devices.

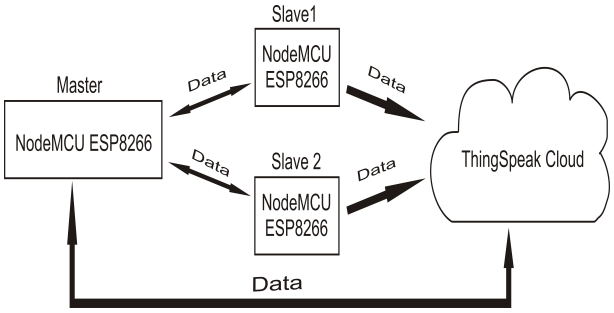


Fig. 4. Master-Slave Communication via IoT

VI. Results and Discussion

This study explored different approaches to implementing swarm behavior using NodeMCU ESP8266 devices and ultrasonic sensors. Three distinct communication topologies were used to facilitate communication within the swarm. These topologies include:

A. Communication Among Swarm Using the Master-Slave Configuration

1) Communication Without Internet: The first topology examined was the master-slave communication model. In this model, data transfer was simulated using the Arduino IDE. The results showed that both the master and slave devices successfully exchanged data without the need for internet connectivity. Communication was established using MAC addresses, as illustrated in Figure 5.

2) Communication Using Internet (IoT): For the IoT-based master-slave communication, the ThingSpeak cloud server was utilized for data storage. The master device could perform both write and read operations on the cloud server, whereas the slave devices were limited to writing data. The simulation results of data transfer between the master and slave devices are shown in Figure 6.

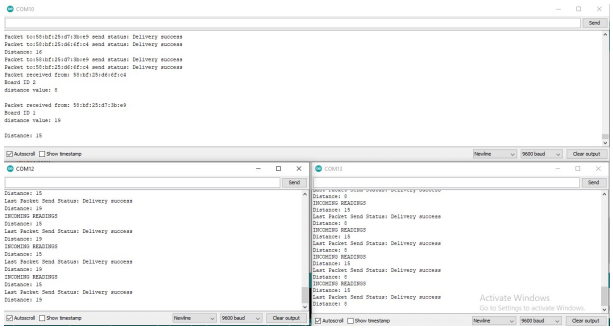


Fig. 5. Master-Slave Communication in Swarm Robotics Without Internet

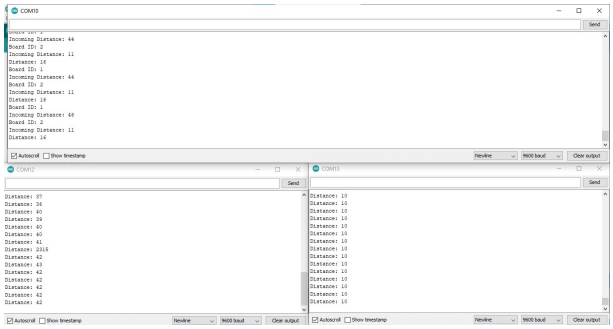


Fig. 6. Master-Slave Communication in Swarm via IoT

B. Communication Among Swarm Using Mesh Topology

The second topology implemented was the mesh network, where IoT-based communication was utilized, and the ThingSpeak cloud server was employed for data storage. The results showed that the master device could both write and read data from the cloud server, while the slave devices could only write data. Real-time data transfer among the master, $slave_1$, and $slave_2$ was observed in Figures 7, 8, and 9, respectively.

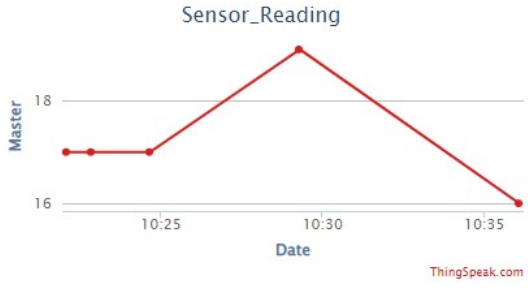


Fig. 7. Sensor Data from the Master Robot

The mesh topology allowed each device to share its ultrasonic sensor data with every other device in the swarm. The simulation results demonstrated that each device could send data to two other devices and also monitor data from these devices. Figure 10 illustrates the data transfer among all three devices.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:38:24 UTC from IEEE Xplore. Restrictions apply.

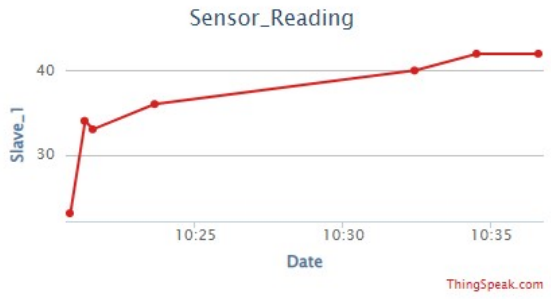


Fig. 8. Sensor Data from $Slave_1$

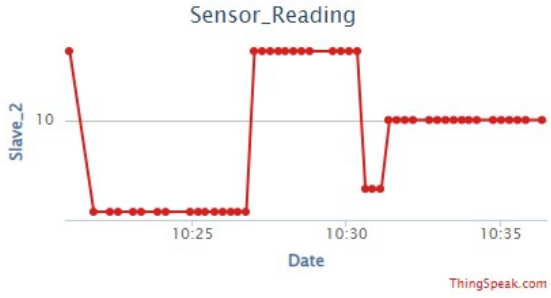


Fig. 9. Sensor Data from $Slave_2$

C. Communication Among Swarm Using Ring Topology

The final topology tested was the ring topology, where the devices were interconnected in a closed-loop. In this setup, information traveled unidirectionally from one device to the next. Each device shared data with one of its adjacent devices and received data from the next device in the loop. The results of data transfer in this configuration are shown in Figure 11.

From the simulated results, it was observed that the master-slave communication topology provided more accurate data transfer compared to the IoT-based swarm communication. The master-slave configuration allows centralized communication, where the master device has control over data exchange and operations. The mesh and ring topologies, however, allow each device to operate autonomously within the swarm.

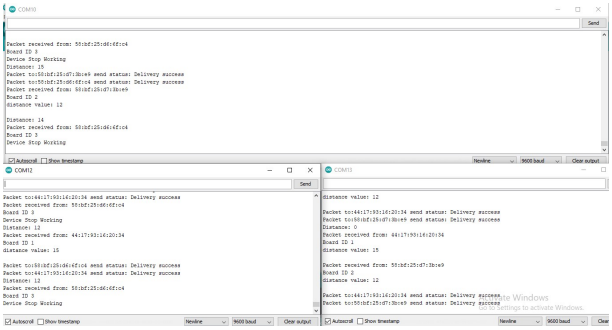


Fig. 10. Communication in the Swarm Using Mesh Topology

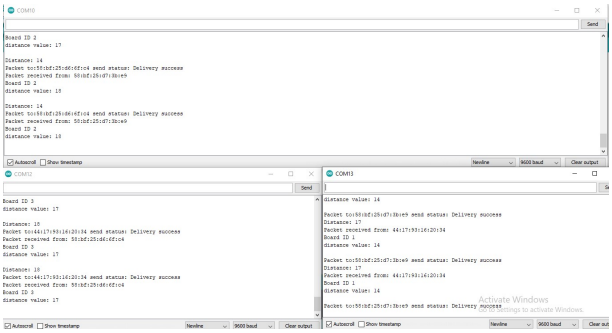


Fig. 11. Communication in the Swarm Using Ring Topology

Figures 6, 10, and 11 illustrate the results for the master-slave, mesh, and ring communication models, respectively, using NodeMCU ESP8266 devices. In the master-slave setup, the master device monitored the slaves' data and issued instructions in case of malfunction. In the mesh topology, each device exchanged data with all other devices and could instruct a malfunctioning device to be replaced. In the ring topology, each device communicated with only one adjacent device, thus managing communication in a more limited manner.

The method described in this paper demonstrates several strengths in its approach to implementing communication topologies in swarm robotics using NodeMCU ESP8266 devices and ultrasonic sensors. One major strength is the flexibility provided by the NodeMCU platform, which allows for both online and offline communication modes. This versatility is crucial for testing various topologies such as master-slave, mesh, and ring networks, each of which can be adapted to different swarm scenarios. The ability to use ultrasonic sensors for proximity detection further enhances the robots' ability to interact with their environment, facilitating more effective swarm behaviors. Additionally, the simplicity of the NodeMCU ESP8266 platform makes it accessible and cost-effective for swarm robotics research.

However, the method also has notable weaknesses. While the study explores a range of communication frameworks, the offline communication using MAC addresses can be prone to latency and data loss, particularly in larger swarms where network congestion is more likely. Moreover, relying on Wi-Fi for communication introduces potential issues with signal interference and range limitations, especially in environments with many obstacles or in expansive areas. Despite these challenges, the results provide valuable insights into the feasibility of implementing these communication methods in swarm robotics. A comparative summary of the communication methods explored is presented in Table I.

VII. Conclusion

In conclusion, this paper has explored and compared various communication topologies for swarm robotics, focusing on master-slave, mesh, and ring configurations. The findings suggest that for large-scale applications, the master-slave communication method combined with IoT proves to be the most

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:38:24 UTC from IEEE Xplore. Restrictions apply.

TABLE I
COMPARISON OF COMMUNICATION APPROACHES IN ROBOTIC SWARMS

Feature	Swarm Communication with Master-Slave Approach		Swarm Communication with Mesh Topology	Swarm Communication with Ring Topology
	Communication without Internet	Communication via Internet (IoT)		
Dependence on Intermediate Medium	No	Yes	No	No
Number of Devices Connected	Few	Many	Few	Few
Coverage Area	Moderate	Large	Small	Small
Device Monitoring Method	Requires physical monitoring only	Can be monitored remotely	Requires physical monitoring only	Requires physical monitoring only
Communication Medium	Uses MAC Address	Router-based	Uses MAC Address	Uses MAC Address
Malfunction Handling	Replacement is easy	Replacement is straightforward	Replacement is complex	Replacement is moderate

suitable due to its centralized nature and scalability, enabling efficient management of multiple robots. For smaller-scale setups, any of the three topologies—master-slave, mesh, or ring—can be employed, as they rely on MAC addresses for device communication. However, it was found that offline communication, without internet access, tends to yield more accurate sensor readings, as it reduces potential latency and network interference.

Both the mesh and ring topologies support decentralization, a crucial aspect of swarm behavior, enabling more flexible and autonomous coordination among robots. On the other hand, the centralized master-slave method provides better control over the swarm's actions, making it ideal for certain structured tasks. In the mesh topology, there is a risk of data collision when a device needs to be replaced, as robots communicate directly with each other. Similarly, in the ring topology, data replacement can take longer due to the sequential nature of communication, as information must pass through every device in the loop.

Future work will focus on advancing decentralized communication methods, which are vital for swarm robotics' full potential. As technology evolves, swarm robotics is poised to transform industries such as manufacturing, agriculture, and healthcare, offering solutions that are more efficient, adaptable, and autonomous.

REFERENCES

[1] Wang, Xiuwang, Xinna Mao, and Hossein Khodaei. "A multi-objective home energy management system based on internet of things and optimization algorithms." Journal of Building Engineering 33 (2021): 101603.

[2] Agrawal, Ayush Kumar, and Manisha Bharti. "Feature, Technology, Application, and Challenges of Internet of Things." In Big Data, IoT, and Machine Learning, pp. 255-276. CRC Press, 2020.

[3] Gazi, Peymon, Mo Jamshidi, Aleksandar Jevtić, and Diego Andina. "A mechatronic system design case study: Control of a robotic swarm using networked control algorithms." In 2010 IEEE International Systems Conference, pp. 169-173. IEEE, 2010.

[4] Sharkey, Amanda JC, and Noel Sharkey. "The application of swarm intelligence to collective robots." In Advances in applied artificial intelligence, pp. 157-185. IGI Global, 2006.

[5] Albani, Dario, Joris IJsselmuiden, Ramon Haken, and Vito Trianni. "Monitoring and mapping with robot swarms for agricultural applications." In 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pp. 1-6. IEEE, 2017.

[6] Sivakumar, T. B., S. Hasan Hussain, A. Kanmani, and MH Anand Babu. "WITHDRAWN: Surveillance robot for health care applications using IoT and wireless sensor network." (2021).

[7] Sangeetha, B. Parvathi, Neeraj Kumar, Ambresh P. Ambalgi, Sulaima Lebbe Abdul Haleem, K. Thilagam, and P. Vijayakumar. "IOT based smart irrigation management system for environmental sustainability in India." Sustainable Energy Technologies and Assessments 52 (2022): 101973.

[8] Agrawal, Abhishek, and Pratyush Kumar Padhi. "Low Cost Approach to a Two Agent Imitating Master-Slave Robotic System." In 2021 International Symposium of Asian Control Association on Intelligent Robotics and Industrial Automation (IRIA), pp. 414-418. IEEE, 2021.

[9] Bharti, Manisha, and Ayush Kumar Agrawal. "War field spy and fighter robot." In Recent Advances in Mechanical Engineering: Select Proceedings of NCAME 2019, pp. 565-574. Springer Singapore, 2020.

[10] Rezeck, Paulo AF, Hector Azpurua, and Luiz Chaimowicz. "HeRo: An open platform for robotics research and education." In 2017 Latin American Robotics Symposium (LARS) and 2017 Brazilian Symposium on Robotics (SBR), pp. 1-6. IEEE, 2017.

[11] Bento, Antonio Carlos. "IoT: NodeMCU 12e X Arduino Uno, Results of an experimental and comparative survey." International Journal 6, no. 1 (2018).

[12] Blender, Timo, Thiemo Buchner, Benjamin Fernandez, Benno Pichlmaier, and Christian Schlegel. "Managing a mobile agricultural robot swarm for a seeding task." In IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society, pp. 6879-6886. IEEE, 2016.

[13] Huang, Shan-Huen, Ying-Hua Huang, Carola A. Blazquez, and Chia-Yi Chen. "Solving the vehicle routing problem with drone for delivery services using an ant colony optimization algorithm." Advanced Engineering Informatics 51 (2022): 101536.

[14] Navarro, Iñaki, and Fernando Matía. "An introduction to swarm robotics." Isrn robotics 2013 (2013): 1-10.

[15] Agrawal, Ayush Kumar, and Manisha Bharti. "Aspects and Use of Digital Agriculture Using IoT." In Harnessing the Internet of Things (IoT) for a Hyper-Connected Smart World, pp. 243-258. Apple Academic Press, 2022.

[16] Liekna, Aleksis, and Janis Grundspenkis. "Towards practical application of swarm robotics: overview of swarm tasks." Engineering for rural development 13 (2014): 271-277.

[17] Bhuiya, Aritra, Aditya Mukherjee, and Ranjit Kumar Barai. "Development of wi-fi communication module for atmega microcontroller based mobile robot for cooperative autonomous navigation." In 2017 IEEE Calcutta Conference (CALCON), pp. 168-172. IEEE, 2017.

[18] Agrawal, Ayush Kumar, Pritam Pidge, Manisha Bharti, M. Prabhat Dev, and Prashant Kaduba Kedare. "Innovations and future of robotics." In

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:38:24 UTC from IEEE Xplore. Restrictions apply.

Advances in Manufacturing and Industrial Engineering: Select Proceedings of ICAPIE 2019, pp. 975-980. Springer Singapore, 2021.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:38:24 UTC from IEEE Xplore. Restrictions apply.

arXiv:2504.09523v1 [astro-ph.CO] 13 Apr 2025

Decaying vacuum energy, matter creation and cosmic acceleration

Lokesh Chander* and C P Singh†
Department of Applied Mathematics,
Delhi Technological University, Delhi-110042, India
(Dated: April 15, 2025)

In this paper, we discuss an interacting decaying vacuum energy and dark matter empowered by gravitationally induced matter creation model, and its impact on structure formation by analysing the growth rate perturbations. Our work is motivated by the possibility that the decaying vacuum is due to quantum field theory and dark matter originates from gravitationally induced matter creation. We delve deeper into our investigation and explore both theoretical and statistical analysis of the cosmological model to test its ability to describe the evolution of the Universe. To achieve this, we use three distinct combinations of datasets from Cosmic Chronometers, Pantheon Type Ia supernovae sample, Baryon acoustic oscillations, Cosmic Microwave Background distance priors and $f(z)\sigma_8(z)$ datapoints to constrain the model parameters. Our statistical analysis employs Markov Chain Monte Carlo (MCMC) methods. The deceleration parameter shows that the model transitions from a decelerated phase to an accelerated phase of expansion. The current Hubble parameter values are estimated to be $H_0 = 67.517 \pm 0.869$ km/s/Mpc, $H_0 = 67.534 \pm 0.874$ km/s/Mpc, and $H_0 = 67.533 \pm 0.884$ km/s/Mpc using DS1, DS2 and DS3 datasets, respectively. These values of H_0 are very close to those derived from the Planck data. The effective equation of state parameter indicates an accelerating phase, with density parameter for vacuum energy exhibiting expected values. The present value of the jerk parameter deviates slightly from the ΛCDM model. However, it approaches the standard model in late time. We analyse the stability characteristics through the selection information criteria. We also perform thermodynamic analysis by studying the evolution of entropy in the Universe for the model and find it to be in agreement with the generalised second law of thermodynamics. These findings support that the proposed model effectively describes the evolutionary features of the Universe at both theoretical and observational levels.

I. INTRODUCTION

The first observational evidence for the accelerated expansion of the Universe was confirmed through Type Ia supernovae (SNIa), rightly called 'standard candles', as these astronomical objects have an intrinsic luminosity [1, 2]. This discovery significantly changed our understanding about the Universe and caused a resurgence in modern cosmology. Subsequent datasets: Baryon Acoustic Oscillations (BAO) [3], the Hubble parameter measurements $H(z)$ [4], and Cosmic Microwave Background (CMB) anisotropies [5] have further validated this phenomenon. According to 'General Relativity' (GR), this accelerated expansion is driven by a new energy density with negative pressure, referred to as 'dark energy'(DE), which makes up around 70% of the Universe's total energy density. Accompanying this is 'dark matter'(DM), whose nature is not completely understood and together, both dark components make up for 95% of the Universe's content [6]. Understanding the origin of DE remains one of the major unsolved issues in modern cosmology as its exact nature is still unknown. The naturally arising candidate for dark energy within the framework of GR is the cosmological constant, denoted by Λ, or equivalently, the vacuum energy density associated with it, $\rho_\Lambda = \Lambda/8\pi G$ [7]. It has a negative pressure, making it the repulsive driving force that causes accelerated expansion under the condition that Λ has a value greater than zero [8].

However, the Λ-based concordance model faces theoretical challenges, most notably the cosmological constant problem and cosmic coincidence [9–13], which remain central issues in cosmology. These unresolved problems have motivated alternative approaches, such as dynamic dark energy models, where Λ evolves as a function of time, and interacting models, where dark energy interacts with other components in the cosmic fluid. In decaying vacuum energy models, Λ is not a constant; rather, it evolves over time and is thus a dynamic component $\Lambda(t)$. This approach addresses the fine-tuning problem by linking the present-day small value of Λ to its evolutionary trajectory during the Universe's history. Early studies [14–29], often motivated by phenomenological considerations, laid the theoretical foundations and proposed certain functional forms of $\Lambda(t)$. More recently, many authors [30–37] have sought to ground these models in quantum field theory (QFT). This approach also works when Λ is considered to be a relic of inflationary dynamics [38].

Besides the dynamical $\Lambda(t)$ approach, recently, a new alternative to address late-time cosmic acceleration has gained traction. It is the 'adiabatic' matter creation which is gravitationally induced, and a non-equilibrium thermodynamical process. A long time ago, Schrödinger [39] conceptualized matter creation in the microscopic realm, which was further developed by Parker [40, 41]. Since Einstein's field equations are the background equations to understand the evolution of the Universe, it was

* lokeshchander˙23phdam04@dtu.ac.in
† cpsingh@dce.ac.in

Prigogine and his collaborators [42–44] who proposed a thermodynamic framework for open systems, allowing the creation of matter via transfer of energy from the gravitational field. In this case, by assuming that DM is being produced due to a time-varying gravitational field, it is plausible for a late-time acceleration to happen in a Universe consisting of only pressureless fluids, like cold dark matter and baryons. The research work carried out by many authors [45–61] suggests that matter creation can mimic the effects of dark energy, producing negative pressure and driving accelerated expansion. These models, rooted in thermodynamics, posit that new matter gets created as the Universe expands. With the energy conservation law, we suggest that this creation comes at the expense of another component in the cosmic fluid. When combined with DE characterized by a decaying vacuum term $\Lambda(t)$, there is a naturally arising coupling between the DE and matter sector, which consists of baryons and DM. The decay of $\Lambda(t)$ thus contributes to the creation of matter. While the exact mechanism of this phenomenon cannot be deduced, the implication is clear that DE and DM are interacting entities. This idea is somewhat similar to that of an exotic fluid like the Chaplygin gas (CG), which behaves like dark matter at early redshifts and like dark energy at later times. Many models based on various forms of the CG have been studied in the literature, and this approach has been promising since it is consistent with observational data [62–64]. Moreover, there are interacting dark energy (IDE) models in which interaction between the DE and other components in the cosmic fluid is assumed, and these models have been extensively explored in the literature. Along with the class of CG models, these models can solve the coincidence problem by allowing dark energy and dark matter to evolve over time. The gradual convergence of the two densities helps to resolve the issue of their seemingly fine-tuned values at present [65–68].

The present study investigates a combined model incorporating decaying vacuum energy and matter creation, offering a unified framework to address the aforementioned challenges. Unlike traditional interacting dark energy models that require an explicitly defined coupling term, this approach inherently links the decay of $\Lambda(t)$ to the rate of matter creation, denoted by Γ, avoiding the need for arbitrary assumptions about the form of coupling. To test the viability of the proposed model, we undertake a rigorous theoretical and observational analysis. With the help of Bayesian Markov Chain Monte Carlo (MCMC) methods, the model parameters are constrained with multiple datasets, including Baryon Acoustic Oscillations (BAO), Cosmic Microwave Background (CMB) distance priors, Pantheon Supernovae Type Ia (SNIa), H(z) measurements from Cosmic Chronometers (CC), and weighted linear growth rate measurements $f(z)\sigma_8(z)$. Further, we employ statistical criteria; the reduced chi-squared test and two model selection criteria, AIC and BIC, which judge the model based on information theory and Bayesian statistics, respectively. This allows for a robust comparison with the standard ΛCDM model, highlighting the potential advantages and limitations of the proposed framework.

The paper is organized in the following way: Section II lays down the groundwork for our cosmological model combining decaying vacuum energy with matter creation, outlining the modifications to the energy-momentum tensor and the Friedmann equations. In Section III, we present analytic solutions to the field equations, focusing on the following cosmological parameters: Hubble parameter $H(z)$, deceleration parameter $q(z)$, transitional redshift $z_{\rm tr}$, and effective equation of state (EoS) parameter $w_{\rm eff}(z)$. Section IV discusses perturbations in the cosmic background with primary focus on matter perturbations and their implications for structure formation. In Section V, we discuss the observational datasets and methodology employed for constraining the model, while in Section VI, we present the results of our analysis and discuss their cosmological implications. Section VII uses the reduced chi-squared test and two model selection criteria to compare our chosen model with a reference model, the ΛCDM Model. Then we focus on the thermodynamic analysis in Section VIII, where we proceed with finding an expression for the dark matter temperature. Then we calculate the total entropy of the universe and its evolution with respect to time, to test the generalized second law (GSL) of thermodynamics. Finally, in Section IX, we summarise our findings and the scope of future research.

II. DECAYING VACUUM MODEL WITH MATTER CREATION

Let us begin with a theoretical framework characterized by a spatially isotropic and homogeneous flat Friedmann-Lemaître-Robertson-Walker line element

$$ds^2 = -dt^2 + a^2(t)\left[dr^2 + r^2\left(d\theta^2 + \sin^2\theta\, d\phi^2\right)\right], \quad (1)$$

where (r, θ, ϕ) represents spherical coordinates and $a(t)$ is the scale factor, t denotes the cosmic time. Hereafter, we assume natural units $c = \hbar = 1$.

We introduce some basic thermodynamic concept of adiabatic matter creation as proposed by Prigogine and his collaborators [42–44] in cosmological context. We consider an N particle system under adiabatic conditions, which means that the total particle number changes and is not conserved. Mathematically, we express it as $N^{\mu}_{;\mu} \neq 0$. Due to this, the energy conservation equation gets modified. The particle flux vector, which measures how the particle number changes, is taken as $N^{\mu} = nu^{\mu}$, where $n = N/V$ gives the number of particles per unit of time (particle number density), and u^{μ} is the fluid's four-velocity.

The non-conservation of the particle flux, i.e., $\nabla_{\mu} n^{\mu} = \psi$, can be written as

$$\dot{n} + 3Hn = \psi = n\Gamma, \quad (2)$$

where $\dot{a}$ denotes derivative with respect to the cosmic time, $H = \dot{a}/a$ is the Hubble parameter and Γ quantifies the rate of change of the particle number in a physical volume V consisting N number of particles. Although the physical nature of Γ is unknown, we have one constraint over Γ, i.e., $\Gamma \geq 0$. For the generalized second law of thermodynamics to be valid, this must be true. Equation (2) gives the evolution of the particle number density. As the total number of particles per comoving volume a^3 is $N = a^3 n$, therefore, the above Eq.(2) can be written as $\dot{N}/N = \Gamma$. We define the entropy flux vector as $s^\mu = n\sigma u^\mu$, where σ is the entropy per particle. The entropy must satisfy the second law of thermodynamics, $s^\mu_{;\mu} \geq 0$. Also, we consider matter creation to be an adiabatic process, i.e., we have a system in which σ remains constant.

As we are assuming the total particle number N to be variable, we have a modification of the first law of thermodynamics:

$$d(\rho V) + p\, dV - \left(\frac{h}{n}\right) d(nV) = 0, \tag{3}$$

where ρ is the energy density and p is the equilibrium pressure of the cosmic matter fluid (dark matter and baryonic matter), and $h = (\rho + p)$ is the enthalpy per unit volume. We observe there to be an extra contribution in (3), and we interpret this as a non-thermal pressure arising due to the particle production process. We call this the 'creation pressure' p_c, which is given in terms of the change in particle number with respect to volume by

$$p_c = -\frac{h}{n}\frac{d(nV)}{dV}. \tag{4}$$

The negative sign indicates that p_c can help drive the accelerated cosmic expansion era. For simplicity's sake, we assume that only dark matter particles are produced since the production of baryons is limited by the local gravity measurement-induced constraints.

From (2) and (4), along with the modified first law in (3), we obtain the following form of the creation pressure [44, 45, 61]:

$$p_c = -(\rho + p)\frac{\Gamma}{3H}. \tag{5}$$

Let us outline the fundamental equations of a general cosmological scenario. The gravitational field equations in Einstein's General Relativity are:

$$G_{\mu\nu} = \kappa \tilde{T}_{\mu\nu}, \tag{6}$$

where $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R$ is the Einstein tensor which describes the spacetime geometry and κ is the Einstein gravitational constant which in this case is, $\kappa = 8\pi G/c^4 = 1$ due to the choice of units. $\tilde{T}_{\mu\nu} \equiv T_{\mu\nu} + g_{\mu\nu}\rho_\Lambda$ is the total energy-momentum tensor, which accounts for the contribution of both matter creation and a time-dependent vacuum term. In the presence of matter creation, it is customary to associate the creation pressure, p_c in energy-momentum tensor. As a result, the usual perfect fluid energy-momentum tensor, $T_{\mu\nu}$ gets modified to

$$T_{\mu\nu} = (\rho_m + \tilde{p}_m)\, u_\mu u_\nu + g_{\mu\nu}\tilde{p}_m, \tag{7}$$

where u_μ is the four-velocity of the matter fluid, ρ_m is the matter energy density, and $\tilde{p}_m$ is the total pressure exerted by the fluid which is the sum of the isotropic pressure p_m and the creation pressure p_c. We assume matter to be dust-like, i.e. it has negligible pressure ($p_m = 0$). In this paper, we assume that the vacuum behaves like a perfect fluid, with an equation of state given by $\rho_\Lambda = -p_\Lambda$. Given these considerations, we can write the Friedmann equations as follows:

$$3H^2 = \rho_m + \rho_\Lambda, \tag{8}$$

$$2\dot{H} + 3H^2 = -(p_c + p_\Lambda), \tag{9}$$

where $H = \dot{a}/a$ gives the expansion rate and is called the Hubble parameter. The conservation of the energy-momentum tensor, $\nabla^\mu \tilde{T}_{\mu\nu} = 0$ gives

$$\dot{\rho}_m + \dot{\rho}_\Lambda + 3H\,(\rho_m + \rho_\Lambda + p_c + p_\Lambda) = 0. \tag{10}$$

Since $\rho_\Lambda = -p_\Lambda$, the equation (10) gets simplified to:

$$\dot{\rho}_m + 3H\,(\rho_m + p_c) = -\dot{\rho}_\Lambda. \tag{11}$$

From (8) and (11), we can write the evolution equation as

$$2\dot{H} + 3H^2 + p_c = \rho_\Lambda. \tag{12}$$

III. SOLUTION OF THE FIELD EQUATIONS

In this section, we find the solution of the evolution equation (12) to constrain the parameters and discuss the results accordingly. Equation (12) can be solved for H if we know ρ_Λ, and Γ which gives us p_c through Eq. (5). Parametrization of Λ has changed over time and some in particular are: $\Lambda \propto a^{-2}$ [69], $\Lambda \propto a^{-m}$ [70] and $\Lambda \propto H^2$ [71]. Although there is no correct form for the decaying VED, it is often assumed as a function of the scale factor a or the expansion rate H. The motivation for taking it as a function of H is based on a quantum field theory (QFT) approach [72, 73]. This has been proposed in the context of the renormalization group (RG) [29, 74–76]. The interesting work by Shapiro and Solà [19], further investigated by many authors [37, 77–79] motivates us to adopt the following parametrization for the vacuum energy density:

$$\rho_\Lambda = c_0 + 3\nu H^2, \tag{13}$$

where c_0 is a constant term, and ν is an independent term associated with the vacuum decay. We expect ν to

be very small ($|\nu| \ll 1$) and interpret it as a perturbation term. In the limit $\nu \to 0$, the vacuum energy density becomes a constant, i.e. it reduces to the cosmological constant. This form is well motivated from the theoretical point of view as this is related to the general form of the effective action of QFT in curved spacetime.

In the proposed model, the cosmic history is based on the rate of expansion given by H and the evolution of energy density, which can define the gravitational creation rate Γ. This is often expressed as function of the expansion rate H, i.e. $\Gamma = \Gamma(H)$. The simplest form, though, would be taking Γ as a constant; $\Gamma = \Gamma_0$ and for matter creation to take place, this would be constrained by $\Gamma_0 > 0$. A direct proportionality to H can give us another form of the creation rate. In this paper, we consider a simple form that scales directly with the Hubble parameter [47] as

$$\Gamma = 3\beta H. \tag{14}$$

The parameter β is associated with the particle production. It is important to note that the matter creation rate Γ drops to zero in the limit $\beta \to 0$, meaning that the standard model with matter conservation (no creation or annihilation) is restored. Thus, like ν, it acts as a perturbation and therefore its value is expected be small and positive.

Since we propose that our framework is interacting, we write the continuity equations for the matter and vacuum components in the following way:

$$\dot{\rho}_m + 3H(1-\beta)\rho_m = Q, \quad \dot{\rho}_\Lambda = -Q. \tag{15}$$

Here Q is the interaction term, and its expression comes directly from the vacuum decay as

$$Q = -6\nu\dot{H}H. \tag{16}$$

The sign of Q determines the direction of energy flow. In the proposed scenario, vacuum energy decaying into dark matter is plausible because it aligns with particle production. Moreover, a model where the energy flow is in the opposite direction is generally not viable, [80] and therefore, we must ensure that Q is positive. We note that $Q \propto -6\dot{H}H = -(\dot{\rho}_m + \dot{\rho}_\Lambda)$ and ν in addition to being the vacuum decay parameter also determines the strength of the interaction. For the vacuum to act as a source, we must have $Q > 0$, and since $\dot{H}$ is negative, the viability condition gives us $\nu > 0$. Additionally, from the above, we interpret matter to have an effective equation of state $w_m = -\beta$ due to the adiabatic matter creation process, which indicates a small deviation in behavior from a dust-like fluid.

Using (13) and (14) into (12), we obtain

$$\dot{H} + \frac{3}{2}(1-\beta)(1-\nu)H^2 = 3H_0^2(1-\beta)(\Omega_\Lambda - \nu), \tag{17}$$

where $c_0 = 3H_0^2(\Omega_\Lambda - \nu)$, H_0 is the value of H at present and is called the Hubble constant, and Ω_Λ is the present value of the dark energy density parameter. Equation(17) is a differential equation in H with respect to cosmic time t. Instead of using the variable t, it is more convenient to write the above equation in terms of the scale factor a or redshift z. Thus, Eq. (17) can be rewritten as

$$\frac{dH^2}{da} + \frac{3}{a}(1-\beta)(1-\nu)H^2 = \frac{3}{a}(1-\beta)(\Omega_\Lambda - \nu)H_0^2. \tag{18}$$

The solution of (18) is given by

$$H(z) = H_0\left[\left(1 - \frac{\Omega_\Lambda - \nu}{1-\nu}\right)(1+z)^{3(1-\beta)(1-\nu)} + \frac{\Omega_\Lambda - \nu}{1-\nu}\right]^{1/2}, \tag{19}$$

where $(1+z) = a_0/a(t)$. Here, a_0 represents the current cosmic scale factor value, typically assumed to be 1. The above expression gives the expansion rate in terms of the four independent model parameters H_0, Ω_Λ, ν and β. Since we have considered only matter and vacuum components to be part of the cosmic fluid, we have $\Omega_m + \Omega_\Lambda = 1$.

From (19), the *scale factor* in terms of t can be found as

$$a = \left[\sqrt{\frac{\tilde{\Omega}_m}{\tilde{\Omega}_\Lambda}}\sinh\left(\frac{3}{2}\sqrt{\tilde{\Omega}_\Lambda}(1-\beta)(1-\nu)H_0 t\right)\right]^{\frac{2}{3(1-\beta)(1-\nu)}}, \tag{20}$$

where $\tilde{\Omega}_m = 1 - \frac{\Omega_\Lambda - \nu}{1-\nu}$ and $\tilde{\Omega}_\Lambda = \frac{\Omega_\Lambda - \nu}{1-\nu}$ are the effective fractional density parameters.

To discuss the decelerating and accelerating phases of the expansion of the Universe, along with the transition, we study the cosmological parameter known as the *deceleration parameter*, q, which is defined as:

$$q = -\frac{\ddot{a}}{aH^2} = -\left(1 + \frac{\dot{H}}{H^2}\right). \tag{21}$$

The equation (19) is used for obtaining an expression for the above, and it comes out to be:

$$q(z) = -1 + \frac{3}{2}\frac{(1-\beta)(1-\nu)\tilde{\Omega}_m(1+z)^{3(1-\beta)(1-\nu)}}{\tilde{\Omega}_m(1+z)^{3(1-\beta)(1-\nu)} + \tilde{\Omega}_\Lambda}. \tag{22}$$

The value of the deceleration parameter at present time, q_0 is given by

$$q_0 = -1 + \frac{3}{2}(1-\beta)(1-\Omega_\Lambda), \tag{23}$$

In late times, as we approach the present epoch $z = 0$, the deceleration parameter transitions from a positive to a negative value. This can be probed with another quantity called the *transitional redshift*. The expression for it is given by

$$z_{\rm tr} = -1 + \left[\frac{2(\Omega_\Lambda - \nu)}{(1-\Omega_\Lambda)\left(3(1-\beta)(1-\nu) - 2\right)}\right]^{\frac{1}{3(1-\beta)(1-\nu)}}. \tag{24}$$

Another key quantity that helps us understand the dynamics and the evolution of content in the Universe is the *effective equation of state* parameter $w_{\rm eff}$. In this work, we have assumed that the decaying vacuum has an EoS parameter equal to -1. We will now compute the model's effective EoS, which includes both decaying VED and matter with dark matter creation. This is calculated by the expression,

$$w_{\rm eff} = -1 - \frac{2}{3}\frac{\dot{H}}{H^2}. \tag{25}$$

On substituting the relevant values into (25), we determine the effective EoS parameter to be as follows:

$$w_{\rm eff}(z) = -1 + \frac{(1-\beta)(1-\nu)\tilde{\Omega}_m(1+z)^{3(1-\beta)(1-\nu)}}{\tilde{\Omega}_m(1+z)^{3(1-\beta)(1-\nu)} + \tilde{\Omega}_\Lambda}, \tag{26}$$

The present value of $w_{\rm eff}$ which we denote by w_0, is given by

$$w_0 = -1 + (1-\beta)(1-\Omega_\Lambda). \tag{27}$$

IV. PERTURBATIONS IN THE COSMIC BACKGROUND

This section gives a general overview of cosmic perturbation theory in the context of a matter creation cosmology with a decaying vacuum. Studying cosmic perturbations is crucial, as they underpin the formation of large-scale structures in the Universe—such as stars, galaxies, clusters, and quasars. The reader is referred to the Refs.[81–84] for a detailed discussion on the theory of cosmological perturbations and the derivation of the perturbation equations. Here, we present only the basic equations and terms that are appropriate and relevant to our discussion.

The differential equation for the matter density contrast, defined as $\delta_m \equiv \delta\rho_m/\rho_m$, in the considered model can be approximated as follows [85]

$$\delta_m'' + \left(3 + a\frac{H'(a)}{H(a)}\right)\frac{\delta_m'}{a} - \frac{3}{2}\Omega_m(a)\frac{\delta_m}{a^2} = 0, \tag{28}$$

where 'prime' denotes differentiation with respect to the scale factor a. This second-order differential equation describes the evolution of the density contrast, as the primary effects arise from the specific form of the Hubble function.

To solve (28), we use the initial condition given in Ref.[86] as

$$\delta_{mi} = 1.5\times10^{-4} \quad \text{at} \quad a_i = 10^{-3}, \quad \text{and} \quad \delta_{mi}' = \delta_{mi}/a_i. \tag{29}$$

The expression for $\Omega_m(a)$ in terms of the model parameters can be derived through $\Omega_m(a) = 1 - \Omega_\Lambda(a)$, where

$$\Omega_\Lambda(a) = \frac{\rho_\Lambda}{3H^2} = \frac{\Omega_\Lambda - \nu}{E(a)^2} + \nu. \tag{30}$$

Here $E(a) = H(a)/H_0$ is the dimensionless Hubble parameter in terms of the scale factor $a = (1+z)^{-1}$. We substitute the required values into (28), allowing us to numerically compute the matter density contrast δ_{mi} using the initial conditions (29). Next, we consider the linear growth rate of δ_{mi}. This quantity is denoted by f, and in the linear theory, it is related to the peculiar velocity [87]. We define this as

$$f(a) = \frac{d\ln D_m(a)}{d\ln a}, \tag{31}$$

where $D_m(a) = \frac{\delta_m(a)}{\delta_m(a=1)}$ represents the linear growth function. The quantity denoted by $f\sigma_8$, called the 'weighted linear growth rate', is obtained by multiplying the linear growth rate $f(z)$, defined in (31), and $\sigma_8(z)$. Here, σ_8 is a parameter that measures the density root-mean-square fluctuation on $8h^{-1}$ Mpc scales [88, 89]. It is defined as [90]

$$\sigma_8(z) = \frac{\delta_m(z)}{\delta_m(z=0)}\sigma_8(z=0). \tag{32}$$

Using (31) and (32), the weighted linear growth rate is computed by the following relation:

$$f\sigma_8(z) = -(1+z)\frac{\sigma_8(z=0)}{\delta_m(z=0)}\frac{d\delta_m}{dz}. \tag{33}$$

V. DATA AND METHODOLOGY

In this section, we consider the well-established and most recent observational samples to constrain the model parameters. We use the Bayesian Markov Chain Monte Carlo (MCMC) method for statistical analysis, which is conducted using the *emcee* package in Python. Optimization of the best fit of parameters is done by maximization of the probability function given by

$$\mathcal{L} \propto \exp(-\chi^2/2), \tag{34}$$

where χ^2 is the chi-squared function. We enlist the χ^2 function for different data samples in the following subsections.

A. Sample of Type Ia supernovae (SNIa)

Often referred to as "standard candles," SNIa are among the most popular cosmological probes used to study the Universe's evolution. These supernovae are extremely luminous, with their peak brightness often rivalling that of their host galaxy [91]. Increasingly precise data from SNIa have solidified the original ground-breaking observations and their implications for the expansion of the universe[92–96].

The comprehensive Pantheon sample, which includes luminosity distance measurements of 1048 supernovae

across a redshift range of $0.01 < z < 2.26$, is used in this analysis. The observations allow us to constrain cosmological models using the observed distance modulus, $\mu_{\rm obs}$. The χ^2 of the SNIa dataset is defined as

$$\chi^2_{\rm SNIa} = \sum_{i=1}^{1048} \Delta\mu^T C^{-1} \Delta\mu, \tag{35}$$

where $\Delta\mu = \mu_{\rm obs} - \mu_{\rm th}$. Here, $\mu_{\rm obs}$ represents the observed distance modulus, defined as

$$\mu_{\rm obs} = m_B - M, \tag{36}$$

where m_B denotes the observed peak magnitude in the rest-frame B-band, and M gives the absolute B-band magnitude of a fiducial SNIa. We take the value of the magnitude M to be $M = -19.38$. The theoretical distance modulus, $\mu_{\rm th}$, is defined as

$$\mu_{\rm th}(z, p) = 5\log_{10}\left(\frac{D_L(z_{\rm hel}, z_{\rm cmb})}{1\ {\rm Mpc}}\right) + 25, \tag{37}$$

where p represents the parameter space, which varies across different cosmological models, and D_L is the luminosity distance, expressed as

$$D_L(z_{\rm hel}, z_{\rm cmb}) = (1 + z_{\rm hel}) r(z_{\rm cmb}). \tag{38}$$

The comoving distance $r(z)$ is calculated as

$$r(z) = cH_0^{-1} \int_0^z \frac{dz'}{E(z', p)}, \tag{39}$$

where $E(z) = H(z)/H_0$ is the dimensionless Hubble parameter, c is the speed of light in vacuum, and H_0 is the Hubble constant. The CMB-frame and heliocentric redshifts are denoted by $z_{\rm cmb}$ and $z_{\rm hel}$, respectively. The covariance matrix C used in the χ^2 calculation is given by

$$C = D_{\rm stat} + C_{\rm sys}, \tag{40}$$

where $D_{\rm stat}$ is a diagonal matrix representing statistical uncertainties, and $C_{\rm sys}$ is the covariance matrix accounting for systematic uncertainties.

B. Baryon Acoustic Oscillations

The measurement of the background expansion history of the Universe can be achieved through the standard ruler provided by the imprint of Baryon Acoustic Oscillations (BAO). The BAO method has been extensively studied and validated in numerous works [97–99], providing a critical tool for constraining cosmological models. For a comprehensive discussion, the reader is referred to these studies. We present the basic equations and terms necessary for working with BAO measuremens.

The comoving sound horizon is given by

$$r_s(z) = \frac{c}{H_0} \int_z^\infty \frac{c_s}{E(z')} dz', \tag{41}$$

where c is the speed of light in vacuum and c_s is the speed of sound, given by

$$c_s = \frac{1}{\sqrt{3(1 + R_b a)}}, \tag{42}$$

where R_b gives the baryon to photon density ratio, calculated as

$$R_b a = \frac{3\rho_b}{4\rho_\gamma}, \quad R_b = 31500\,\omega_b \left(\frac{T_{\rm CMB}}{2.7\ {\rm K}}\right)^{-4}. \tag{43}$$

Here, $\omega_b = \Omega_b h^2$ is the baryon density parameter with $h = H_0/100$, and $T_{\rm CMB}$ is the current measured value of the CMB temperature.

One crucial period in our cosmic history is the baryon drag epoch, which refers to the time when baryons decoupled from photons. The redshift at drag epoch is calculated by the fitting function:

$$z_d = \frac{1345\,\omega_m^{0.251}}{1 + 0.659\,\omega_m^{0.828}} \left(1 + b_1\,\omega_b^{b_2}\right), \tag{44}$$

where $b_1 = 0.313\,\omega_m^{-0.419}\left(1 + 0.607\,\omega_m^{0.674}\right)$ and $b_2 = 0.238\,\omega_m^{0.223}$. Here $\omega_m = \Omega_m h^2$ is the matter density parameter. By evaluating Eq. (41) at the drag epoch redshift z_d, we obtain a key quantity r_d which is a crucial term in BAO measurements. This gives the maximum distance sound waves could travel in the early Universe till the time of the last scattering (before the photon-baryon decoupling).

We use the latest BAO measurements from the Dark Energy Spectroscopic Instrument (DESI). The 12 data points used have been shown in Table I. The complete data, along with the survey details, can be found in Ref.[100, 101]. We constrain the model parameters from the BAO data using three observable within the redshift range $0.295 \le z \le 2.330$.

The first observable is the comoving angular distance $D_M(z)$ which is given by

$$D_M(z) = \int_0^z \frac{c}{H(z')}\, dz'. \tag{45}$$

Then we have the Hubble distance parameter $D_H(z)$ which is given by

$$D_H(z) = \frac{c}{H(z)}. \tag{46}$$

From the former two, we calculate the angle-average distance, which is given by

$$D_V(z) = (z D_M(z)^2 D_H(z))^{1/3}. \tag{47}$$

These quantities are normalized by the sound horizon r_d, which provides a standard ruler for measuring on large

TABLE I. Summary of tracer redshift and other observable parameters from the latest DESI BAO 2024 survey. The complete data is available at [101]

Tracer	**Redshift Range**	$z_{\mathbf{eff}}$	D_M/r_d	D_H/r_d	D_V/r_d
BGS	0.1 - 0.4	0.295	—	—	7.93 ± 0.15
LRG1	0.4 - 0.6	0.510	13.62 ± 0.25	20.98 ± 0.61	—
LRG2	0.6 - 0.8	0.706	16.85 ± 0.32	20.08 ± 0.60	—
LRG3+ELG1	0.8 - 1.1	0.930	21.71 ± 0.28	17.88 ± 0.35	—
ELG2	1.1 - 1.6	1.317	27.79 ± 0.69	13.82 ± 0.42	—
QSO	0.8 - 2.1	1.491	—	—	26.07 ± 0.67
Lya QSO	1.77 - 4.16	2.330	39.71 ± 0.94	8.52 ± 0.17	—

scales. Thus, we calculate our theoretical values against the observed values $D_M/r_d, D_H/r_d$ and D_V/r_d obtained from the BAO measurements, which can be referred to from Table I. The χ^2 functions for the BAO data are:

$$\chi^2_{D_M/r_d} = \sum_{i=1}^{5} \frac{\left(D_M(z_i,p) - D_{M,\mathrm{obs}}(z_i)\right)^2}{\sigma^2_{D_M}(z_i)}, \tag{48}$$

$$\chi^2_{D_H/r_d} = \sum_{i=1}^{5} \frac{\left(D_H(z_i,p) - D_{H,\mathrm{obs}}(z_i)\right)^2}{\sigma^2_{D_H}(z_i)}, \tag{49}$$

$$\chi^2_{D_V/r_d} = \sum_{i=1}^{2} \frac{\left(D_V(z_i,p) - D_{V,\mathrm{obs}}(z_i)\right)^2}{\sigma^2_{D_V}(z_i)}, \tag{50}$$

where $D_M(z_i,p), D_H(z_i,p)$ and $D_V(z_i,p)$ are the theoretical values of the observable based on the model, and $D_{M,\mathrm{obs}}$, $D_{H,\mathrm{obs}}$ and $D_{V,\mathrm{obs}}$ are the observed values of the DESI BAO measurements.

Thus, the total χ^2 value for the BAO data is

$$\chi^2_{BAO} = \chi^2_{D_M/r_d} + \chi^2_{D_H/r_d} + \chi^2_{D_V/r_d}. \tag{51}$$

C. Cosmic Microwave Background

The Cosmic Microwave Background (CMB) measurements of the locations and amplitudes of peaks of acoustic oscillations give information about the distance to the surface of last scattering given by $r(z_*)$. This can be evaluated through Eq. (39) at z_*, which is the redshift at the time of recombination or last scattering. It marks the epoch when photons last scattered off free electrons before being able to freely travel through the universe, leaving behind their imprint in the form of the CMB radiation. The reader is referred to the Refs. [102, 103] for a more detailed discussion on CMB. We now proceed with the basic equations employed in the data analysis. The fitting function for recombination redshift z_* is

$$z_* = 1048\left(1 + 0.00124\,\omega_b^{-0.738}\right)\left(1 + g_1\,\omega_m^{g_2}\right), \tag{52}$$

where

$$g_1 = \frac{0.0783\,\omega_b}{1 + 39.5\,\omega_b^{0.763}}, \quad g_2 = \frac{0.560}{1 + 21.1\,\omega_b^{1.81}}. \tag{53}$$

Here ω_m and ω_b are the matter and baryon density parameters respectively. We are particularly interested in two quantities that characterize the CMB spectrum: the shift parameter and acoustic length scale.

It is observed that through the acoustic oscillations, doppler peaks are produced in the photon (radiation) spectrum. The placement of the peaks in the CMB spectrum is affected by the matter and radiation density, along with the dark energy. The shift in these peaks is probed by the shift parameter $\mathcal{R}$. The spectrum is affected in the line-of-sight direction, leading to the peak heights. We obtain $\mathcal{R}$ through the relation:

$$\mathcal{R} = \sqrt{\Omega_m} H_0 \frac{r(z_*)}{c}. \tag{54}$$

The acoustic scale l_a determines the spacing between the peaks and characterizes the CMB temperature power spectrum in the transverse direction. This is calculated by

$$l_A = \frac{\pi r(z_*)}{r_s(z_*)}. \tag{55}$$

When working with CMB, we have to consider the radiation component as in early-times it was dominant. We use the relation $\Omega_r = \Omega_m/(1+z_{\mathrm{eq}})$ to determine the radiation density, where z_{eq} is the epoch of matter-radiation equality and is given by

$$z_{\mathrm{eq}} = 2.5 \times 10^4\,\omega_m \left(\frac{T_{\mathrm{CMB}}}{2.7\,\mathrm{K}}\right)^{-4}, \tag{56}$$

where T_{CMB} represents the present-day temperature of the CMB. We adopt the best-fit value $T_{\mathrm{CMB}} = 2.7255\,\mathrm{K}$, measured by the COBE-FIRAS instrument [104]. CMB observation contribution in the constraint of cosmological parameters can be converted to CMB shift parameters for breaking degeneracy of parameters, instead of using the entire CMB power spectrum, [105, 106] This allows us to consider the data vector $\mathbf{v} = (R, l_A, \omega_b)^T$ which summarizes the Planck 2018 TT,TE,EE+lowE results [107]. The vector $\mathbf{v}$, along with the covariance matrix $C_{\mathbf{v}}$ have been reported in the Ref. [108].

The χ^2 function for the CMB data is given by:

$$\chi^2_{\mathrm{CMB}} = (\mathbf{v} - \mathbf{v}_{\mathrm{th}})^T C_{\mathbf{v}}^{-1} (\mathbf{v} - \mathbf{v}_{\mathrm{th}}), \tag{57}$$

where $\mathbf{v}_{\rm th}$ is the vector for the theoretical values of R, l_A, ω_b; and $C_{\mathbf{v}}^{-1}$ is the inverse of the covariance matrix.

D. Sample of Cosmic Chronometer (CC)

The cosmic chronometer (CC) approach provides measurements of the Hubble parameter based on the galaxy differential age method using the largest galaxies that are passively evolving [109]. In our analysis, we use 32 CC data points from different surveys that span a redshift range of $0.07 \leq z \leq 1.965$. The χ^2 function for CC data is given by

$$\chi^2_{CC} = \sum_{i=1}^{32} \frac{\left[H(z_i,p) - H_{\rm obs}(z_i)\right]^2}{\sigma_H^2(z_i)}, \tag{58}$$

where $H(z_i,p)$ represents the theoretical value of the Hubble parameter for the model parameters, $H_{\rm obs}(z_i)$ is the observed value of the Hubble parameter at redshift z_i, and $\sigma_H(z_i)$ is the associated uncertainty with $H_{\rm obs}(z_i)$.

E. $f(z)\sigma_8(z)$ Data

We introduced the background evolution of matter perturbations in Section 4 and used Eq.(33) to define the weighted linear growth rate. To provide a more comprehensive analysis of the present model in the context of perturbation evolution, we focus on the observable quantity $f(z)\sigma_8(z)$. For this purpose, we use 18 data points from the "Gold" compilation, which contains measurements of $f(z)\sigma_8(z)$ obtained from various galaxy surveys. The data has been summarized in Table III of Ref. [110].

We compare the observational data with the theoretical values calculated from our models using the χ^2 function for $f(z)\sigma_8(z)$, which is defined as

$$\chi^2_{f\sigma_8} = \sum_{i=1}^{18} \frac{\left[f\sigma_{8,\rm theo}(z_i,p) - f\sigma_{8,\rm obs}(z_i)\right]^2}{\sigma^2_{f\sigma_8}(z_i)}, \tag{59}$$

where $f\sigma_{8,\rm theo}(z_i,p)$ is the theoretical value computed using (33), and $f\sigma_{8,\rm obs}(z_i)$ represents the observed data.

Using three different combinations of the data discussed above, we employ Bayesian MCMC methods to explore the parameter space of the proposed model described in Section III, using the `emcee` Python package [111, 112]. We use the following chi-squared functions:

- **DS1**: This dataset consists of a combination of data from the observations of SNIa, BAO, and CMB. It contains 1063 independent data points. The total χ^2 function for this dataset is given by

$$\chi^2_{\rm tot} = \chi^2_{\rm SNIa} + \chi^2_{\rm BAO} + \chi^2_{\rm CMB}. \tag{60}$$

- **DS2**: The first dataset is augmented with the cosmic chronometer (CC) data containing 32 data points making the total to be 1095. In this case, the total χ^2 function is defined as

$$\chi^2_{\rm tot} = \chi^2_{\rm SNIa} + \chi^2_{\rm BAO} + \chi^2_{\rm CMB} + \chi^2_{H(z)}. \tag{61}$$

- **DS3**: We further expand DS2 by including the weighted linear growth rate data from the $f\sigma_8$ measurements. The total data points are 1113 and for this combination, the total χ^2 function becomes

$$\chi^2_{\rm tot} = \chi^2_{\rm SNIa} + \chi^2_{\rm BAO} + \chi^2_{\rm CMB} + \chi^2_{H(z)} + \chi^2_{f\sigma_8}. \tag{62}$$

VI. RESULTS AND DISCUSSION

We now discuss over the outcomes from various data sets to constrain our model with Hubble parameter, given in Eq.(19). For the parameters of the ΛCDM and the $\Lambda(t)+\Gamma$ models, Figures 1-3 display the combined $1-\sigma(68.3\%)$ and $2-\sigma(95.4\%)$ confidence level (CL) contours with marginalized likelihood distributions. Table II displays the matching parameter constraints for the datasets DS1, DS2, and DS3 that were acquired from the MCMC sampler. The best fit values of Hubble constant for $\Lambda(t)+\Gamma$ come out to be $H_0 = 67.517 \pm 0.869$ km/s/Mpc for DS1 data, $H_0 = 67.534 \pm 0.874$ km/s/Mpc for DS2 data, and $H_0 = 67.533 \pm 0.884$ km/s/Mpc for DS3 data. We know that the discrepancies in the measurement of H_0 between the Planck collaboration [113] and independent cosmological probes [114] refer to H_0 tension. The Planck collaborator estimates $H_0 = 67.27 \pm 0.06$ km/s/Mpc, whereas R22 reported $H_0 = 73.04 \pm 1.04$ km/s/Mpc which leads to a tension of 4.89σ. Our estimated values of H_0 from three the datasets show very little tension with the Planck result upon comparison. We plot the evolution of $H(z)$ for ΛCDM and decaying VED with matter creation model using 32 data points of CC as shown in Figs. 4-6. Both models are in agreement for the three datasets, as the hybrid model and the standard model coincide for nearly the whole evolution period and the model paths cover most of the CC dataset with the error bars.

In Figs. 7-9, we present the evolution of the deceleration parameter with respect to redshift. The transition redshift $z_{\rm tr}$ at which the Universe transits from a decelerating phase to an acceleration phase occurs at $z_{\rm tr} = 0.675 \pm 0.019$ with DS1, $z_{\rm tr} = 0.675 \pm 0.019$ with DS2, and at $z_{\rm tr} = 0.677 \pm 0.019$ with DS3 which doesn't vary much from the standard model but the transition happens earlier. The analysis shows that the Universe is currently in an accelerating phase of cosmic expansion, whereas it was in a decelerating phase in the past. We find the constraints of the present value of the deceleration parameter as $q_0 = -0.550 \pm 0.010$, -0.550 ± 0.010 and -0.551 ± 0.010 for DS1, DS2 and DS3 datasets, respectively. Comparing with the values of q_0 obtained for

TABLE II. Constraints on parameters for ΛCDM and $\Lambda(t)+\Gamma$ model for different observational data sets.

Parameter	**DS1**		**DS2**		**DS3**	
	ΛCDM	$\Lambda(t)+\Gamma$	ΛCDM	$\Lambda(t)+\Gamma$	ΛCDM	$\Lambda(t)+\Gamma$
H_0	67.278 ± 0.355	67.517 ± 0.869	67.301 ± 0.353	67.534 ± 0.874	67.993 ± 0.414	67.533 ± 0.884
Ω_Λ	0.68362 ± 0.00482	0.69973 ± 0.00685	0.68398 ± 0.00479	0.69989 ± 0.00684	0.69337 ± 0.00540	0.70035 ± 0.00680
ν	-	0.00069 ± 0.00692	-	0.00058 ± 0.00687	-	0.00067 ± 0.00694
β	-	0.00109 ± 0.00694	-	0.00122 ± 0.00685	-	0.00120 ± 0.00694
ω_b	0.02234 ± 0.00013	0.02237 ± 0.00014	0.02235 ± 0.00013	0.02237 ± 0.00015	0.02247 ± 0.00013	0.02237 ± 0.00014
σ_8	-	-	-	-	0.755 ± 0.026	0.762 ± 0.028
S_8	-	-	-	-	0.763 ± 0.026	0.762 ± 0.027
$z_{\rm tr}$	0.629 ± 0.012	0.675 ± 0.019	0.630 ± 0.012	0.675 ± 0.019	0.654 ± 0.014	0.677 ± 0.019
q_0	-0.525 ± 0.007	-0.550 ± 0.010	-0.526 ± 0.007	-0.550 ± 0.010	-0.540 ± 0.008	-0.551 ± 0.010
w_0	-0.684 ± 0.005	-0.700 ± 0.007	-0.684 ± 0.005	-0.700 ± 0.007	-0.684 ± 0.005	-0.700 ± 0.007
t_0 (Gyr)	13.80 ± 0.02	13.98 ± 0.18	13.80 ± 0.02	13.98 ± 0.19	13.78 ± 0.02	13.99 ± 0.19

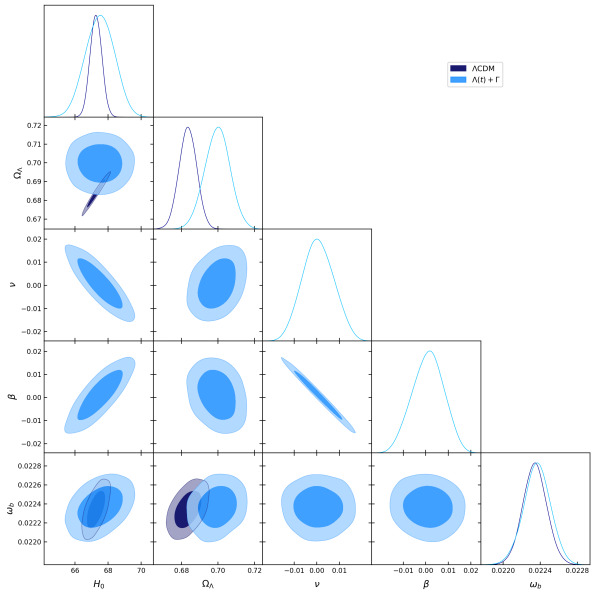


FIG. 1. The one-dimensional marginalized distribution and two-dimensional contour plots for ΛCDM and $\Lambda(t)+\Gamma$ models at 68.3% and 95.4% confidence levels using DS1 dataset.

the ΛCDM model, we see that the expansion of the Universe is accelerating at a slightly faster rate. Overall, the values of both q_0 and $z_{\rm tr}$ obtained for the model agree with the values obtained for the standard model.

In the considered cosmological scenario, the calculated values of the age of the Universe t_0 based on the parameter constraints obtained for the three datasets are 13.98 ± 0.18 Gyr, 13.98 ± 0.19 Gyr, and 13.99 ± 0.19 Gyr, respectively which when compared with the mean value for the standard model 13.79 ± 0.02 Gyr suggest a longer cosmic history however the difference is not that significant. This can be attributed to the perturbations through matter creation and vacuum decay. In this study, we have initially assumed the EoS of vacuum energy as $w_\Lambda = -1$. Therefore, we investigate the total effective EoS given by (26). The evolution of EoS is depicted in Figs. 10-12. It is obvious from the figures that $w_{\rm eff}$ has a negative value such that $w_{\rm eff} < -1/3$, which indicates an accelerating phase in the expansion. In the future, we can see that the model asymptotically approaches negative unity, i.e. $w_{\rm eff} = -1$, implying complete vacuum domination. The present values of $w_{\rm eff}$, represented by w_0, strongly favor a value around -0.700 ± 0.007 across all three datasets for the presented model which indicates the transition towards a dark energy era. The values are also consistent with those reported for the ΛCDM model.(cf. Table II)

Let us now discuss σ_8 and S_8, which are crucial for structure formation and are obtained by probing cosmological perturbations, as discussed in Section 4. The DS3 column in Table II reports the parameter constraints for these quantities. For the $\Lambda(t)+\Gamma$ model, we have $\sigma_8 = 0.762 \pm 0.028$, which is a good result considering the slight deviation with ΛCDM. We rephrase this in terms of the fitting value of the related LSS observable $S_8 = \sigma_8\sqrt{\Omega_m}/0.3$, which gives $S_8 = 0.762 \pm 0.0297$. Comparing it with $S_8 = 0.763 \pm 0.026$ of ΛCDM model for the same DS3 dataset and $S_8 = 0.737 \pm 0.036$ of weak

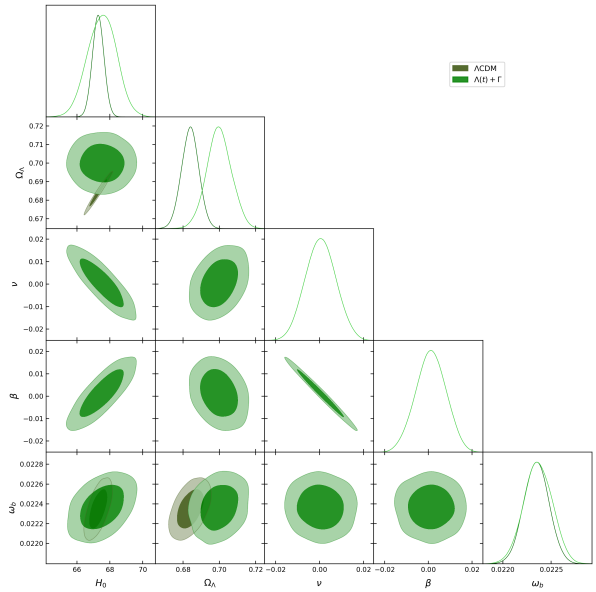


FIG. 2. The one-dimensional marginalized distribution and two-dimensional contour plots for ΛCDM and $\Lambda(t) + \Gamma$ models at 68.3% and 95.4% confidence levels using DS2 dtatset.

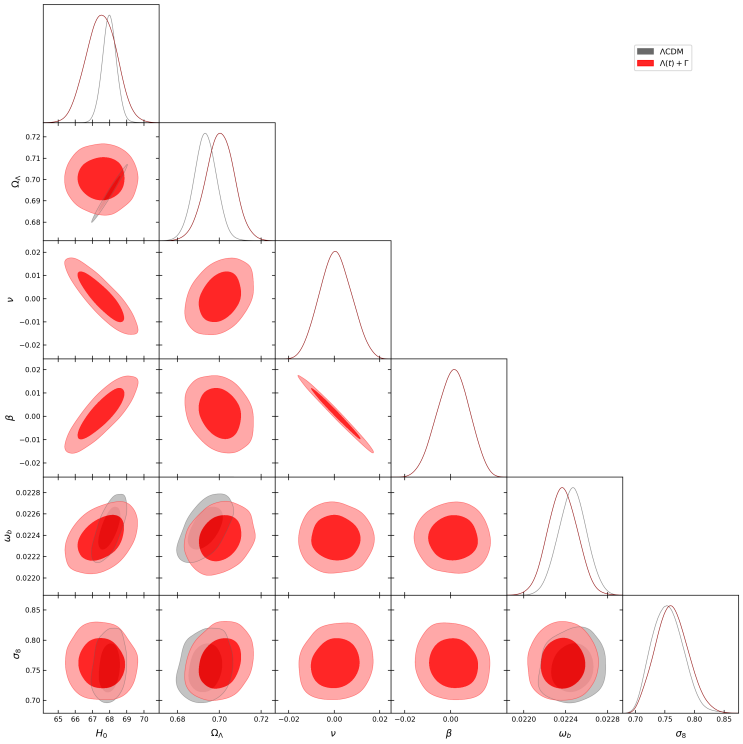


FIG. 3. The one-dimensional marginalized distribution and two-dimensional contour plots for ΛCDM and $\Lambda(t) + \Gamma$ models at 68.3% and 95.4% confidence levels using DS3 dataset.

lensing measurement, we find that there is no residual tension left as mismatch is less than 1σ. We have also plotted $f(z)\sigma_8(z)$ data corresponding to both the standard model and our model in Fig. 13. We have used 18 data points from the $f(z)\sigma_8(z)$ Gold data along with their uncertainty measures. The curves provide a fairly good fit for the data, implying our model agrees with the observational data in the context of structure formation and cosmic perturbations.

We have found that the best-fit values of decaying pa-

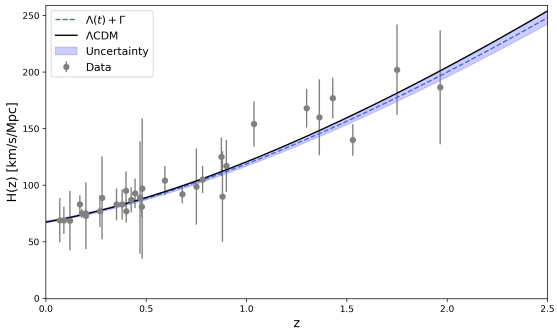


FIG. 4. Best-fit using DS1 over H(z) data for the $\Lambda(t)+\Gamma$ model (blue dotted line) and standard model (black line). The shaded band represents the uncertainty for the former. The grey points with error bar represent the 32 CC data points.

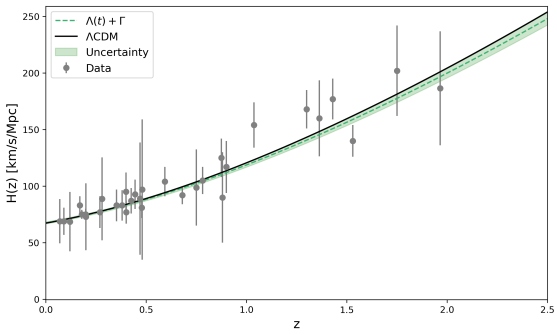


FIG. 5. Best-fit using DS2 over H(z) data for the $\Lambda(t)+\Gamma$ model (green dotted line) and standard model (black line). The shaded band represents the uncertainty for the former. The grey points with error bar represent the 32 CC data points.

rameter are $\nu = 0.00069 \pm 0.00692$, $\nu = 0.00058 \pm 0.00687$ and $\nu = 0.00067 \pm 0.00694$, and of the matter creation parameter are $\beta = 0.00109 \pm 0.00694$, $\beta = 0.00122 \pm 0.00685$ and $\beta = 0.00120 \pm 0.00694$ using DS1, DS2 and DS3 datasets, respectively. The results confirm that these parameters act as a perturbation to the ΛCDM model.

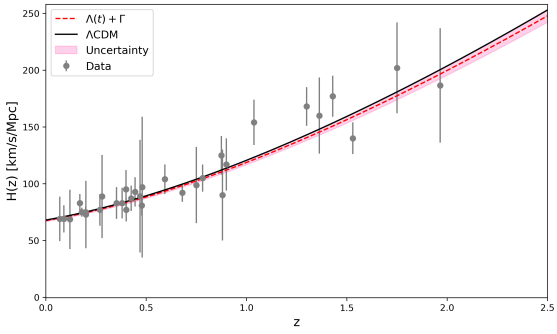


FIG. 6. Best-fit using DS3 over H(z) data for the $\Lambda(t)+\Gamma$ model (red dotted line) and standard model (black line). The shaded band represents the uncertainty for the former. The grey points with error bar represent the 32 CC data points.

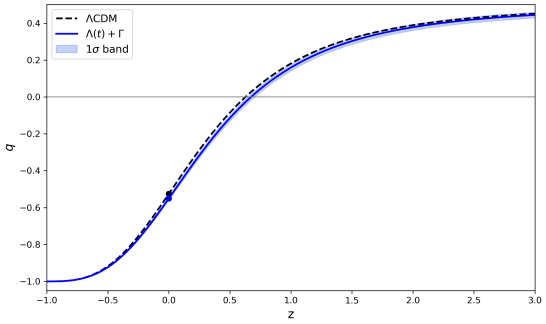


FIG. 7. The deceleration parameter versus redshift for the $\Lambda(t)+\Gamma$ model using DS1 dataset. The standard model's evolution is depicted by the dashed curve.

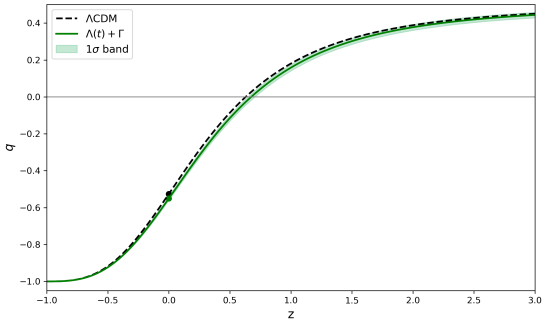


FIG. 8. The deceleration parameter versus redshift for the $\Lambda(t)+\Gamma$ model with DS2 dataset. The standard model's evolution is depicted by the dashed curve.

The values obtained for the baryon density parameter $\omega_b = \Omega_b h^2$ are also in good agreement with the Planck values across the three datasets.

Furthermore, we report the cosmographic parameter, known as the jerk parameter, which is defined by $j = \dddot{a}/aH$. It gives information about the dynamics of the DE EoS. In terms of the deceleration parameter, it is given by the expression:

$$j = q(2q+1) + (1+z)\frac{dq}{dz}. \tag{63}$$

It can be verified that for the standard model, the value of the jerk parameter is a constant, i.e. $j = 1$. This parameter is a dimensionless third derivative of the scale factor and can help us study how our model deviates from the standard model. For the proposed model, the jerk parameter (63) gives

$$j = 3(1-\beta)(1-\nu)(q+1) - 3q - 2. \tag{64}$$

We see that in the limit where the model parameters β and ν are zero, j reduces to one. In the case of the deceleration parameter, negative values imply an

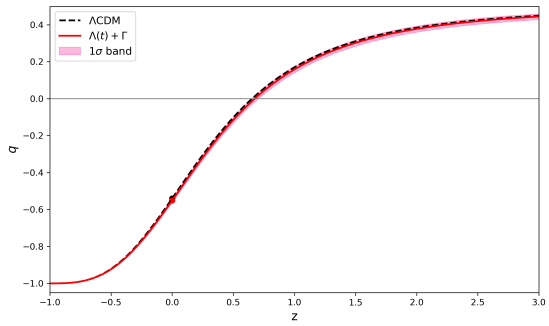


FIG. 9. The deceleration parameter versus redshift for the $\Lambda(t)+\Gamma$ model with DS3 dataset. The standard model's evolution is depicted by the dashed curve.

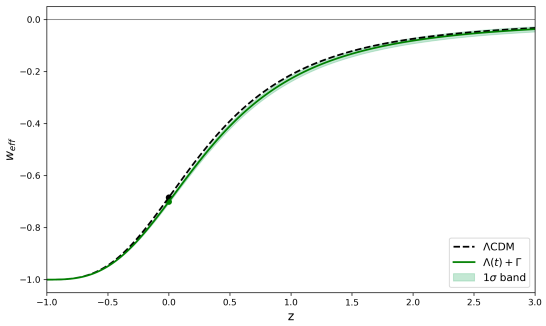


FIG. 11. Effective EoS parameter for the $\Lambda(t)+\Gamma$ model with DS2 as a function of redshift z. The standard model's evolution is depicted by the dashed curve.

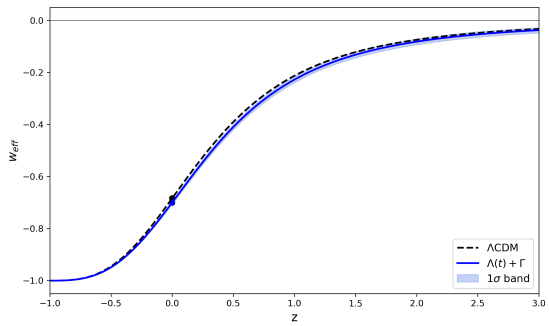


FIG. 10. Effective EoS parameter for the $\Lambda(t)+\Gamma$ model with DS1 as a function of redshift z. The standard model's evolution is depicted by the dashed curve.

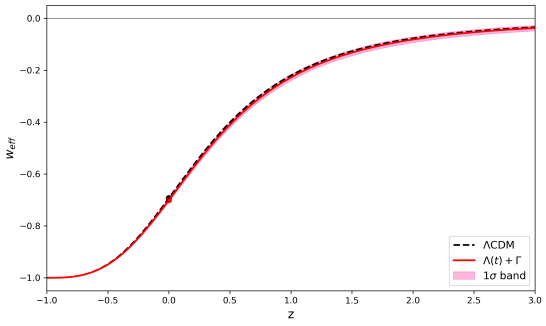


FIG. 12. Effective EoS parameter for the $\Lambda(t)+\Gamma$ model with DS3 as a function of redshift z. The standard model's evolution is depicted by the dashed curve.

accelerating expansion of the Universe. However, in the case of the jerk parameter, a positive value represents the same, and from (64), we can see that for the considered model, the value will remain positive. In Fig.14, the evolution of this parameter is shown for our model using the parameter constraints obtained for all three datasets. The figure confirms the positive nature, and we see that it remains close to but less than unity at early redshifts, eventually tending to 1 in the future. Thus, in the past, the parameter deviates from the standard model; however, in late time, it converges to a similar behaviour.

We also plot the evolution of matter density in ΛCDM and the proposed model in Fig.15 using the cosmological parameter estimates of DS1, DS2 and DS3 datasets. We consider a single case for the evolution in the case of the standard model represented by a black line. We see that in late time, the matter density tends towards zero for all the cases, resulting from the accelerated expansion causing complete vacuum domination. We also observe that because of our particular selection of the particle production rate, the creation pressure p_c, which contributes to this acceleration, has the expression $p_c = -\beta\rho_m$. The evolution of this pressure is plotted in Fig.16, which shows that it increases from negative infinity at higher redshifts and converges towards zero in late time. Thus, particle production will cease in the dark energy domination era.

VII. SELECTION CRITERIA

In this section, we are interested in determining the models that could perform better with respect to the observational data used in the analysis. We utilize three selection statistical criteria, namely, reduced chi-squared, Akaike information criteria and Bayesian information criteria, which consider the presence of an extra degree of freedom.

Table III presents the χ^2 and the reduced χ^2 values for the standard model and the presented model, respectively. The reduced χ^2, denoted by χ^2_{red}, is calculated by the formula ${\chi^2}_{red} = \chi^2/(N-d)$, where N is the total number of data points (here, it is 1063, 1095 and 1113 for DS1, DS2 and DS3, respectively), and d is the number of

TABLE III. Comparison of χ^2, AIC, and BIC values for ΛCDM and the $\Lambda(t)+\Gamma$ model across different datasets.

Values	**DS1**		**DS2**		**DS3**	
	ΛCDM	$\Lambda(t)+\Gamma$	**ΛCDM**	$\Lambda(t)+\Gamma$	**ΛCDM**	$\Lambda(t)+\Gamma$
χ^2	1057.31	1057.00	1072.26	1071.72	1085.06	1084.58
d	3	5	3	5	4	6
N	1063	1063	1095	1095	1113	1113
$\chi^2_{\rm red}$	0.997	0.999	0.982	0.983	0.978	0.980
AIC	1063.33	1067.06	1078.29	1081.78	1093.10	1096.65
BIC	1078.22	1091.85	1093.26	1106.72	1113.12	1126.66
ΔAIC	—	3.73	—	3.49	—	3.55
ΔBIC	—	13.63	—	13.46	—	13.54

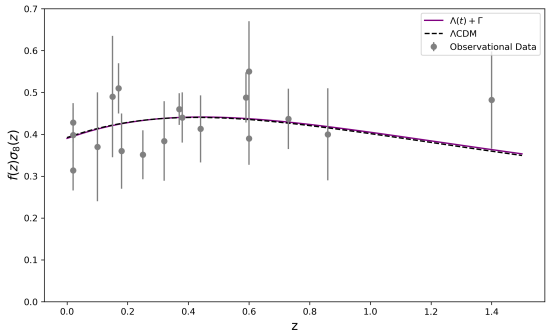


FIG. 13. Theoretical curves for the $f(z)\sigma_8(z)$ data corresponding to the standard model and the $\Lambda(t)+\Gamma$ model obtained from the parameter constraints for the cosmological parameters.

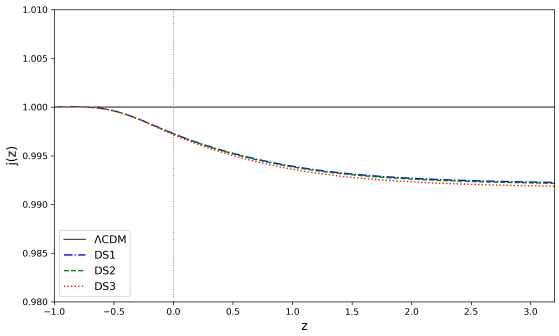


FIG. 14. Jerk parameter j with redshift z using parameter constraints obtained for the $\Lambda(t)+\Gamma$ model. The black line depicts the value for the standard model, i.e. 1.

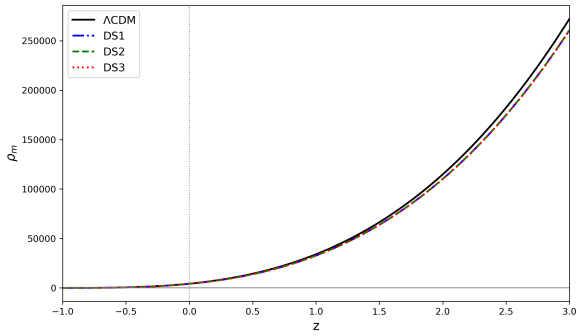


FIG. 15. Plot of matter density with the redshift z for the $\Lambda(t)+\Gamma$ corresponding to the three combined datasets. The black line represents the evolution of the standard model.

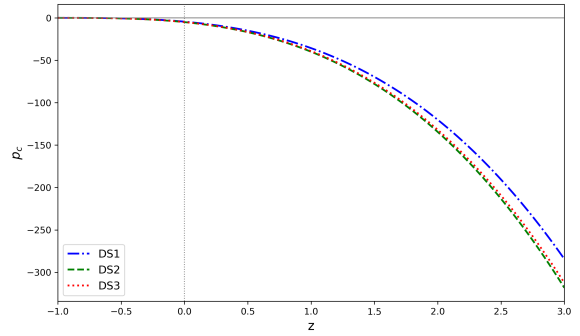


FIG. 16. Evolution of the creation pressure with redshift z corresponding to the three different datasets. The pressure eventually tends towards zero in the future.

free parameters ($d=3$ for ΛCDM and $d=5$ for $\Lambda(t)+\Gamma$ model for DS1 and DS2, and one additional parameter σ_8 is added to both for DS3). It is to be noted that a reduced χ^2 value less than one gives a good fit, and greater than one indicates a poor fit. From Table III, we see that the reduced χ^2 values are less than 1, indicating that our model fits well with the observational datasets and the data fits well with the current models. This is just one criterion to test the viability of our model. We will utilize two selection criteria for further analysis.

The Akaike information criteria (AIC) value is calculated as [115]

$$AIC = \chi^2_{min} + \frac{2d\,N}{N-d-1}, \tag{65}$$

where χ^2_{min} is the minimum value of the χ^2-function. Equation(65) makes it clear that the penalty for the extra number of free parameters is represented by the second term on the right hand side. In cases where d is significantly lesser than N, Eq.(65) reduces to $AIC = \chi^2_{min}+2d$. To test the proposed model M_1 against the reference model M_0, we calculate the difference in their respective AIC values as $\Delta AIC_{M_1M_0} = AIC_{M_1} - AIC_{M_0}$,

which is interpreted as the "evidence in favour" of the model M_1 as compared to the reference M_0. In this paper, the concordance model has been considered as the reference.

The value $0 \leq AIC_{M_1M_0} < 2$ indicates "strong evidence in favour" of the proposed model M_1. For $2 \leq AIC_{M_1M_0} \leq 4$, there is "averagely strong evidence in favour" of M_1. If $4 < AIC_{M_1M_0} \leq 7$, then there is "little evidence in favour", whereas for $AIC_{M_1M_0} > 7$, there is "no evidence in favour" of the proposed model.

On the other hand, the Bayesian information criterion (BIC) is defined as [116]

$$BIC = \chi^2_{min} + d\, lnN \tag{66}$$

Likewise $\Delta AIC_{M_1M_0}$, we compute $\Delta BIC_{M_1M_0}$ as $\Delta BIC_{M_1M_0} = BIC_{M_1} - BIC_{M_0}$. This gives the "evidence against the model M_1 with respect to the reference model M_0. Thus, the ranges of the values for BIC and their interpretation are different from those of AIC. The value $0 \leq BIC_{M_1M_0} < 2$ indicates "not enough evidence in favour" of the proposed model M_1. For $2 \leq BIC_{M_1M_0} < 6$, there is "evidence against" the model M_1, and for $6 \leq BIC_{M_1M_0} < 10$, we have "strong evidence against" model M_1. If $BIC_{M_1M_0} \geq 10$ then there is "overwhelmingly strong evidence" against the proposed model, and it is not the best model.

The values of $\Delta AIC_{M_1M_0}$ and $\Delta BIC_{M_1M_0}$ are shown in Table III. For the three datasets, DS1, DS2 and DS3, we obtain $\Delta AIC_{M_1M_0}$ values to be 3.73, 3.49, and 3.55, respectively, indicating averagely strong evidence in favour of our proposed model with respect to the standard model for all the datasets. The $\Delta BIC_{M_1M_0}$ values obtained for the datasets are 13.63, 13.46, and 13.54, respectively. This indicates that the evidence against our model is overwhelmingly strong and it is not the best fit for the data according to BIC.

VIII. THERMODYNAMIC ANALYSIS

In this section, we focus on the thermodynamic analysis of a universe under the proposed scenario. First we use the generalized first law of thermodynamics to obtain an expression for the dark matter temperature in terms of our model parameters. Then, we test the validity of generalized second law (GSL) of thermodynamics by computing the total entropy of the universe and its evolution w.r.t time and assess the results.

A. Dark Matter Temperature

Let us consider the modified first law of thermodynamics, given by Eq.(3). This can be rewritten as

$$d(\rho_m V) + \tilde{p}_m\, dV - \frac{(\rho_m + \tilde{p}_m)}{n} d(nV) = 0, \tag{67}$$

where $h = (\rho_m + \tilde{p}_m)$. Taking time derivative of Eq. (67), we have

$$\dot{\rho_m} = \frac{\dot{n}}{n}(\rho_m + \tilde{p}_m). \tag{68}$$

Since the temperature is defined by Gibbs equation

$$TdS = d\left(\frac{\rho_m}{n}\right) + \tilde{p}_m\, d\left(\frac{1}{n}\right) \tag{69}$$

We have $T = T(n, \rho_m)$ which gives

$$\dot{T} = \frac{\partial T}{\partial n}\dot{n} + \frac{\partial T}{\partial \rho_m}\dot{\rho_m}. \tag{70}$$

Using (68) into (70), we get

$$\dot{T} = \frac{\dot{n}}{n}\left[\frac{\partial T}{\partial n}n + \frac{\partial T}{\partial \rho_m}(\rho_m + \tilde{p}_m)\right]. \tag{71}$$

Using the integrability condition, $\partial^2 S/\partial T \partial n = \partial^2 S/\partial n \partial T$, one gets

$$\frac{\partial T}{\partial n}n + \frac{\partial T}{\partial \rho_m}(\rho_m + \tilde{p}_m) = T\frac{\partial \tilde{p}_m}{\rho_m}. \tag{72}$$

Using (72) into (71), we obtain the evolution equation for the temperature which is given by

$$\frac{\dot{T}}{T} = \frac{\dot{n}}{n}\frac{\partial \tilde{p}_m}{\partial \rho_m} = w_{\text{eff}}\frac{\dot{n}}{n}. \tag{73}$$

Using Eq.(2) in above equation, we get

$$\frac{\dot{T}}{T} = -3w_{\text{eff}}H\left(1 - \frac{\Gamma}{3H}\right), \tag{74}$$

or equivalently,

$$\frac{\dot{T}}{T} = 3\beta(1-\beta)H. \tag{75}$$

Using (20) into (75), we obtain the temperature of dark matter

$$T = T_0\left[\sqrt{\frac{\tilde{\Omega}_m}{\tilde{\Omega}_\Lambda}}\sinh\left(\frac{3}{2}\sqrt{\tilde{\Omega}_\Lambda}(1-\beta)(1-\nu)H_0 t\right)\right]^{\frac{2\beta}{(1-\nu)}}. \tag{76}$$

Using (20) into (2) we can obtain the particle number density which is given by

$$n(t) = n(t_0)\left[\sqrt{\frac{\tilde{\Omega}_m}{\tilde{\Omega}_\Lambda}}\sinh\left(\frac{3}{2}\sqrt{\tilde{\Omega}_\Lambda}(1-\beta)(1-\nu)H_0 t\right)\right]^{-\frac{2}{(1-\nu)}}. \tag{77}$$

B. Generalized Second Law of Thermodynamics

This section is devoted to examine and analyse the validity of generalized second law (GSL) of thermodynamics for the presented model. According to the GSL of thermodynamics, the total entropy of the Universe, i.e., sum of the horizon entropy and entropy of the fluid always increases with time [117]. Mathematically, we can express it as

$$\frac{d}{dt}(S_H + S_m + S_\Lambda) \geq 0, \tag{78}$$

where S_H is the entropy associated with the Hubble horizon (apparent horizon) for flat Universe, S_m is the entropy of dark matter produced by gravitationally induced matter creation and S_Λ is the entropy associated with vacuum energy. The GSL thermodynamics was first formulated for black hole by Bekestein and Hawking [117]. They stated that the black hole horizons have entropy and temperature associated with them. According to Bekenstein, the horizon entropy is proportional to the area of the horizon, which is given by

$$S_H = \frac{k_B A}{4 l_{\rm Pl}^2}, \tag{79}$$

where k_B is the Boltzmann constant, $A = 4\pi {R_H}^2$ is the area of the Hubble horizon and $l_{\rm Pl} = \sqrt{\frac{\hbar G}{c^3}}$ is the Planck length. Here, $R_H = \frac{1}{\sqrt{(H^2+\kappa a^{-2})}}$ is the radius of the Hubble horizon. In a spatially flat FLRW Universe, we have $R_H = \frac{1}{H}$. Assuming $k_B = \hbar = c = 8\pi G = 1$, the above equation can be rewritten as

$$S_H = \frac{8\pi^2}{H^2} \tag{80}$$

Therefore, the rate of change of entropy of the Hubble horizon is given by

$$\dot{S}_H = -\frac{16\pi^2}{H^3}\dot{H} \tag{81}$$

The entropy of cosmic components (matter and dark energy) inside the Hubble horizon can be calculated by the Gibb's equation [118]

$$T\, dS = d(\rho V) + p\, dV, \tag{82}$$

where $V = 4\pi/3H^3$ is volume enclosed by horizon and T is temperature associated with both matter and dark energy and due to their mutual interaction, we assume that T is equal. We assume that the system bounded by the Hubble horizon remains in equilibrium so that the temperature distribution is uniform and its value equal to the temperature of the horizon. In this case, Gibbons-Hawking temperature, $T = H/2\pi$ is a natural choice for the horizon temperature. It is evident from (82) that the vacuum energy density in the present model doesn't contribute to the entropy as $\rho_\Lambda = -p_\Lambda$. The non-zero contribution to the entropy is only due to dark matter produced by gravitationally induced matter creation, which is given by

$$S = \frac{(\rho_m + \tilde{p}_m)V}{T}, \tag{83}$$

where $\tilde{p}_m = p_m + p_c$. Substituting $V = 4\pi/3H^3$ and $T = H/2\pi$ in above equation, we get matter entropy as

$$S_m = \frac{8\pi^2}{3H^4}(\rho_m + p_c). \tag{84}$$

Using $p_c = -\beta\rho_m$, Equation (84) can be written as

$$S_m = \frac{8\pi^2(1-\beta)\rho_m}{3H^4}, \tag{85}$$

where matter energy density ρ_m is obtained from the Friedmann equation (8), which is given by

$$\rho_m = 3H_0^2(1-\nu)\tilde{\Omega}_m a^{-3(1-\beta)(1-\nu)}. \tag{86}$$

Using (80) and (85) the total entropy of the Universe, i.e., $S = S_H + S_m$ is calculated as

$$S = \frac{8\pi^2}{H^2}\left[1 + \frac{(1-\beta)\rho_m}{3H^2}\right]. \tag{87}$$

On substituting the value of ρ_m from (86), we see that the total entropy can be simplified in terms of deceleration parameter q, given in (22), as

$$S = \frac{8\pi^2}{3H^2}(5 + 2q). \tag{88}$$

Now, to check the validity of GSL, given by (78) for this model, we have the combined change in entropy for the dark fluid

$$\dot{S}_m = -\frac{8\pi^2(1-\beta)\rho_m}{H^3}\left(1 + \frac{\dot{H}}{H^2}\right). \tag{89}$$

Combining (81) and (89), and noting Eq. (21), we obtain the total change in entropy as

$$\dot{S} = \frac{8\pi^2}{H}\left[2(1+q) + \frac{(1-\beta)\rho_m}{H^2}q\right], \tag{90}$$

in which we substitute the value of ρ_m and solve further to obtain the simplified expression

$$\dot{S} = \frac{16\pi^2}{H}(1+q)^2. \tag{91}$$

In equation (91) it is noted that $\dot{S} \geq 0$ as all the terms are positive and hence the GSL is satisfied. In the asymptotic limit, the change in entropy tends to zero in late time which shows that the entropy is extremized and thus the end phase is in an equilibrium state. As $q \to -1$ in the

future, the total entropy S becomes equal to the horizon entropy S_H, given by

$$S_{\rm future} = \frac{8\pi^2}{H_0^2 \tilde{\Omega_\Lambda}} \tag{92}$$

The change in entropy, $\dot{S} \to 0$ implies that the total entropy stabilizes at the value $S_{\rm future}$ and does not increase any further, i.e. it reaches an asymptotic limit [119].

The evolution of $\dot{S}$ with respect to redshift z is shown in Fig.17 using the best-fit values obtained for DS3. Figure shows the time derivative of entropy, $\dot{S}$ remains positive throughout the cosmic evolution and tends towards zero in the future. Further, the equilibrium state may be stable or unstable according as $\ddot{S} < 0$ or $\ddot{S} > 0$ at least in the far future. We plot the evolution of $\ddot{S}$ with respect to redshift which is shown in Fig.18. It is observed that $\ddot{S} > 0$ in the early phase of evolution and made a transition to $\ddot{S} < 0$ in the recent past. Thus, a thermodynamic stable equilibrium state is achieved in late time. As $z \to -1$, $\ddot{S} \to 0$ which shows that any system satisfying the extremum of entropy and convexity condition behaves like an ordinary macroscopic system. Therefore, we can say that the evolution of the Universe is like the evolution of an ordinary macroscopic system.

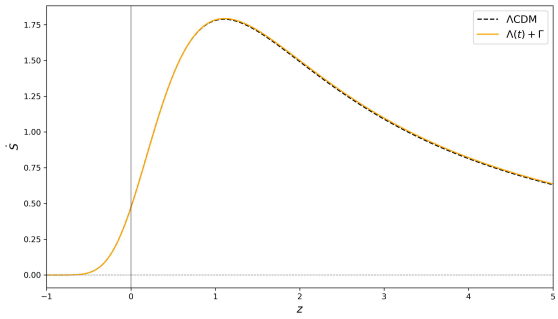


FIG. 17. Rate of change of entropy with respect to time from the best-fit values obtained for DS3. The positive value ensures that the entropy always increases. The dashed line shows the evolution for the standard model.

IX. CONCLUSION

In this work, we have studied the Universe's expansion powered by gravitationally induced 'adiabatic' matter creation and decaying vacuum in a spatially flat homogeneous and isotropic FLRW spacetime. We have demonstrated how matter creation affects the expanding Universe with a decaying vacuum, especially the late time evolution. As matter creation and varying $\Lambda(t)$ are phenomenological and have various forms in the literature, we have considered the specific forms for the vacuum energy density and matter creation rate, viz., $\rho_\Lambda = c_0 + 3\nu H^2$ and $\Gamma = 3\beta H$. Based on

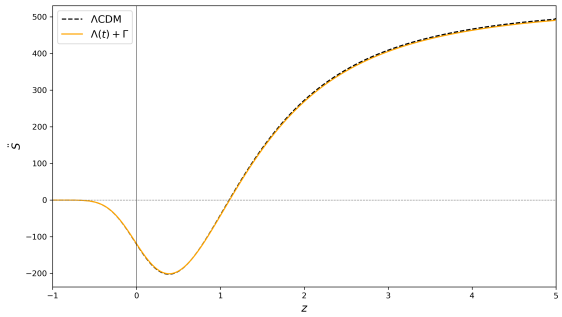


FIG. 18. Second derivative of entropy with respect to time from the best-fit values obtained from DS3 showing how the entropy accelerates with time. The dashed line shows the evolution for the standard model.

the latest observational data, the motivation was to study the dynamics of a cosmological model that has an interplay between the vacuum (dark energy) and the matter sector. The dynamical analysis confirms that the Universe's expansion is going through a cosmic acceleration phase. We have described this phenomenon using analytic solutions and observational constraints of the Hubble function and deceleration parameter. Our model's parameter constraints, which we obtained through the MCMC sampler, largely agreed with the ΛCDM model. We now present a summary of the key findings.

We have found the analytic solutions of the various cosmological parameters like $H(z), q(z), w_{\rm eff}(z)$ and $z_{\rm tr}$ and then tested them with the latest observational data. Our results show that the Universe is currently in an accelerated expansion phase, having previously been in a decelerating phase. The transition redshift $z_{\rm tr} = 0.675 \pm 0.019$ for DS1, $z_{\rm tr} = 0.675 \pm 0.019$ for DS2 and $z_{\rm tr} = 0.677 \pm 0.019$ for DS3 are consistent with ΛCDM and other recent constraints reported in the literature. The values for the effective EoS parameter for the three datasets were less than $-1/3$, which affirm the accelerating phase. Thus, the parameter constraints we obtained align with our understanding of DE in the Universe.

One thing we note is that our model favors strongly favors a value close to 0.700 ± 0.007 for the current value of the fractional density parameter Ω_Λ which is slightly higher than the values reported for the standard model and the Planck estimate 0.6889 ± 0056. The tension between the two is around 1.24σ which is not signficant. The residual tensions for the Hubble parameter constraints for our model in respect of the latest measurement of $H_0 = 73.04 \pm 1.04$ km/s/Mpc [114] for the three chosen datasets DS1, DS2, and DS3 are 4.07σ, 4.06σ, and 4.03σ respectively. We have explored cosmic perturbations and structure formation through the σ_8 and S_8 parameters, which were constrained through DS3 (Table II). The evolution of $f\sigma_8$ has been shown in

Figure 14, where theoretical curves of the weighted linear growth rate function for the ΛCDM and the $\Lambda(t) + \Gamma$ model have been plotted against the 18 $f(z)\sigma_8(z)$ data points. The curves are in agreement and consistent with the observational data. Our results align with the weak gravitational lensing observations [120] for both the standard model and the proposed model. However, there is some discrepancy between the best-fit values and the values obtained from the Planck 2018 results [121]: $\sigma_8 = 0.8102 \pm 0.0060$ and $S_8 = 0.825 \pm 0.011$.

We have also observed the evolution of the jerk parameter for our model (Figure 13), and for all three datasets, the parameter is positive and less than unity at early times. It tends towards unity in the future, mimicking the behaviour of the standard model, for which the parameter is always equal to one. The reduced χ^2 values for all three data combinations have also been calculated, and the values come out to be less than one, implying the model fits the observational data well and vice-versa. These values have been displayed in Table III. We have also used AIC and BIC selection criteria to discriminate between our chosen and the reference model, ΛCDM. According to the values of ΔAIC obtained for all three datasets, our model fits the data reasonably well, and there is averagely strong evidence in favour of it compared to the reference model. However, as per the ΔBIC values, the evidence is stacked against our model and is probably not a very good fit.

Following that, we have discussed a thermodynamic analysis for a Universe in our proposed scenario of matter creation and decaying vacuum. We found expressions for the dark matter temperature and the particle number density. We have discussed the GSL of thermodynamics for the proposed model. We have observed that the proposed model satisfies the GSL and convexity condition for the stability. We have shown these evolutions through the graphs of the first derivative of entropy $\dot{S}$ and its acceleration $\ddot{S}$ in Figs. 17 and 18, respectively.

In conclusion, the phenomenological model depicting the interaction between dark matter with matter creation and decaying vacuum dark energy, as explored in the paper, effectively describes the evolutionary features of the Universe at both background and perturbation levels. The thermodynamic features are also in accordance with established laws. Thus, exploring the interaction between dark matter empowered by matter creation and decaying vacuum stems from diverse perspectives, and holds promise for future study. This involves employing dynamical system analysis to explore deeper insights into the evolution of the Universe.

CRediT authorship contribution statement

Lokesh Chander: Formal analysis, software, writing-original draft and **C P Singh:** Conceptualization, Methodology, Writing, review & editing.

Declaration of Competing interest

The authors declare that they have no known competing financial interests that could have appeared to influence the work reported in this paper.

Data availability

There are no new data associated with this research.

ACKNOWLEDGMENTS

One of the authors, L.C. would like to thank University Grant commission (UGC), India for granting Junior Research Fellowship to carry out this work.

[1] A. G. Riess, et al., Observational evidence from supernovae for an accelerating universe and a cosmological constant, Astron. J. **116**, 1009 (1998).
[2] S. Perlmutter, et al., Measurements of Ω and Λ from 42 high-redshift supernovae, Astrophys. J. **517**, 565 (1999).
[3] D. J. Eisenstein, et al., Detection of the baryon acoustic peak in the large-scale correlation function of SDSS luminous red galaxies, Astrophys. J. **633**, 560 (2005).
[4] O. Farooq and B. Ratra, Hubble parameter measurement constraints on the cosmological deceleration-acceleration transition redshift, Astrophys. J. Lett. **766**, L7 (2013).
[5] D. Larson, et al., Seven-year wilkinson microwave anisotropy probe (WMAP*) observations: power spectra and WMAP-derived parameters, Astrophys. J. Suppl. **192**, 16 (2011).
[6] V. Sahni, Dark matter and dark energy, Lect. Notes Phys. **653**, 141 (2004). arXiv:astroph/0403324
[7] C. ORaifeartaigh, et al., One hundred years of the cosmological constant: from superfluous stunt to dark energy, Eur. Phys. J. H **43**, 73 (2018).
[8] V. Sahni and A. Starobinsky, The case for a positive cosmological Λ-term, Int. J. Mod. Phys. D **9**, 373 (2000).
[9] L. M. Krauss and M. S. Turner, The cosmological constant is back, Gen. Relativ. Gravit. **27**, 1137 (1995).
[10] S. M. Carroll, The cosmological constant, Living Rev. Relativ. **4**, 1 (2001).
[11] R. Schützhold, On the cosmological constant and the cosmic coincidence problem, Int. J. Mod. Phys. A **17**, 4359 (2002).
[12] S. Weinberg, The cosmological constant problem, Rev. Mod. Phys. **61**, 1 (1989).
[13] H. E. S. Velten, R. F. vom Marttens and W. Zimdahl, Aspects of the cosmological coincidence problem, Eur. Phys. J. C **74** 1 (2014).
[14] O. Bertolami, Time-dependent cosmological term, Il

Nuovo Cimento B **93**, 36 (1986).
[15] K. Freese, F. C. Adams, J. A. Frieman and E. Mottola, Cosmology with decaying vacuum energy, Nucl. Phys. B **287**, 797 (1987).
[16] J. A. S. Lima and M. Trodden, Decaying vacuum energy and deflationary cosmology in open and closed universes, Phys. Rev. D **53**, 4280 (1996).
[17] J. M. Overduin and F. I. Cooperstock, Evolution of the scale factor with a variable cosmological term, Phys. Rev. D **58**, 043506 (1998).
[18] I. L. Shapiro and J. Solà, Scaling behavior of the cosmological constant and the possible existence of new forces and new light degrees of freedom, Phys. Lett. B **475**, 236 (2000).
[19] I. L. Shapiro and J. Solà, The scaling evolution of the cosmological constant, J. High Energy Phys. **02**, 006 (2002).
[20] J. S. Alcaniz and J. M. F. Maia, Current and future supernova constraints on decaying Λ cosmologies, Phys. Rev. D **67**, 043502 (2003).
[21] R. Opher and A. Pelinson, Studying the decay of the vacuum energy with the observed density fluctuation spectrum, Phys. Rev. D **70**, 063529 (2004).
[22] F. Bauer, The running of the cosmological and the newton constants controlled by the cosmological event horizon, Class. Quant. Grav. **22**, 3533 (2005).
[23] S. Carneiro and J. A. S. Lima, Time -dependent cosmological term and holography, Int. J. Mod. Phys. A **20**, 2465 (2005).
[24] J. S. Alcaniz, J. A. S. Lima, Interpreting cosmological vacuum decay, Phys. Rev. D **72**, 063516 (2005).
[25] J. Solà and H. Štefančić, Effective equation of state for dark energy: Mimicking quintessence and phantom energy through a variable Λ, Phys. Lett. B **624**, 147 (2005).
[26] J. D. Barrow and T. Clifton, Cosmologies with energy exchange, Phys. Rev. D **73**, 103520 (2006) .
[27] A. E. Montenegro Jr and S. Carneiro, Exact solutions of Brans–Dicke cosmology with decaying vacuum density, Class. Quant. Grav. **24**, 313 (2006).
[28] S. Basilakos, Cosmological implications and structure formation from a time varying vacuum, Mon. Not. R. Astron. Soc. **395**, 2347 (2009).
[29] I. L. Shapiro and J. Solà, On the possible running of the cosmological constant, Phys. Lett. B **682**, 105 (2009).
[30] M. Tong and H. Noh, Observational constraints on decaying vacuum dark energy model, Eur. Phys. J. C **71**, 1 (2011).
[31] J. Solà, Cosmologies with a time dependent vacuum, J. Phys.: Conf. Ser. **283**, (2011) 012033.
[32] J. A. S. Lima, S. Basilakos and J. Solà, Expansion history with decaying vacuum: a complete cosmological scenario, Mon. Not. R. Astron. Soc. **431**, 923 (2013).
[33] M. Szydlowski, A. Stachowski and K. Urbanowski, Cosmology with a decaying vacuum energy parametrization derived from quantum mechanics, J. Phys.: Conf. Ser. **626**, 012033 (2015).
[34] J. A. S. Lima, E. L. D. Perico and G. J. M. Zilioti, Decaying vacuum inflationary cosmologies: Searching for a complete scenario including curvature effects, Int. J. Mod. Phys. D **24**, 1541006 (2015).
[35] G. J. M. Zilioti, R. C. Santos and J. A. S. Lima, From de Sitter to de Sitter: Decaying vacuum models as a possible solution to the main cosmological problems, Adv. High Energy Phys. **2018** 6980486 (2018).
[36] C. P. Singh and J. S. Peracaula, Friedmann cosmology with decaying vacuum density in Brans-Dicke theory, Eur. Phys. J. C **81**, 1 (2021).
[37] C. P. Singh and V. Khatri, Viscous fluid dynamics with decaying vacuum energy density, Phys. Rev. D **109**, 023508 (2024).
[38] J. Solà, Dark energy: a quantum fossil from the inflationary universe?, J. Phys. A: Math. Theor. **41**, 164066 (2008).
[39] E. Schrödinger, The proper vibrations of the expanding universe, Physica **6**, 899 (1939).
[40] L. Parker, Particle creation in expanding universes, Phys. Rev. Lett. **21**, 562 (1968).
[41] L. Parker, Quantized fields and particle creation in expanding universes. I, Phys. Rev. **183**, 1057 (1969).
[42] I. Prigogine and J. Geheniau, Entropy, matter, and cosmology, Proc. Natl. Acad. Sci. U.S.A. **83**, 6245 (1986).
[43] I. Prigogine, et al., Thermodynamics of cosmological-matter creation, Proc. Natl. Acad. Sci. U.S.A. **85**, 7428 (1988).
[44] I. Prigogine, Thermodynamics and cosmology, Int. J. Theor. Phys. **28**, 927 (1989).
[45] M. O. Calvão, J. A. S. Lima and I. Waga, On the thermodynamics of matter creation in cosmology, Phys. Lett. A **162**, 223 (1992).
[46] J. A. S. Lima and A. S. M. Germano, On the equivalence of bulk viscosity and matter creation, Phys. Lett. A **170**, 373 (1992).
[47] J. A. S. Lima, A. S. M. Germano and L. R.W. Abramo, FRW-type cosmologies with adiabatic matter creation, Phys. Rev. D **53**, 4287 (1996).
[48] W. Zimdahl, J. Triginer and D. Pavón, Collisional equilibrium, particle production, and the inflationary universe, Phys. Rev. D **54**, 6101 (1996).
[49] V. Johri and K. Desikan, An extended class of FRW models with creation of particles out of gravitational energy, Astrophys. Lett. Commun. **33**, 287 (1996).
[50] J. S. Alcaniz and J. A. S. Lima, Closed and open FRW cosmologies with matter creation: kinematic tests, Astron. Astrophys. **349**, 729 (1999).
[51] W. Zimdahl, D. J. Schwarz, A. B. Balakin and D. Pavón, Cosmic antifriction and accelerated expansion, Phys. Rev. D **64**, 063501 (2001).
[52] Y. Qiang, T.-J. Zhang and Z.-L. Yi, Constraint on cosmological model with matter creation using complementary astronomical observations, Astrophys. Space Sci. **311**, 407 (2007).
[53] G. Steigman, R. C. Santos and J. A. S. Lima, An accelerating cosmology without dark energy, J. Cosmol. Astropart. Phys. **09**, 033 (2009).
[54] C. P. Singh and A. Beesham, Early universe cosmology with particle creation: kinematics tests, Astrophys. Space Sci. **336**, 469 (2011).
[55] C. P. Singh, FRW models with particle creation in Brans-Dicke theory, Astrophys. Space Sci. **338**, 411 (2012).
[56] J. A. S. Lima, S. Basilakos and F. E. M. Costa, New cosmic accelerating scenario without dark energy, Phys. Rev. D **86**, 103534 (2012).
[57] J. A. S. Lima, L. L. Graef, D. Pavón and S. Basilakos, Cosmic acceleration without dark energy: Background tests and thermodynamic analysis, J. Cosmol. Astropart. Phys. **14**, 042 (2014).

[58] R. O. Ramos, M. V. dos Santos and I. Waga, Matter creation and cosmic acceleration, Phys. Rev. D **89**, 083524 (2014).

[59] S. Pan, B. K. Pal and S. Pramanik, Gravitationally influenced particle creation models and late-time cosmic acceleration, Int. J. Geom. Methods Mod. Phys. **15**, 1850042 (2018).

[60] C. P. Singh and A. Kumar, Quintessence behavior via matter creation cosmology, Eur. Phys. J. C **80**, 106 (2020).

[61] Y. Bhardwaj and C. P. Singh, Matter creation cosmology with generalized Chaplygin gas, Astrophys. Space Sci.**369**, 2 (2024).

[62] Z.-K. Guo and Y.-Z. Zhang, Cosmology with a variable Chaplygin gas, Phys. Lett. B **645**, 326 (2007).

[63] J. Zheng, et al., Revisiting chaplygin gas cosmologies with the recent observations of high-redshift quasars, Eur. Phys. J. C **82**, 582 (2022).

[64] F. Salahedin, R. Pazhouhesh and M. Malekjani, Cosmological constrains on new generalized Chaplygin gas model, Eur. Phys. J. Plus **135**, 429 (2020).

[65] L. P. Chimento, et al., Interacting quintessence solution to the coincidence problem, Phys. Rev. D **67**, 083513 (2003).

[66] L. Amendola, G. C. Campos and R. Rosenfeld, Consequences of dark matter-dark energy interaction on cosmological parameters derived from type Ia supernova data, Phys. Rev. D **75**, 083506 (2007).

[67] R. Chowdhury and P. Rudra, Interacting generalized cosmic chaplygin gas in loop quantum cosmology:A singularity free universe, Int. J. Theor. Phys. **52**, 489 (2013).

[68] H. Zhang, H. Yu, Z.-H. Zhu and Y. Gong, A quantitative criteria for the coincidence problem, Phys. Lett. B **678**, 331 (2009).

[69] M. Özer and M.O. Taha, A model of universe free of cosmological problems, Nucl. Phys. B **287**, 776 (1987).

[70] M. Gasperini, A thermal interpretation of the cosmological constant, Class. Quantum Gravity **5**, 521 (1988).

[71] J. C. Carvalho, J. A. S. Lima and I. Waga, Cosmological consequences of a time-dependent Λ term, Phys. Rev. D **46**, 2404 (1992).

[72] S. L. Adler, Einstein gravity as a symmetry breaking effect in quantum field theory, Rev. Mod. Phys. **54**, 729 (1982).

[73] L. Parker and D. J. Toms, Explicit curvature dependence of coupling constants, Phys. Rev. D **31**, 2424 (1985).

[74] S. Basilakos, M. Plionis and J. Solà, Hubble expansion and structure formation in time varying vacuum models, Phys. Rev. D **80**,083511 (2009).

[75] J. Solà, Cosmological constant and vacuum energy:old and new ideas, J. Phys.: Conf. Ser. **453**, (2013) 012015.

[76] F. Andrade-Oliveira, F. E. M. Costa and J. A. S. Lima, Decaying vacuum cosmology and its scalar field description, Class. Quantum Grav. **31**, 045004 (2014).

[77] J. Grande, J. Solà, S. Basilakos and M. Plionis, Hubble expansion and structure formation in the "running FLRW model" of the cosmic evolution, J. Cosmol. Astropart. Phys. **11** 007 (2011).

[78] D. Bessada and O. D. Miranda, Probing a cosmological model with a $\Lambda = \Lambda_0 + 3\beta H^2$ decaying vacuum, Phys. Rev. D **88**, 083530 (2013).

[79] A. P. Jayadevan, et al., Probing the dynamical system and thermal behaviors of the model, $\Lambda_0 + 3\beta H^2$, Astrophys. Space Sci. **364**, 1 (2019).

[80] M. A. van der Westhuizen and A. Abebe, Interacting dark energy: clarifying the cosmological implications and viability conditions, J. Cosmol. Astropart. Phys. **2024**, 048 (2024).

[81] R. Brandenberger, R. Kahn and W. H. Press, Cosmological perturbations in the early universe, Phys. Rev. D **28**, 1809 (1983).

[82] V. F. Mukhanov, H. A. Feldman and R. H. Brandenberger, Theory of cosmological perturbations, Phys. Rep. **215**, 203 (1992).

[83] K. A. Malik and D. Wands, Cosmological perturbations, Phys. Rep. **475** 1 (2009).

[84] J. Solà Peracaula, J. de Cruz Pérez and A. Gómez-Valent, Possible signals of vacuum dynamics in the universe, Mon. Not. R. Astron. Soc. **478**, 4357 (2018).

[85] J. de Cruz Pérez and J. S. Peracaula, Running vacuum in Brans - Dicke theory: A possible cure for the σ_8 and H_0 tensions, Phys. Dark Universe **43**, 101406 (2024).

[86] W. J. C. da Silva and R. Silva, Growth of matter perturbations in the extended viscous dark energy models, Eur. Phys. J. C **81**, 403 (2021).

[87] P. J. E. Peebles, Principles of Physical Cosmology, Princeton University Press, 1993.

[88] Y.-S. Song and W. J. Percival, Reconstructing the history of structure formation using redshift distortions, J. Cosmol. Astropart. Phys. **09**, 004 (2009).

[89] D. Huterer, et al., Growth of cosmic structure: Probing dark energy beyond expansion, Astropart. Phys. **63**, 23 (2015).

[90] S. Nesseris and L. Perivolaropoulos, Testing ΛCDM with the growth function $\delta(a)$: Current constraints, Phys. Rev. D **77**, 023504 (2008).

[91] B. Leibundgut, Type Ia supernovae, Astron. Astrophys. Rev. **10**, 179 (2000).

[92] A. G. Riess, et al., Type Ia supernova discoveries at $z > 1$ from the hubble space telescope: Evidence for past deceleration and constraints on dark energy evolution, Astrophys. J. **607**, 665 (2004).

[93] W. M. Wood-Vasey, et al., Observational constraints on the nature of dark energy: first cosmological results from the essence supernova survey, Astrophys. J. **666**, 694 (2007).

[94] A. Conley, et al., Supernova constraints and systematic uncertainties from the first three years of the supernova legacy survey, Astrophys. J. Suppl. **192**, 1 (2010).

[95] N. Suzuki, et al., The Hubble space telescope cluster supernova survey V improving the dark-energy constraints above $z > 1$ and building an early-typehosted supernova sample, Astrophys. J. **746**, 85 (2012).

[96] M. Betoule, et al., Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples, Astron. Astrophys. **568**, 1 (2014).

[97] D. H. Weinberg, et al., Observational probes of cosmic acceleration, Phys. Rep. **530**, 87 (2013).

[98] D. J. Eisenstein, Dark energy and cosmic sound, New Astron. Rev. **49**, 360 (2005).

[99] B. Bassett and R. Hlozek, Baryon acoustic oscillations, Dark Energy: Observational and Theoretical Approaches, Cambridge University Press, pp. 246 - 278 (2010), DOI: https://doi.org/10.1017/CBO9781139193627.010.

[100] A. G. Adame, et al., Desi 2024 III: Baryon acoustic oscillations from galaxies and quasars, arXiv preprint (2024), doi:10.48550/arXiv.2404.03000.

[101] A. G. Adame, et al., Desi 2024 VI: Cosmological constraints from the measurements of baryon acoustic oscillations, arXiv preprint (2024), doi:10.48550/arXiv.2404.03002.

[102] E. Gawiser and J. Silk, The cosmic microwave background radiation, Phys. Rep. **333**, 245 (2000).

[103] D. J. Fixsen, The temperature of the cosmic microwave background, Astrophys. J. **707**, 916 (2009).

[104] D. J. Fixsen, et al., The cosmic microwave background spectrum from the full COBE FIRAS data set, Astrophys. J. **473**, 576 (1996).

[105] Y. Wang and P. Mukherjee, Observational constraints on dark energy and cosmic curvature, Phys. Rev. D **76**, 103533 (2007).

[106] Y. Wang and S. Wang, Distance priors from Planck and dark energy constraints from current data, Phys. Rev. D **88**, 043522 (2013).

[107] M. Tristram *et al.*, Cosmological parameters derived from the final Planck data release (PR4), Astron. Astrophys. **682**, A37 (2024).

[108] G. Liu, Y. Wang and W. Zhao, Testing the consistency of early and late cosmological parameters with BAO and CMB data, Phys. Lett. B **854**, 138717 (2024).

[109] M. Moresco, et al., Unveiling the universe with emerging cosmological probes, Living Rev. Relativ. **25**, 6 (2022).

[110] S. Nesseris, G. Pantazis and L. Perivolaropoulos, Tension and constraints on modified gravity parametrizations of $G_{eff}(z)$ from growth rate and planck data, Phys. Rev. D **96**, 023542 (2017).

[111] D. Foreman-Mackey, et al., emcee: the MCMC hammer, Publ. Astron. Soc. Pac. **125**, 306 (2013).

[112] R. Trotta, Bayes in the sky: Bayesian inference and model selection in cosmology, Contemp. Phys. **49**, 71 (2008).

[113] N. Aghanim, et al., Planck 2018 results-VI. cosmological parameters, Astron. Astrophys. **641**, A6 (2020).

[114] A. G. Riess, et al., A comprehensive measurement of the local value of the hubble constant with 1 $km\ s^{-1}\ Mpc^{-1}$ uncertainty from the Hubble space telescope and the SH0ES team, Astrophys. J. Lett. **934**, L7 (2022).

[115] H. Akaike, A new look at the statistical identification, IEEE transactions on automatic control. **19**, 716 (1974).

[116] G. Schwarz, Estimating the dimension of a model, Ann. Stat. **6**, 461 (1978).

[117] J.D. Bekenstein, Black holes and entropy, Phys. Rev. D **7**, 2333 (1973).

[118] G. Izquierdo and D. Pavón, Phys. Lett. B **639**, 1 (2006).

[119] S. Frautschi, Entropy in an expanding universe, Science **217**, 593–599 (1982).

[120] C. Heymans, et al., KiDS-1000 Cosmology: Multi-probe weak gravitational lensing and spectroscopic galaxy clustering constraints, Astron. Astrophys. **646**, A140 (2021).

[121] N. Aghanim *et al.*, Planck 2018 results—I. Overview and the cosmological legacy of Planck, Astron. Astrophys. **641**, A1 (2020).



IETE Journal of Research

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tijr20

Design and Analysis of Substrate Integrated Folded Waveguide and Filter

Sudarshan Kumar & Asok De

To cite this article: Sudarshan Kumar & Asok De (13 Apr 2025): Design and Analysis of Substrate Integrated Folded Waveguide and Filter, IETE Journal of Research, DOI: 10.1080/03772063.2025.2486576

To link to this article: https://doi.org/10.1080/03772063.2025.2486576

Published online: 13 Apr 2025.

Submit your article to this journal

Article views: 75

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tijr20

IETE JOURNAL OF RESEARCH
https://doi.org/10.1080/03772063.2025.2486576

Check for updates

Design and Analysis of Substrate Integrated Folded Waveguide and Filter

Sudarshan Kumar[1] and Asok De[2]

[1]Maharaja Agrasen Institute of Technology, New Delhi, India; [2]Delhi Technological University, New Delhi, India

ABSTRACT
A substrate integrated folded waveguide (SIFW) and filters are proposed in this article. Analytical formulations of electric and magnetic fields, as well as cut-off frequencies, are obtained for the proposed SIFW. The dispersion characteristics of the first two fundamental modes of SIFW are considered and found to be similar to those of an ordinary substrate integrated waveguide (SIW). The SIFW is 50% narrower than the SIW. The filter is obtained by making *H*-plane slots in the central partial *H*-plane metal vane of the SIFW. The proposed SIFW filters have an application in the 26 GHz, 5 G frequency band. Both the SIFW and the SIFW filter are much smaller than their unfolded equivalents. The simulated results and the measured results match very well.

KEYWORDS
Filter; Rectangular waveguides (RWG); Substrate integrated waveguide (SIW); Substrate integrated folded waveguide (SIFW)

1. INTRODUCTION

Conventional metallic waveguides have the attractive characteristics of low insertion loss, high power handling capability, and high-quality factor, so their components have been widely employed in microwave and millimeter wave systems. However, due to their 3-dimensional design, they are challenging to fabricate and integrate with other planar components. Some researchers have proposed substrate integrated waveguide (SIW) as a solution to the difficulties in recent years. It incorporates many of the advantages of conventional rectangular waveguides (RWG) [1,2], such as simple conversion into filters and antennas and easy connection with micro-strip or coplanar circuits. Compared to micro-strip and coplanar lines, SIW's structure provides additional benefits, including low profile, high Q, low cost, high power handling capability, and light weight [3]. SIW uses metal cladding on the top and bottom of the substrate to act as the top and bottom walls, much like a RWG filled with dielectric. The metal vias on either side of the SIW serve to construct the vertical sidewalls [4,5]. However, compared with micro strip or strip line components, the width of SIW may be too large for some circuits. Therefore, a reduction in the SIW width is still required. Therefore, the concept and the geometry of a SIFW have been proposed [6–8]. Che *et al.* [1] studied SIFW and deduced the formulas for cut-off frequency, propagation constant, and gap width of the central metal vane. Trong *et al.* [9] described the variational analysis of SIFW.

Various SIW and SIFW filters are available in the literature with enhanced performance [10–16]. Greater bandwidth is achieved in these smaller-sized filters by sacrificing insertion loss. Kumar *et al.* [4] designed substrate integrated waveguide filters using the periodicity in the designed structure. Nan Song *et al.* [17] studied the tunable millimeter wave equalizer, which is mainly composed of a micro strip through line and substrate integrated folded waveguide (SIFW) resonant cavities. In [18], a SIFW filter is designed for the Wi-Fi 6E application. L. Yang *et al.* [19] designed a dual mode filter using the SIFW. The high-order modes in the SIFW cavity are employed to realize the function of filters. Fang Zhu *et al.* [20] designed a planar mode composite folded substrate integrated waveguide that enables simultaneous propagation of the $TE_{10}$and TE_{10} folded modes. Some SIFW band pass filters with good performance and compact size are reported in [21–26].

In this paper, a SIFW is designed and its electric and magnetic field components and dispersion characteristics are derived. Secondly, band pass filters are designed by making slots in the central partial *H*-plane metal vane. A FEM based simulator (Ansys HFSS) is used for the simulation of SIW, SIFW, and SIFW filters.

2. DESIGN AND ANALYSIS OF SIFW

The geometrical configuration of the SIW and the SIFW is shown in Figure 1. The SIW is designed with a dielectric substrate with a relative dielectric constant of 3.48. The height of the substrate is 0.762 mm. The width of the proposed SIW is calculated for the cut-off frequency of 10.5 GHz according to the following formula [27]. The

© 2025 IETE

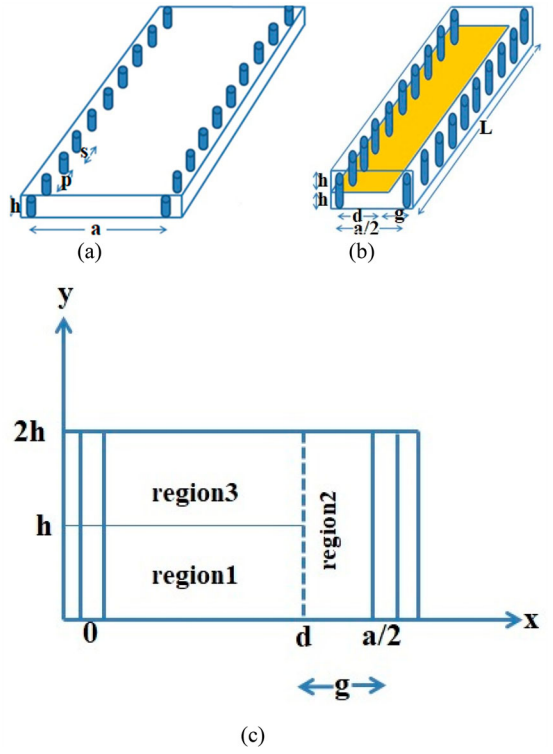


Figure 1: (a) Configuration of SIW (b) Configuration of SIFW (c) Cross section view of SIFW ($\boldsymbol{a} = 8$ mm,$\boldsymbol{s} = 0.61$ mm, $\boldsymbol{p} = 0.91$mm, $\boldsymbol{d} = 3.4$ mm,$\boldsymbol{g} = 0.6$ mm and $\boldsymbol{L} = 20$ mm)

width of the designed SIW is given by:

$$a = a_{RWG} + \frac{s^2}{0.95p} \tag{1}$$

where a is the width of the designed SIW, a_{RWG} is the width of the equivalent rectangular waveguide, s is the diameter of the vias, and p is the pitch of the vias. The calculated width of SIW is 8 mm at a cut – off frequency of 10.5 GHz. By transversely folding the SIW, the width becomes $a/2$ and the height becomes $2h$ and by inserting a partial H plane metal vane of width d, the gap spacing between the metal vane and the via is g. The thickness of the H plane metal vane is supposed to be negligible. The wave propagation direction is considered along the positive Z axis. The SIW and SIFW are both periodic devices. The unit cell of these devices comprises a dielectric substrate supported by two ground planes with two front metallic vias in SIW and an additional metal vane in SIFW. These periodic patterns have a complicated analysis. Therefore, these can be considered as regular structures to increase precision in the analysis [28,29].

The SIFW is divided into three regions labeled 1, 2, and 3 by a metal vane. As SIW only supports TE mode, the dominant mode TE_{10} and second mode TE_{20} are considered. To derive the field's components, we solve Maxwell's equation using a potential function [30].

$$E_x = -\frac{1}{\varepsilon_0\varepsilon_r}\frac{\partial F_z}{\partial y},\ E_y = \frac{1}{\varepsilon_0\varepsilon_r}\frac{\partial F_z}{\partial x},\ E_z = 0$$

$$H_x = -j\frac{1}{j\omega\mu_0\varepsilon_0\varepsilon_r}\frac{\partial^2 F_z}{\partial x\partial z},\ H_y = -j\frac{1}{j\omega\mu_0\varepsilon_0\varepsilon_r}\frac{\partial^2 F_z}{\partial y\partial z},$$
$$H_z = -j\frac{1}{j\omega\mu_0\varepsilon_0\varepsilon_r}\left(\frac{\partial^2}{\partial z^2} + k^2\right)F_z \tag{2}$$

where k is a propagation constant and $F_z(x, y, z)$ is a scalar potential function, it represents the z component of the vector potential function $\vec{F}$. The potential $\vec{F}$, and in turn F_z must satisfy Helmholtz's equation.

$$\nabla^2 F_z(x, y, z) + k^2\nabla^2 F_z(x, y, z) = 0 \tag{3a}$$

This can be reduced to

$$\frac{\partial^2 F_z}{\partial x^2} + \frac{\partial^2 F_z}{\partial y^2} + \frac{\partial^2 F_z}{\partial z^2} + k^2 F_z = 0 \tag{3b}$$

The necessary and sufficient boundary conditions for the region 1 are given by

$$E_x(0 \le x \le d, y = 0, z) = E_x(0 \le x \le d, y = h, z) = 0 \tag{4a}$$

$$E_y(x = 0, 0 \le y \le h, y = 0, z) = 0 \tag{4b}$$

and those for the region 2 are given by

$$E_x(d \le x \le a/2, y = 0, z) = E_x(d \le x \le a/2, y = b, z) = 0 \tag{5a}$$

$$E_y(x = a/2, 0 \le y \le h, z) = 0 \tag{5b}$$

The solution of the Equations (3a) or (3b) is obtained by the methods of separation of variables.

The derived electric and magnetic field components for the region 1 are as ($k_{y1} = 2m\pi/2h, m = 0, 1, 2, \ldots\ldots\ldots$),

$$E_x = \frac{k_{y1}}{\varepsilon_0\varepsilon_r}A_1\cos(k_{x1}x)\sin(k_{y1}y)e^{-jk_z z} \tag{6a}$$

$$E_y = \frac{-k_{x1}}{\varepsilon_0\varepsilon_r}A_1\sin(k_{x1}x)\cos(k_{y1}y)e^{-jk_z z} \tag{6b}$$

$$H_x = \frac{k_{x1}k_z}{\omega\mu_0\varepsilon_0\varepsilon_r}A_1\sin(k_{x1}x)\cos(k_{y1}y)e^{-jk_z z} \tag{6c}$$

$$H_y = \frac{k_{y1}k_z}{\omega\mu_0\varepsilon_0\varepsilon_r}A_1\cos(k_{x1}x)\sin(k_{y1}y)e^{-jk_z z} \tag{6d}$$

$$H_z = -j\frac{k_0^2 - k_z^2}{\omega\mu_0\varepsilon_0\varepsilon_r}A_1\cos(k_{x1}x)\cos(k_{y1}y)e^{-jk_zz} \quad (6e)$$

The derived electric and magnetic fields components for the region 2 are as ($k_{y2} = n\pi/2h, m = 0, 1, 2, \ldots\ldots\ldots$),

$$E_x = \frac{k_{y2}}{\varepsilon_0\varepsilon_r}A_2\cos\left[k_{x2}\left(\frac{a}{2} - x\right)\right]\sin(k_{y2}y)e^{-jk_zz} \quad (7a)$$

$$E_y = \frac{k_{y2}}{\varepsilon_0\varepsilon_r}A_2 sin\left[k_{x2}\left(\frac{a}{2} - x\right)\right]\cos(k_{y2}y)e^{-jk_zz} \quad (7b)$$

$$H_x = \frac{-k_{x2}k_z}{\omega\mu_0\varepsilon_0\varepsilon_r}A_2 sin\left[k_{x2}\left(\frac{a}{2} - x\right)\right]\cos(k_{y2}y)e^{-jk_zz} \quad (7c)$$

$$H_y = \frac{k_{y2}k_z}{\omega\mu_0\varepsilon_0\varepsilon_r}A_2\cos\left[k_{x2}\left(\frac{a}{2} - x\right)\right]\sin(k_{y2}y)e^{-jk_zz} \quad (7d)$$

$$H_z = -j\frac{k_0^2 - k_z^2}{\omega\mu_0\varepsilon_0\varepsilon_r}A_2\cos\left[k_{x2}\left(\frac{a}{2} - x\right)\right]\cos(k_{y2}y)e^{-jk_zz} \quad (7e)$$

The subscripts 1 and 2 define region 1 and region 2, respectively, while the wave numbers along the X and Y axis are denoted by the letters $k_x$andk_y respectively and k_z represent the wave number in the z direction as well as the propagation constant. A is an arbitrary constant of amplitude.k_0 is a free space wave number, and k is a SIFW wave number.

$$k^2 = k_{x1}^2 + \left(\frac{2m\pi}{2h}\right)^2 + k_z^2 = k_{x2}^2 + \left(\frac{n\pi}{2h}\right)^2 + k_z^2 = \omega^2\mu_0\varepsilon_0\varepsilon_r \quad (8)$$

Since the fields components E_y and H_z are continuous at $x = d$, the characteristics equation for the TE mode is derived as

$$-k_{x1}\tan(k_{x1}d) = k_{x2}\tan\left[k_{x2}\left(\frac{a}{2} - d\right)\right] \quad (9)$$

As the gap width $a/2 - d = g$, so the Equation (9) becomes

$$-k_{x1}\tan(k_{x1}d) = k_{x2}\tan(k_{x2}g) \quad (10)$$

where k_{x1} is obtained from Equations (8)–(10). The propagation constant k_z and the cut-off frequency are given by

$$k_z = \sqrt{k^2 - k_{x1}^2 - \left(\frac{2m\pi}{2h}\right)^2}, (m = 0, 1, 2, \ldots) \quad (11)$$

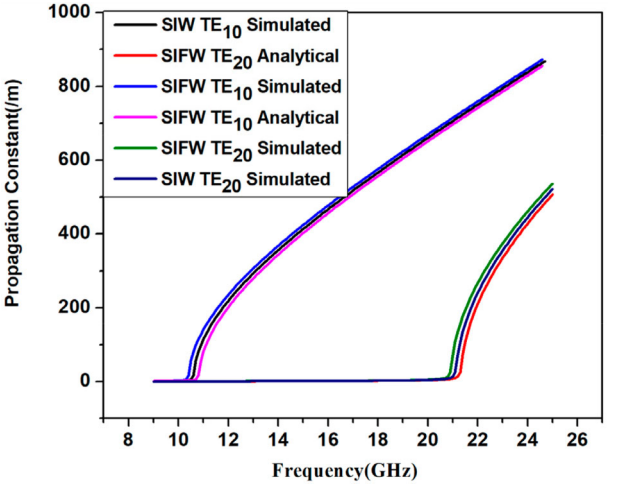


Figure 2: Dispersion characteristics of SIW and SIFW

$$f_c = \frac{c}{2\pi\sqrt{\varepsilon_r}}\sqrt{k_{x1}^2 + \left(\frac{2m\pi}{2h}\right)^2}, (m = 0, 1, 2, \ldots) \quad (12)$$

where c is the velocity of light. To validate our analytical expressions, the simulation results are compared with the analytical results of the TE_{10} and TE_{20} modes obtained from (1) to (12). The dimensions of the SIFW are $a/2 = 4$ mm,$d = 3.4$ mm and $g = 0.6$ mm. As shown in Figure 2, the simulated dispersion curves of SIW are compared with the simulated and analytical dispersion curves of SIFW.

The result shows that the cut-off frequencies and dispersion characteristics are almost the same for the modes TE_{10} and TE_{20} in both SIW and SIFW. So, at a gap width g of 0.6 mm, the propagation constants are almost equal to each other. In addition, the dispersion curves are also calculated and simulated for two different gap widths $g = 0.7$ mm and $g = 0.5$ mm. The gap width is different in these two cases, but all the other parameters are the same as in Figure 1. The results are shown in Figures 3 and 4.

Good agreement can also be observed in these two Figures 3 and 4, again representing the validity of our analysis. Table 1 shows the simulated and analytical cut-off frequencies of the TE_{10} and TE_{20} modes in SIFW. The cut-off frequencies are shown at different gap widths.

Where f_c_sim is the simulation frequency; f_c_anal is the analytical frequency and $\triangle f_c$ is the difference between the simulation and analytical frequency. The port field distribution of the TE_{10} and TE_{20} is shown in Figure 5. The field intensity is maximum in the region 2 and 3.

Table 1: SIFW analytical and simulation cut-off frequencies for various gap widths

	TE_{10}				TE_{20}			
g(mm)	f_c_sim (GHz)	f_c_anal (GHz)	Δf_c (GHz)	$\Delta f_c/f_c$_sim %	f_c_sim (GHz)	f_c_anal (GHz)	Δf_c (GHz)	$\Delta f_c/f_c$_sim %
0.5	10.45	10.476	0.026	0.248	20.90	20.942	0.042	0.200
0.6	10.55	10.578	0.028	0.265	21.10	21.147	0.047	0.227
0.7	10.66	10.691	0.031	0.290	21.32	21.372	0.052	0.243

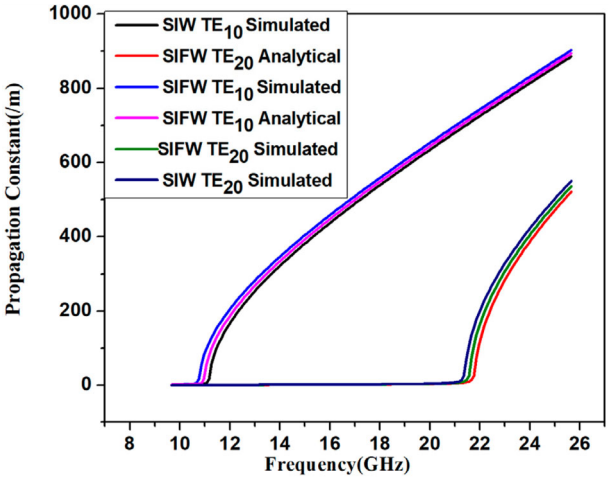


Figure 3: Dispersion characteristics of SIW and SIFW ($g = 0.7$ mm)

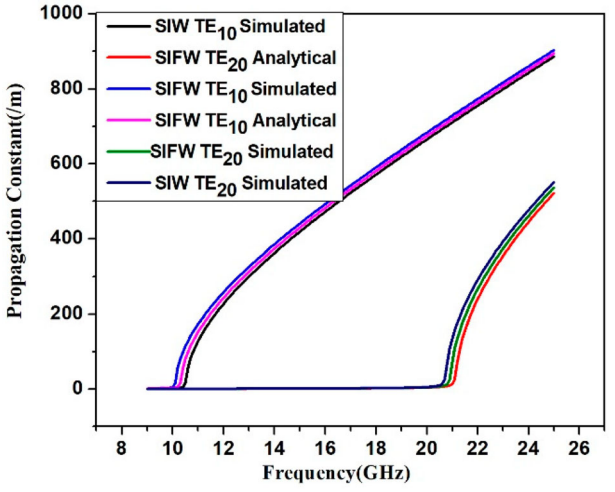


Figure 4: Dispersion characteristics of SIW and SIFW ($g = 0.5$ mm)

3. SIFW FILTER DESIGN

Figure 6 shows an isometric view of the proposed SIFW filter. Figure 7 shows the top view of a partial H-plane metal vane.

The SIFW band pass filter is a direct coupled resonator filter with a partial H plane metal vane as its basic component. The proposed SIFW filters are made up of alternating resonators and evanescent waveguide regions. The evanescent regions are constructed by inserting H plane slots into a partial H plane metal vane, and they can be represented by an admittance inverter (J inverter) circuit, as shown in Figure 8.

The jB_a and jB_b are the evanescent region reactive components and are a function of the evanescent regions length ϕ. The electrical length ϕ and normalized inverter values J are given by [29].

$$\phi = -\pi - \tan^{-1}\left(\frac{2B_b}{Y_g} + \frac{B_a}{Y_g}\right) - \tan^{-1}\frac{B_a}{Y_g} \tag{13}$$

$$J = Y_g \left| tan\left(-\frac{1}{2}\tan^{-1}\left(\frac{2B_b}{Y_g} + \frac{B_a}{Y_g}\right) + \frac{1}{2}\tan^{-1}\frac{B_a}{Y_g}\right)\right| \tag{14}$$

where Y_g, represents the wave admittance of a partial H plane SIFW filter. The normalized values of each evanescent region are obtained by (15)

$$\frac{J_{01}}{Y_g} = \sqrt{\frac{\pi\,\omega_\lambda}{2g_0g_1\omega_1}} \tag{15a}$$

$$\frac{J_{i,i+1}}{Y_g} = \frac{\pi\,\omega_\lambda}{2\omega_i}\frac{1}{\sqrt{g_i,g_{i+1}}} \tag{15b}$$

$$\frac{J_{n,n+1}}{Y_g} = \sqrt{\frac{\pi\,\omega_\lambda}{2g_n,g_{n+1}\omega_1}} \tag{15c}$$

where $g_0, g_1, g_2, \ldots\ldots g_{n+1}$ are the equal ripple low pass prototype element values, $\omega_1$is the normalized cut off frequency, $\lambda_{g0}, \lambda_{g1}, \lambda_{g2}$ is the guided wavelengths at the center frequency, lower and upper pass band edge frequencies, respectively, ω_λ is the relative band width at guided wavelength λ and n is the order of the filter. The proposed filter has specifications of 26 GHz center frequency, 0.01 dB pass band ripple, and fractional bandwidth of 4%. Simulation based SIFW design is divided into four steps. Initially, a unit cell simulation model with an evanescent region in the center and two propagating regions on both sides is investigated. In the unit cell, the dominant mode TE_{10} is considered to be propagated. For the 26 GHz center frequency, the S – parameters of the unit cell with varying widths (W_i) and depths (S_i) are produced. Second, the derived S-parameters are translated into an $ABCD$ matrix using (16). The S – parameters of the unit cell with varied widths (W_i) and depths (S_i) are calculated for the 26 GHz center frequency. Second,

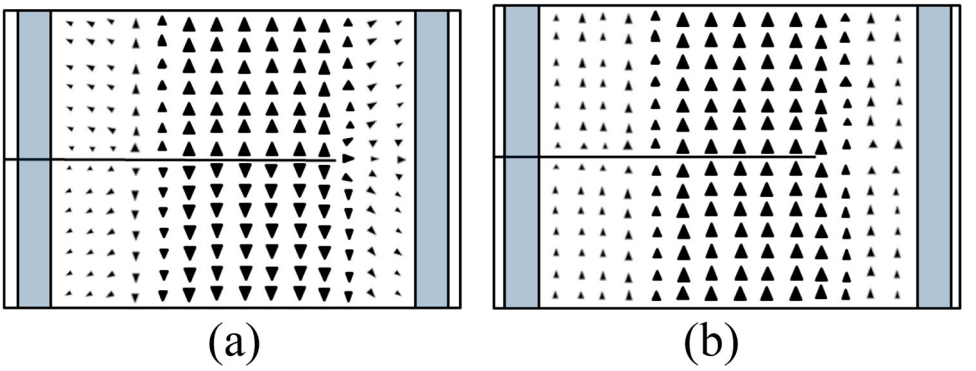


Figure 5: Electric field distribution of (a) TE_{10} and (b) TE_{20}mode

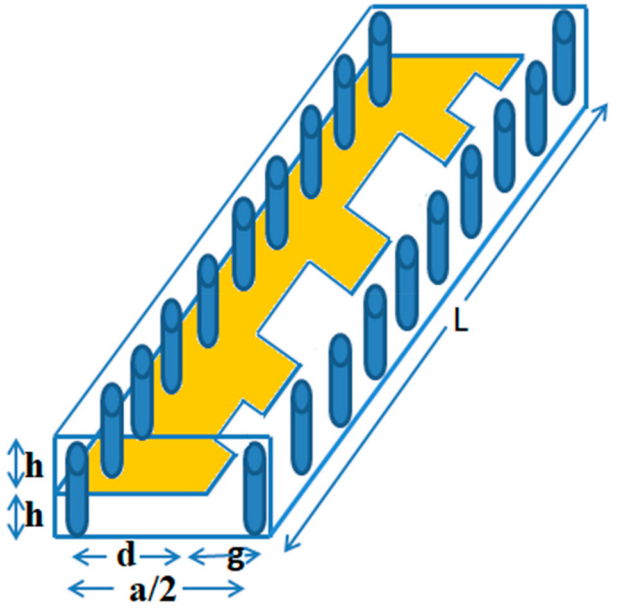


Figure 6: Isometric view proposed SIFW filter ($g = 0.6$ mm)

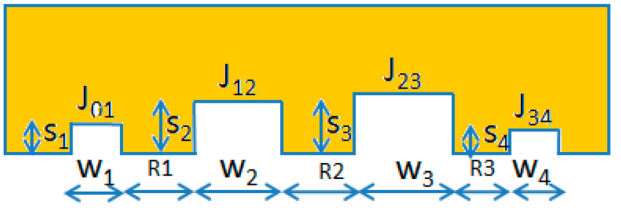


Figure 7: Top view of H plane metal vane ($S_1 = 0.4$, $S_2 = 0.8$, $S_3 = 1$, $S_4 = 0.3$, $W_1 = 2.09$, $W_2 = 3$, $W_3 = 3.74$, $W_4 = 1.8$, $R1 = 2.9$ and $R2 = 3.5$, and $R3 = 1.8$) all are in mm

the derived S parameters are used to produce an ABCD matrix (16).

$$[Overall] = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \tag{16a}$$

$$A = D = \frac{1 - S_{11}^2 + S_{21}^2}{2S_{21}} \tag{16b}$$

$$B = Z_g \frac{(1 + S_{11})^2 - S_{21}^2}{2S_{21}} \tag{16c}$$

$$C = Y_g \frac{(1 - S_{11})^2 - S_{21}^2}{2S_{21}} \tag{16d}$$

where $[Overall]$ is $ABCD$ matrix of the overall unit cell and$Z_{g=}1/Y_g$. Since the unit cell is symmetric structure so it must follows $S_{11} = S_{22}, S_{12} = S_{21}$ and $A = D$. Thirdly, the evanescent region is obtained by (17).

$$[PROP][EVAN][PROP] = [Overall] \tag{17a}$$

$$[EVAN] = [PROP]^{-1}[Overall][PROP]^{-1} \tag{17b}$$

$$[EVAN] = \begin{bmatrix} A_E & B_E \\ C_E & D_E \end{bmatrix} \tag{17c}$$

where $[PROP]$, $[EVAN]$ are the propagation and evanescent region, matrixes respectively, and $[PROP]^{-1}$ is the inverse matrix of $[PROP]$. The evanescent region values can be found using (18).

$$jB_a = \frac{A_E - 1}{B_E}, jB_b = \frac{1}{B_E} \tag{18}$$

The length of the resonators are given by (19)

$$R_n = \frac{1}{K_z}\left[\pi + \frac{1}{2}(\phi_i + \phi_{i+1})\right] \tag{19}$$

where K_z is the propagation constant and is defined in (11). The slot depths S_i are also considered based on the weight binomial and the Chebyshev's coefficient of the four slots. The dimensions of the slots in the H plane metal vane are shown in Table 2.

The frequency characteristics of the proposed SIFW filters are shown in Figure 9.

For the measurement of the SIFW filter, a tapered transition and a feeding line are designed into the prototype.

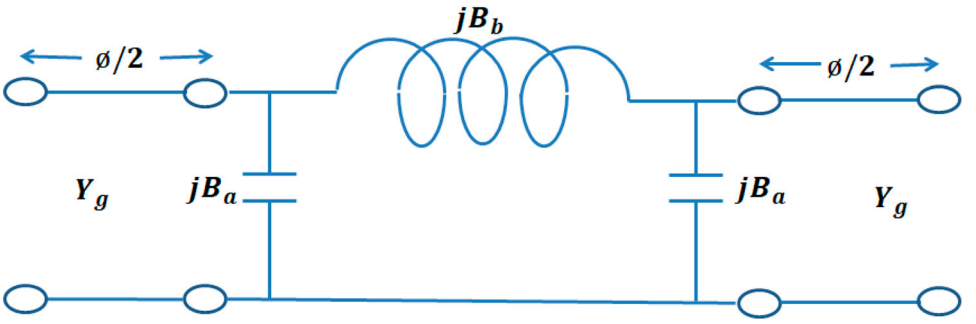


Figure 8: Admittance J inverter circuit for evanescent waveguide region

Table 2: Dimensions of the slots in the *H* Plane metal vane (all are in mm)

	S_1	S_2	S_3	S_4	W_1	W_1	W_3	W_4	R_1	R_2	R_3
Analytical	0.4	0.8	1	0.3	2.03	3	3.74	1.8	2.9	3.5	1.8
Binomial	0.4	1.2	1.2	0.4	2.03	3	3.74	1.8	2.9	3.5	1.8
Chebyshev's	0.4	0.7	0.7	0.4	2.03	3	3.74	1.8	2.9	3.5	1.8

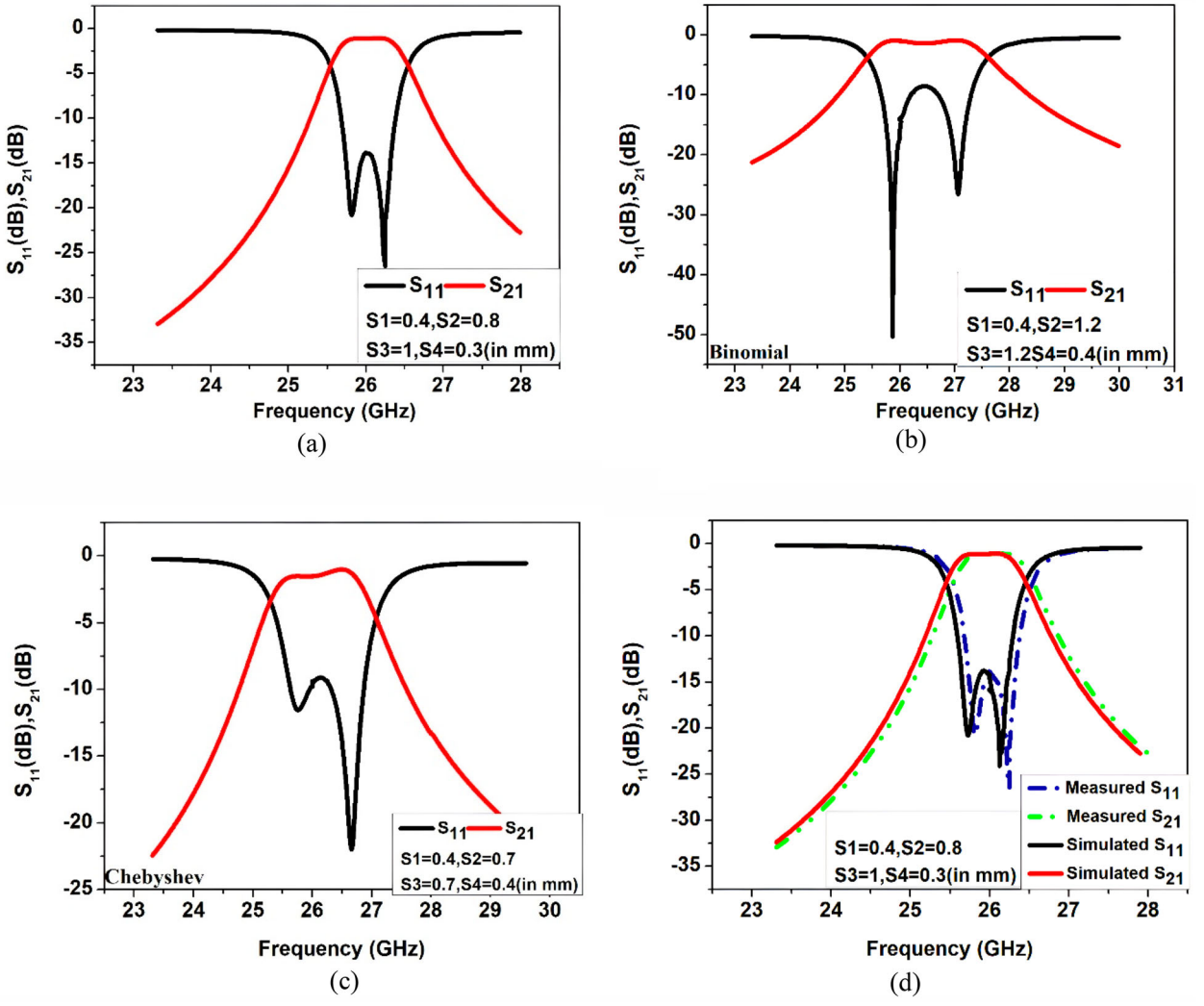


Figure 9: S parameters of (a) an analytically derived filter (b) a binomial filter (c) a Chebyshev's filter (d) an analytically derived filter simulation and measurement comparison (dimensions are in mm)

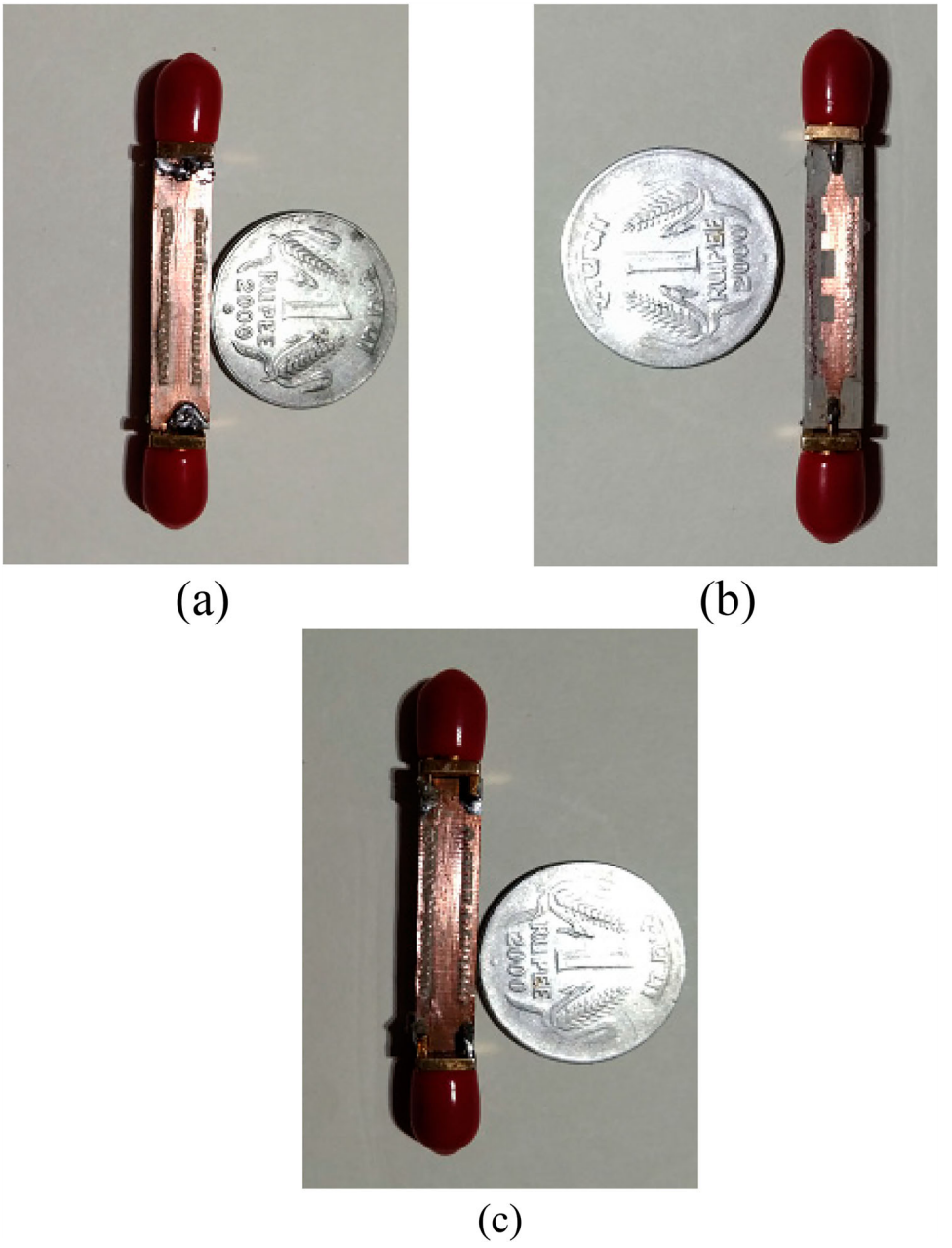


Figure 10: Designed prototype (a) Top (b) Middle (c) Bottom

A variety of transitions have been reported [31,32]. The designed prototype is shown in Figure 10.

The measurement of the designed SIFW filter is performed using an Agilent PNA-X N5247A microwave network analyzer. Calibration planes are established at the connectors where the designed prototypes are connected to the network analyzer. In our measurement setup, de-embedding was employed to isolate the parasitic effects introduced by the connectors and cabling from the actual filter characteristics. The comparison of analytical, simulated, and measured curves is depicted in Figure 9. The simulated and measured curves closely align, with any minor discrepancies attributed to fabrication errors.

The SIFW filter at 26 GHz has a 3-dB fractional bandwidth of 4% with a low insertion loss of 0.6 dB in the passband, indicating efficient signal transmission and minimal signal degradation. Insertion loss, a measure of signal attenuation through the filter, is crucial in ensuring effective signal transmission. When modifying the design to 26.5 GHz and widening the bandwidth to 7.5% using a binomial fourth-order coefficient, there's a slight increase in insertion loss to 0.8 dB, implying a bit more signal

Table 3: Performance comparison of the designed SIFW filter with various reported band pass filters

Ref. No.	Filter Type	CF(GHz)	3-dB FBW[%]	Min. RL[dB]	Min. IL[dB]	Size $\lambda_g X \lambda_g X \lambda_g$
[4]	Sinusoidal SIW	12.1,14.8	60.3,18	32,29	0.8,0.7	-
[10]	Narrow band SIW	8.25	1.1	30	3.3	1.45 × 2.7 × 0.04
[11]	DGS SIW	38.2	8.4	19	1.7	2.64 × 1.04 × 0.13
[14]	K band	22.2	1.44	30	2.9	0.9 (diameter) x 0.2
[15]	Slots with DGS	14.74	76	17.6	1.5	1.05 × 0.67 × 0.064
[16]	Inline	28	3.6	16	2.2	0.71 × 2.05 × 0.06
[21]	SIFW	11.185	4.2	21	0.82	1.85 × 0.42 × 0.1
[22]	SIFW	9.3	4.9	35	1.2	1.28 × 1.28 × 0.05
[23]	SIFW	22.2	5.8	22	3.57	1.89 × 2.14 × 0.1
[24]	SIFW	19.2	20.8	18	1.35	2.73 × 1.16 × 0.32
This Work	SIFW	26	4.11	25	0.6	1.95 × 0.355 × 0.09

loss in the passband. Alternatively, a Chebyshev fourth-order coefficient achieves a 7% bandwidth with a minimal insertion loss of 1.2 dB at a center frequency within the 26 GHz range. The slight increase in insertion loss with this configuration is balanced with achieving the desired center frequency and bandwidth characteristics, showcasing the trade-offs involved in filter design. The roll-off is the highest in the analytical filter and the lowest in the binomial. Table 3 describes the comparison of the proposed SIFW filter and some of the cited filters.

4. CONCLUSIONS

A partial H-plane SIFW and SIFW band pass filter have been designed in this article. The field components and cut-off frequencies have been derived and investigated. The accuracy achieved in the cut off frequency of SIW and SIFW is very high. SIFW band pass filters have been designed by making slots in the partial *H* plane metal vane. The filter has the advantages of compact size and small insertion loss. The filter has an application as a 5G, 26 GHz wireless broad band service in the K band.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the author(s).

REFERENCES

1. W. Che, L. Geng, K. Deng K, and Y. L. Chow, "Analysis and experiments of compact folded substrate-integrated waveguide," *IEEE Trans. Microwave Theory Tech.*, Vol. 56, no. 1, pp. 88–93, 2008.

2. D. Deslandes, and K. Wu, "Integrated microstrip and rectangular waveguide in planar form," *IEEE Microwave Wireless Components Lett*, Vol. 11, no. 2, pp. 68–70, 2001. DOI: 10.1109/7260.914305

3. C. Tomassoni, L. Silvestri, M. Bozzi, and L. Perregrini. "Substrate integrated waveguide filters and stacked cavities," IEEE MTT-S International Wireless Symposium (IWS):pp. 1–3, 2020.

4. S. Kumar, and A. De, "Design and analysis of sinusoidally modulated substrate integrated waveguide and filters," *Int. J of RF microw comput Aided Eng*, Vol. 32, no. 1, pp. 1–8, 2022. DOI: 10.1002/mmce.22912

5. A. Iqbal, J. J. Tiang, S. K. Wong, M. Alibakhshikenari, F. Falcone, and E. Limiti, "Miniaturization trends in substrate integrated waveguide (SIW) filters: a review," *IEEE. Access.*, Vol. 8, pp. 223287–305, 2020. DOI: 10.1109/ACCESS.2020.3044088

6. L. Huang, H. Cha, and S. Zhang, "Compact wideband-folded ridge substrate integrated waveguide filter," *IEEE Microwave Compon. Lett.*, Vol. 30, no. 3, pp. 241–44, 2020. DOI: 10.1109/LMWC.2020.2971659

7. C. Zhao, C. Fumeaux, and C. C. Lim, "Folded substrate-integrated waveguide band-pass post filter," *IEEE Microw. Wireless Compon. Lett.*, Vol. 27, no. 1, pp. 22–24, 2017. DOI: 10.1109/LMWC.2016.2630004

8. N. Grigoropoulos, B. Sanz-Izquierdo, and P. R. Young, "Substrate integrated folded waveguides (SIFW) and filters," *IEEE Microw. Wireless Compon. Lett*, Vol. 15, no. 12, pp. 829–31, 2005. DOI: 10.1109/LMWC.2005.860027

9. N. N. Trong, T. Kaufmann, L. Hall, and C. Fumeaux, "Variational analysis of folded substrate-integrated waveguides," *IEEE Microwave Compon. Lett.*, Vol. 25, no. 6, pp. 352–54, 2015. DOI: 10.1109/LMWC.2015.2421251

10. A. A. Khan, and M. K. Mandal, "Narrowband substrate integrated waveguide band pass filter with high selectivity," *IEEE Microwave Compon. Lett.*, Vol. 28, no. 5, pp. 416–18, 2018. DOI: 10.1109/LMWC.2018.2820605

11. J. Li, Y. Huang, H. Wang, and P. Wang, "38 GHz SIW filter based on the stepped-impedance face-to-face E-shaped DGSs for 5G application," *Microwave Optical Technology Letters*, Vol. 61, no. 6, pp. 1500–04, 2019. DOI: 10.1002/mop.31799

12. D. Tharani, R. K. Barik, Q. S. Cheng, K. Selvajyothi, and S. S. Karthikeyan, "Compact dual-band SIW filters loaded with double ring D-shaped resonators for sub-6 GHz applications," *J. Electromagn. Waves Appl.*, Vol. 35, no. 7, pp. 923–36, 2021. DOI: 10.1080/09205071.2020.1865211

13. L. F. Qiu, B. L. Xie, S. Wu, and J. F. Mao, "A flat-pass band pre distorted band pass filter with folded substrate integrated waveguide," *IEEE Trans. Circuits Syst. Express Briefs*, Vol. 69, no. 2, pp. 324–28, 2022.

14. B. Lee, T. Lee, K. Lee, M. Uham, and J. Lee, "K-band substrate integrated waveguide resonator filter with suppressed higher-Order mode," *IEEE Microwave Compon. Lett.*, Vol. 25, no. 6, pp. 367–69, 2015. DOI: 10.1109/LMWC.2015.2421313

15. T. Duraisamy, S. Kamakshy, S. S. Karthikeyan, R. K. Barik, and Q. S. Cheng, "Compact wideband SIW based bandpass filter for X, Ku and K band applications," *Radioengineering*, Vol. 3, no. 2, pp. 288–95, 2021. DOI: 10.13164/re.2021.0288

16. Z. He, C. J. You, S. Leng, and X. Li, "Compact inline substrate integrated waveguide filter with enhanced selectivity using new non-resonating node," *Electron. Lett.*, Vol. 52, pp. 1778–80, 2016. DOI: 10.1049/el.2016.2712

17. N. Song, Z. B. Wang, J. M. Zhang, and H. Chen. "Tunable milimiter wave equalizer based on folded substrate integrated waveguide" International Applied Computational Electromagnetics Society (ACES-China) Symposium, pp. 1_2, 2021.

18. S. Kumari, V. R. Gupta, and S. Srivastava, "Folded substrate integrated waveguide based multiband filter for Wi-Fi6E application," *Wireless Pers Commun*, Vol. 119, pp. 1607–18, 2021. DOI: 10.1007/s11277-021-08297-3

19. L. Yang, F. Xu, Q. Liu, J. Qiang, and J. Zhan, "Novel dual-mode filters implemented by quarter mode folded substrate integrated waveguide," *Electron. Lett*, Vol. 57, no. 17, pp. 656–58, 2019. DOI: 10.1049/ell2.12217

20. F. Zhu, M. R. Jiao, X. Zhao, K. Fan, G. Q. Luo, and K. Wu, "Mode composite folded substrate integrated waveguide," *IEEE Microwave Compon. Lett.*, Vol. 32, no. 9, pp. 1–4, 2022. DOI: 10.1109/LMWC.2022.3198198

21. N. Muchhal, and S. Srivastava, "Design of miniaturized high selectivity folded substrate integrated waveguide band pass filter with Koch fractal," *Electromagnetics.Vol*, Vol. 39, no. 8, pp. 571–81, 2021. DOI: 10.1080/02726343.2019.1675440

22. L. Yang, L. F. Xu, Q. Liu, J. Qiang, and J. Zhan, "Novel dual-mode filters implemented by quarter mode folded substrate integrated waveguide," *Electron. Lett.*, Vol. 57, no. 17, pp. 656–58, 2021. DOI: 10.1049/ell2.12217

23. X. Zhang, W. Ou, and Y. Ou, "Design and optimization of a planar folded substrate integrated waveguide bandpass filter exploiting aggressive space mapping," *Int. J. RF Microw. Comput. Aided Eng.*, Vol. 29, no. 7, pp. 1–7, 2019.

24. N. Muchhal, A. Chakraborty, M. Vishwakarma, and S. Srivastava, "Slotted folded substrate integrated waveguide band pass filter with enhanced bandwidth for Ku/K band applications," *Progress In Electromagnetics Research M*, Vol. 70, pp. 51–60, 2018. DOI: 10.2528/PIERM18041804

25. N. Claus, Y. K. Kapusuz, J. Verhaevert, and H. Rogier, "Compact and hybrid dual-band bandpass filter using folded multimode resonators and second-mode suppression," *Electronics. (Basel)*, Vol. 13, no. 10, pp. 1–23, 2024. DOI: 10.3390/electronics13101921.

26. Q. Liu, D.-W. Zhang, K. Gong, H. Qian, W.-Z. Qu, and N. An, "Single- and dual-band bandpass filters based on multiple-mode folded substrate-integrated waveguide cavities," *IEEE Trans. Microwave Theory Tech.*, Vol. 71, no. 12, pp. 5335–45, 2023. DOI: 10.1109/TMTT.2023.3276058

27. M. Bozzi, A. Georgiadis, and K. Wu, "Review of substrate integrated waveguide circuits and antennas," *IET Microwaves. Antennas & Propagation*, Vol. 5, no. 8, pp. 909–20, 2011. DOI: 10.1049/iet-map.2010.0463

28. D. Deslandes and K. Wu, "Design consideration and performance analysis of substrate integrated waveguide components" 32nd European Microwave Conference. pp. 1–4, 2002.

29. D. Deslandes, and K. Wu, "Accurate modeling, wave mechanisms, and design considerations of a substrate integrated waveguide," *IEEE Trans. Microwave Theory Tech.*, Vol. 54, no. 6, pp. 2516–26, 2006. DOI: 10.1109/TMTT.2006.875807

30. C. A. Balanis, Advanced Engineering Electromagnetic. 2nd ed. New York: Wiley, 2012, p. 274.

31. D. Deslandes, "Design equations for tapered microstrip-to-substrate integrated waveguide transitions. IEEE MTT-S International Microwave Symposium. pp. 704–7, 2010.

32. Z. Kordiboroujeni, and J. Bornemann, "New wideband transition from microstrip line to substrate integrated waveguide," *IEEE Trans. Microwave Theory Tech.*, Vol. 62, no. 12, pp. 2983–89, 2014. DOI: 10.1109/TMTT.2014.2365794

AUTHORS



Sudarshan Kumar received his BE in electronics and communication engineering from Meerut University, Mtech in digital Communication from UPTU, Lucknow, and PhD from Delhi Technological University, New Delhi. He is currently working as an assistant professor in Maharaja Agrasen Institute of Technology, New Delhi. He has published more than 20 research articles in international journals and conferences.

Corresponding author. Email: sudarshanmanav@gmail.com



Asok De obtained his BE and ME degrees with specialization in electronic and communication engineering from Jadavpur University, Kolkata, in 1978 and 1980, respectively. He earned his PhD from the Indian Institute of Technology Kharagpur in 1986. He served as a professor in the Department of Electronics and Communication Engineering. Additionally, he was the founder director at NIT Patna. His areas of expertise include antennas, transmission lines, microwave circuits, and computational electromagnetics. De has contributed significantly to his field, publishing over 150 research articles in international journals and more than 150 research articles in international/national conference proceedings. He has successfully supervised 25 PhD theses, with four currently in progress. De is recognized for his achievements and holds memberships in esteemed professional organizations, including being a senior member of IEEE (USA), a fellow of IEI (India), a fellow of IETE (India), and a life member of CSI (India).

PLOS ONE

RESEARCH ARTICLE

Development of intelligent hybrid controller for torque ripple minimization in electric drive system with adaptive flux estimator: An experimental case study

Surya Kant[1], Mini Sreejeth[2], Madhusudan Singh[2], Ambrish Devanshu[3], Majed A. Alotaibi[4,5]*, Hasmat Malik[6,7]*, Fausto Pedro García Márquez[8]

1 Department of Electronics and Communication Engineering, Graphic Era Hill University, Bhimtal, Uttrakhand, India, **2** Department of Electrical Engineering, Delhi Technological University, Delhi, India, **3** Department of Electrical Engineering, National Institute of Technology, Silchar, Assam, India, **4** Saudi Electricity Company Chair in Power System Reliability and Security, King Saud University, Riyadh, Saudi Arabia, **5** Department of Electrical Engineering, College of Engineering, King Saud University, Riyadh, Saudi Arabia, **6** Department of Electrical Power Engineering, Faculty of Electrical Engineering, University Technology, Malaysia (UTM), Johor, Malaysia, **7** Department of Electrical Engineering, Graphic Era (Deemed to be University), Dehradun, Uttrakhand, India, **8** Ingenium Research Group, Universidad Castilla-La Mancha, Ciudad Real, Spain

* majesdalotaibi@ksu.edu.sa (MAA); hasmat.malik@gmail.com (HM)

Check for updates

OPEN ACCESS

Citation: Kant S, Sreejeth M, Singh M, Devanshu A, Alotaibi MA, Malik H, et al. (2025) Development of intelligent hybrid controller for torque ripple minimization in electric drive system with adaptive flux estimator: An experimental case study. PLoS ONE 20(3): e0312946. https://doi.org/10.1371/journal.pone.0312946

Editor: Mohit Bajaj, Graphic Era Deemed to be University, INDIA

Received: November 9, 2023

Accepted: October 16, 2024

Published: March 28, 2025

Copyright: © 2025 Kant et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: There is no data related to this manuscript. Moreover, the reader may request the corresponding author to get any relevant information.

Funding: Intelligent Prognostic India and KSU University.

Abstract

In order to ensure optimal performance of permanent magnet synchronous motors (PMSMs) across many technical applications, it is imperative to minimize torque fluctuations and reduce total harmonic distortion (THD) in stator currents. Hence, this study proposes the utilization of an adaptive flux estimator (AFE) in conjunction with an Intelligent Hybrid Controller (IHC) to mitigate the ripples and total harmonic distortion (THD). The IHC system is constructed by integrating PI and fuzzy logic controllers (FLC) in a cascade configuration, alongside a new switching unit that facilitates automatic switching between the two controllers during various operations of the PMSM. AFE estimates accurate flux which is required to achieve ripple free high dynamic performance of the PMSM drive by using a limiter to fix the flux at reference flux value of the drive. The proposed controller with AFE has achieved its originality through the refinement of membership functions located at the center of the universe of discourse (UOD) and the enhancement of the switching function. These improvements have resulted in increased sensitivity in the proximity to the reference speed. The Fuzzy Logic Controller (FLC) demonstrates superior performance when operating in a transient state, whereas the Proportional-Integral (PI) controller of the proposed system exhibits satisfactory performance under steady-state situations. The efficacy of AFE with IHC is substantiated by the simulation and experimental analysis reported in this study. A significant reduction in both total harmonics distortion (THD) and torque ripples are found.

Introduction

Alternating current (AC) motor drive systems now can use advanced control techniques due to the development of the latest technologies in power semiconductors and microprocessors.

Competing interests: The authors have declared that no competing interests exist.

In many industrial applications, it is a necessary requirement that the drive system should be more effective and precise in controlling the speed and torque over a wide range of operating speed. PMSMs are frequently employed in elevators, air compressors, and new energy vehicles due to their benefits of high efficiency, high power density, and low maintenance [1–4].

High switching frequency enhances the output current waveform's quality, but it also dramatically increases power loss and decreases system efficiency. While low switching frequency minimizes power loss and advances system effectiveness it also generates torque and current

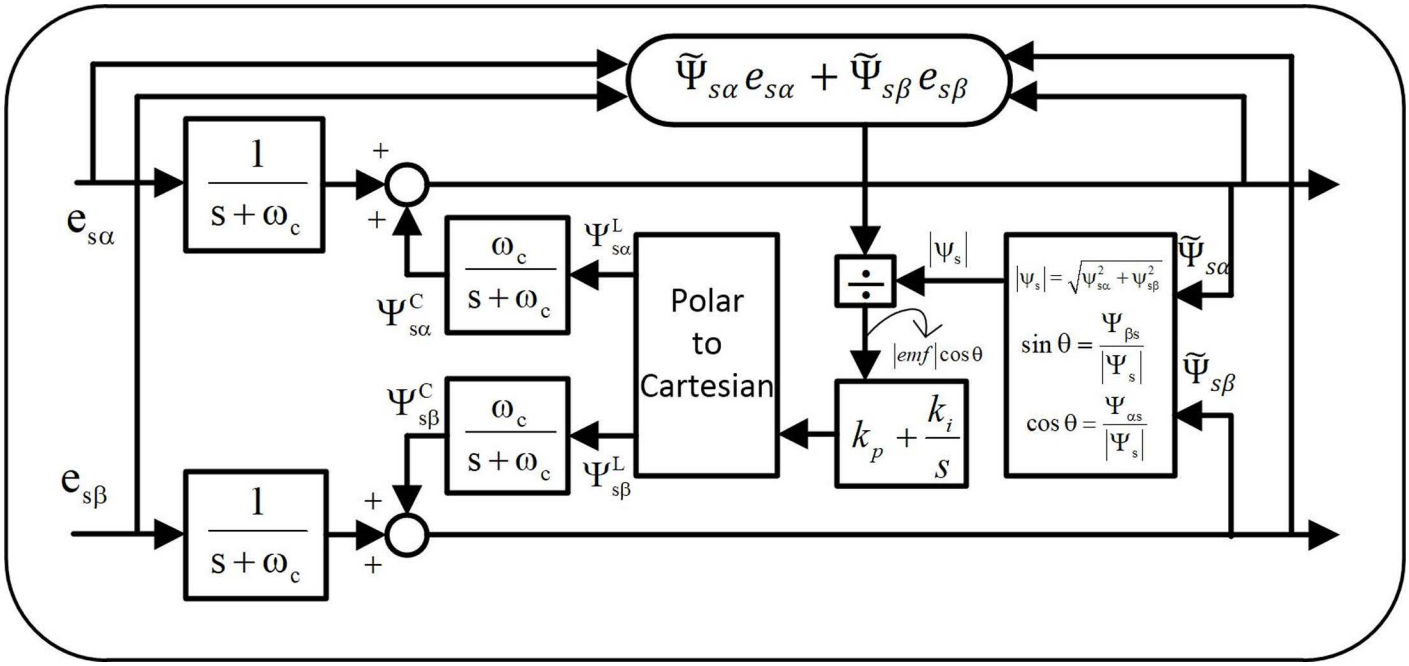


Fig 1. Adaptive flux estimator.

https://doi.org/10.1371/journal.pone.0312946.g001

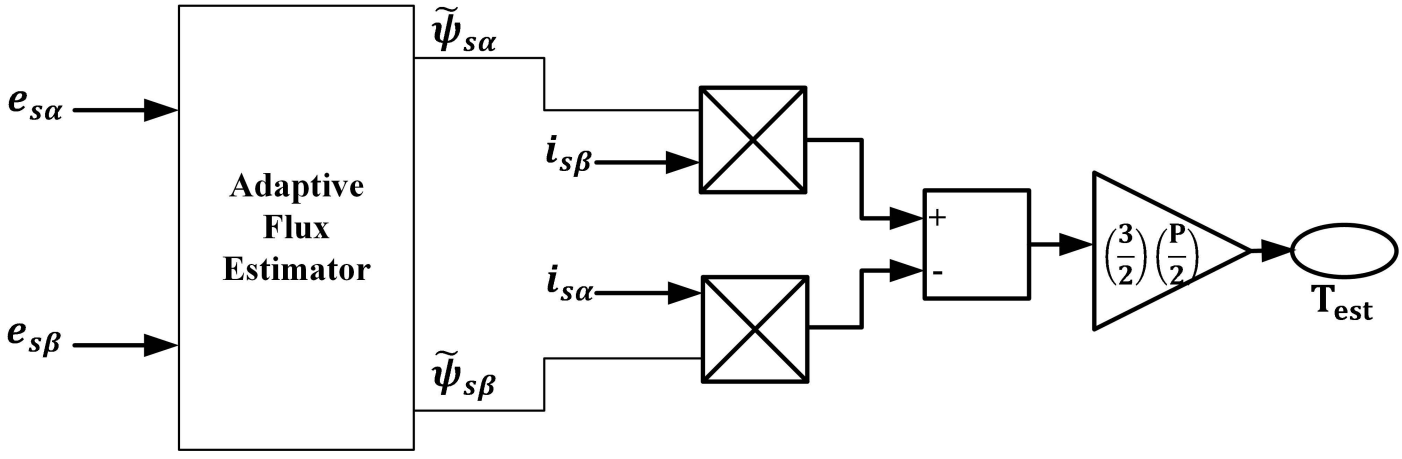


Fig 2. Adaptive torque estimator.

https://doi.org/10.1371/journal.pone.0312946.g002

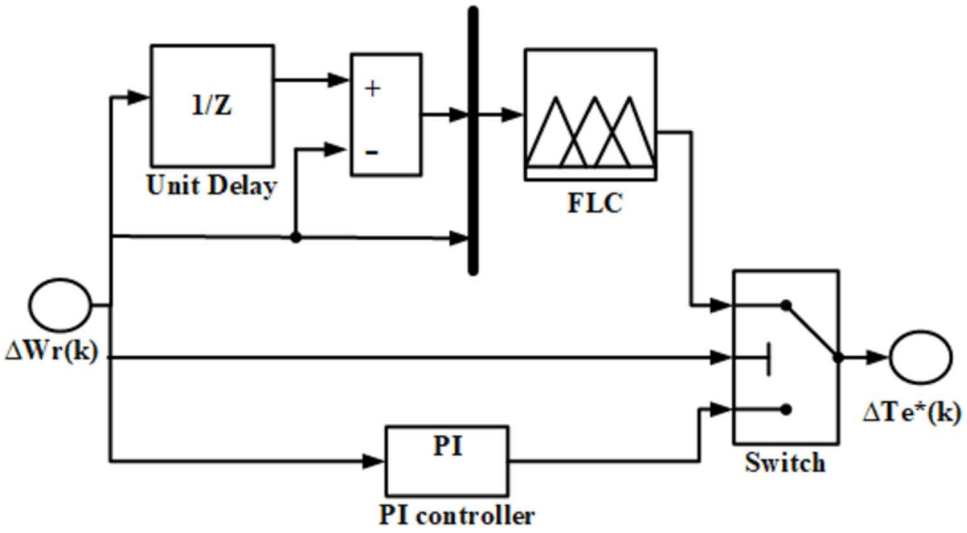


Fig 3. Intelligent hybrid controller.

https://doi.org/10.1371/journal.pone.0312946.g003

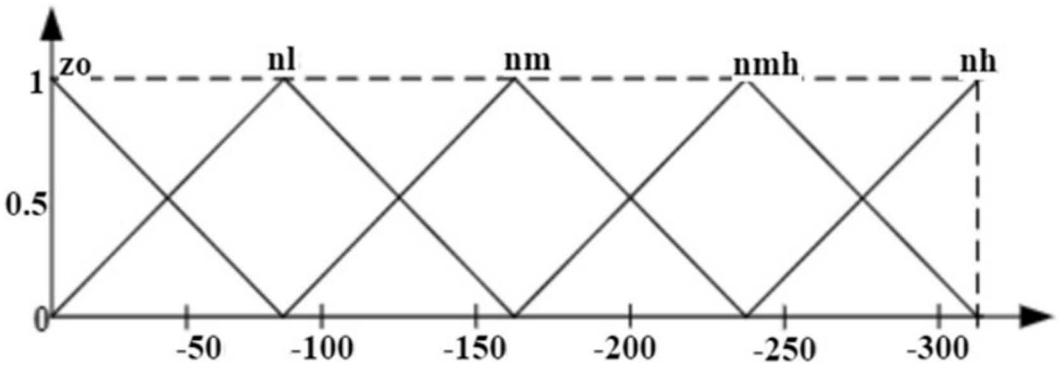


Fig 4. MFs of $\Delta e(k)$.

https://doi.org/10.1371/journal.pone.0312946.g004

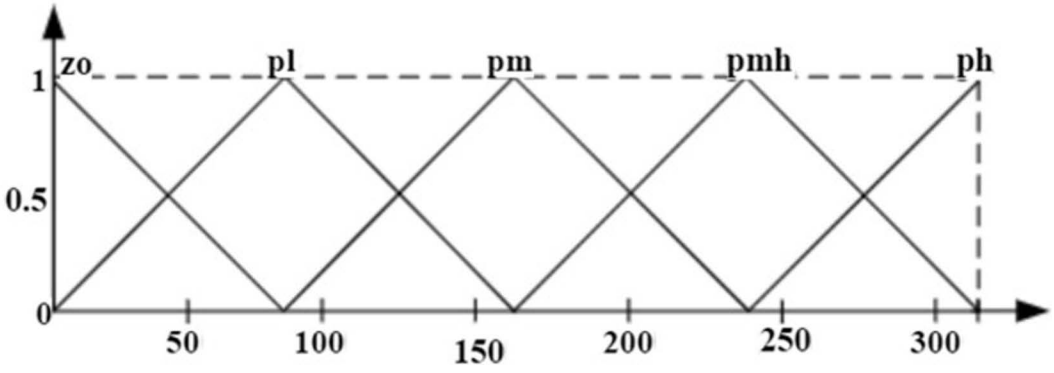


Fig 5. MFs of $\Delta\omega_r(k)$.

https://doi.org/10.1371/journal.pone.0312946.g005

ripple. Therefore, it is vital to strike a balance between control performance and effective operation.

Numerous traditional and reliable control methodologies have been documented in the literature for PMSMs. The motor performance is influenced by external disturbances and nonlinear changes in motor characteristics. Current vector control is the most often used control strategy for PMSM drive systems. Predictive control [5–7], sliding mode control [8,9], Adaptive control [10] etc. are used to remove the causes that affect nonlinearity of system. The two most popular methods for regulating the position and speed of PMSM drives are direct torque control (DTC) and field-oriented control (FOC) [11]. Although DTC-based drives produce high torque and flux ripples, they have a greater dynamic response. In FOC-based drives the current, speed, and position loops are independently controlled by controllers, and they offer better steady state performance. In the early days PI controllers were used to control current, speed and position in FOC-based PMSM drive but it degrades drive performance due

Table 1. Fuzzy rules of IHC.

$\Delta T_e^*(k)$		$\Delta e(k)$				
		zo	***nl***	***nm***	***nmh***	***nh***
$\Delta\omega_r(k)$	*zo*	*zo*	*pl*	*pmh*	*pmh*	*ph*
	pl	*pl*	*zo*	*pl*	*pmh*	*ph*
	pm	*pmh*	*pl*	*zo*	*pl*	*pmh*
	pmh	*pm*	*pmh*	*pl*	*pl*	*pl*
	ph	*ph*	*pm*	*pmh*	*pl*	*zo*

https://doi.org/10.1371/journal.pone.0312946.t001

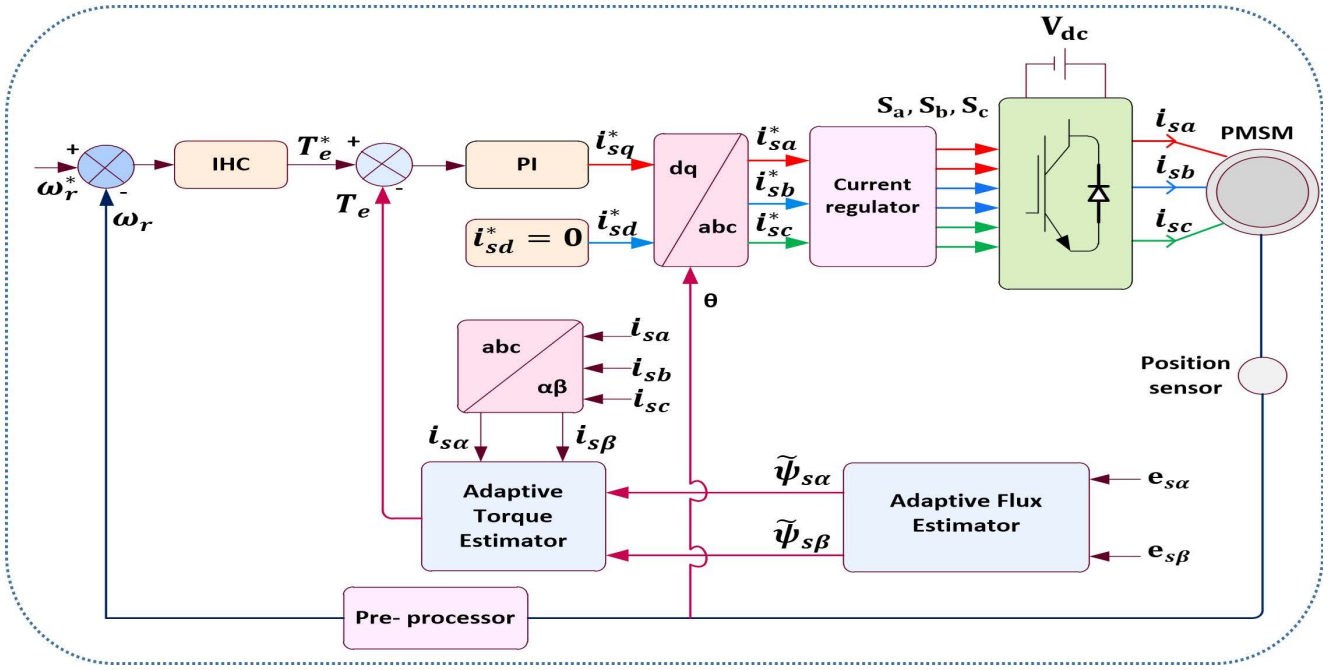


Fig 6. PMSM Drive with IHC.

https://doi.org/10.1371/journal.pone.0312946.g006

to complexity in tuning the controller parameters (Kp and Ki) in PI controller. FLC has the capability to remove this constraint of PI controllers [12]. There is no mathematical modelling required in designing FLC because it performs on the basis of user defined linguistic rules [13,14]. FLC is more complicated, computationally expensive, and memory intensive than PI controller. While PI controller offers better performance in steady state condition, FLC works greatly in transient condition of the drive [15–19]. So, an IHC can be constructed which

Table 2. Parameters of PMSM.

S.No.	Parameters	Rating
1	Power, P_{in}	3.4kW
2	Voltage, V	380V
3	Current, I	11 Nm
4	Torque, T	6.9 A
5	Speed, N	314r/sec
6	Poles, P	8
7	Resistance, R_s	1.93 Ω
8	Inductance, L_s	0.0114 H
9	Motor Inertia, J	0.11 Kg-M^2

https://doi.org/10.1371/journal.pone.0312946.t002

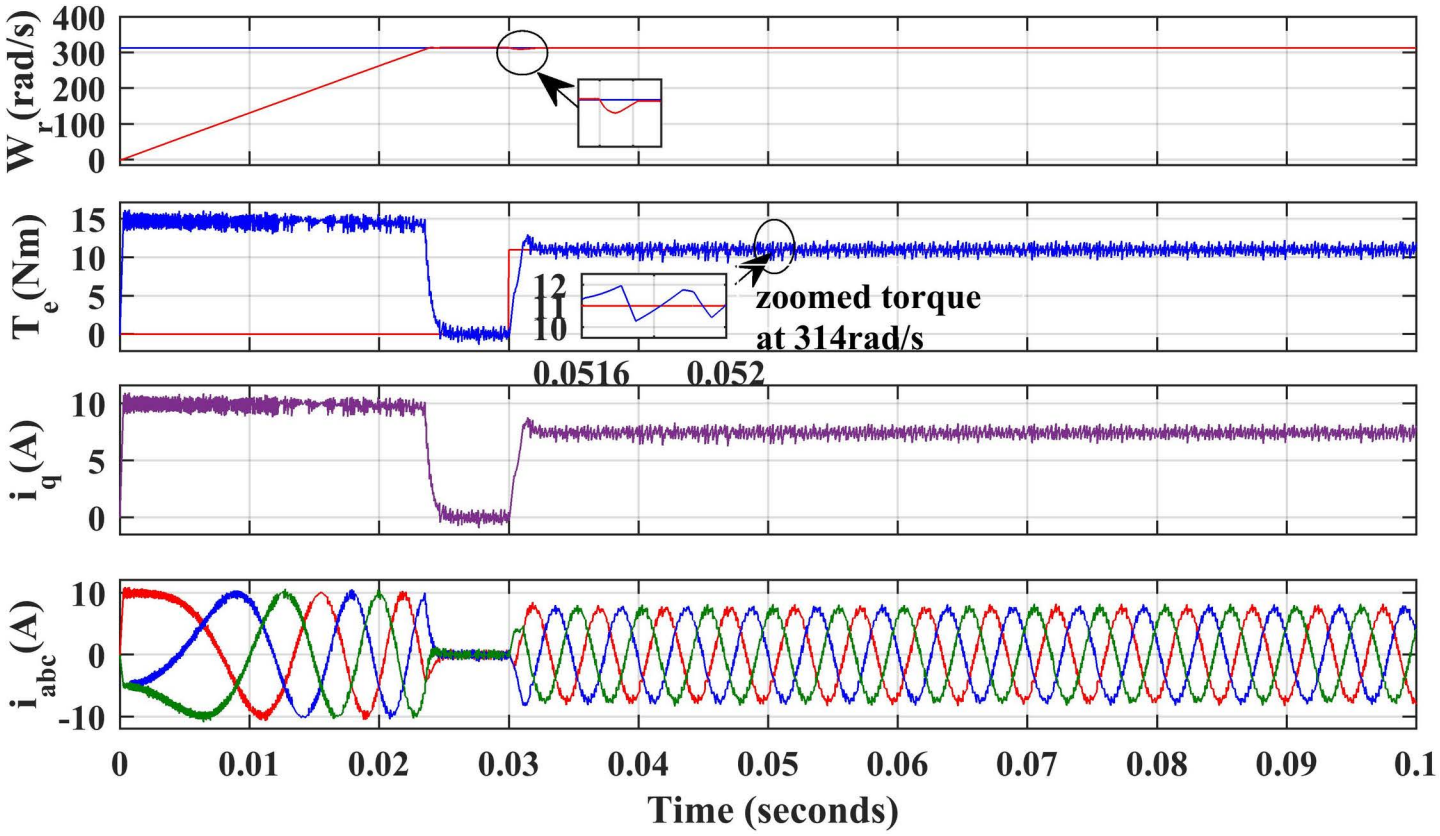


Fig 7. Simulation waveforms of W_r, T_e, i_q and i_{abc} of PMSM Drive at 314r/sec for using PI.

https://doi.org/10.1371/journal.pone.0312946.g007

adopts the qualities of PI and FLC for improved dynamic behavior of drive. Torque ripples initiated in motor by cogging torque, armature current harmonics, the magnetic field, and so on. It affects the operational life of motor. Torque ripples can be reduced by improving motor design or employing appropriate motor control technologies. Improving the distribution of winding to increase the spatial magnetic field is a typical strategy used in motor design. The use of skewed slot and fractional slot approaches in the design of stators has proven to be effective in reducing cogging torque [20]. Control algorithms optimize the excitation current of the motor, resulting in lessened torque ripples and a smooth torque output. [21–23].

Main contribution of this paper is given as follows:

- In this paper, an AFC with IHC is realized with main objective of reducing the torque ripples of the PMSM.
- The proposed method includes cascading fuzzy logic and PI controllers, as well as a switching unit that can switch FLC in the case of transient operations and the cascaded PI controller in the case of steady state operations.
- The variation in I/O and O/P variables generate the membership functions (MFs) for FLC. The FLC's design was made novel by tuning the MFs precisely at the center to offer improved sensitivity in the vicinity of rated speed while optimizing the universe of discourse. The drive's performance can also be improved by fine-tuning the MFs to account for deviations in i/o and o/p variables. Investigation and comparison of the dynamic performance of the PI controller and the PMSM drive employing the IHC are conducted.

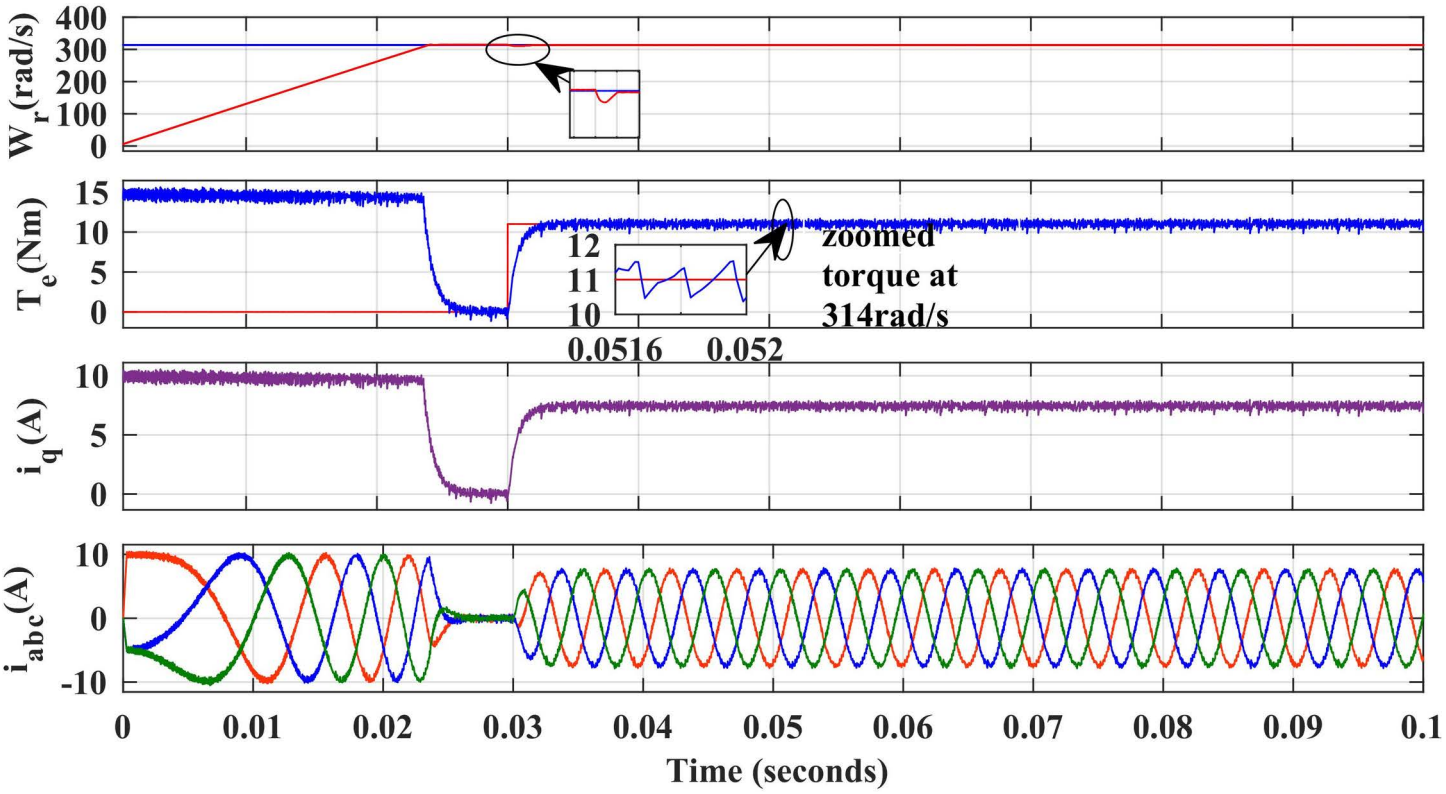


Fig 8. Simulation waveforms of W_r, T_e, i_q and i_{abc} of PMSM Drive at 314r/sec using IHC.

https://doi.org/10.1371/journal.pone.0312946.g008

This paper is arranged into five sections comprising Introduction, Principle, and modelling of PMSM through Torque Estimation, Proposed Control Scheme of PMSM Drive, simulation and experimental results followed by Conclusion.

Principle and modelling of PMSM

To provide decoupled control of stator current and rotor flux, the modelling of PMSM is derived in the α-β [24] as below.

1. Voltage and current equation of PMSM

The voltage ($v_{s\alpha}$, $v_{s\beta}$) equations in the α-β axes are represented below:

$$v_{s\alpha} = R_s i_{s\alpha} + \frac{d\psi_{s\alpha}}{dt} \tag{1}$$

$$v_{s\beta} = R_s i_{s\beta} + \frac{d\psi_{s\beta}}{dt} \tag{2}$$

Where $\mathbf{R_s}$ is the stator resistance, $\mathbf{i_{s\alpha}}$ and $\mathbf{i_{s\beta}}$ are the per phase stator currents in the α-β axes and $\psi_{s\alpha}$, $\psi_{s\beta}$ are the stator flux in the α-β axes, which is represented as

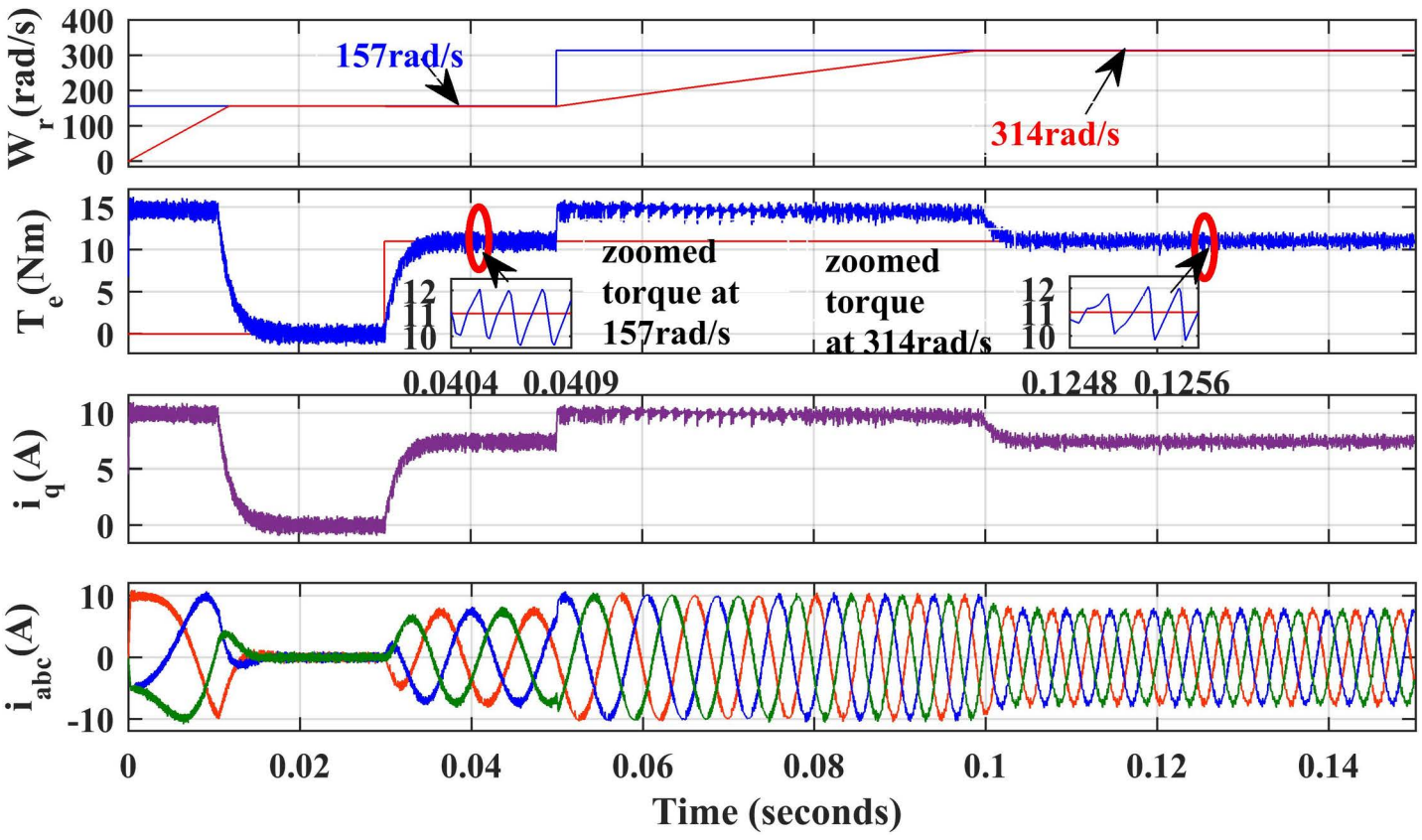


Fig 9. Simulation waveforms of W_r , T_e , i_q and i_{abc} of PMSM Drive at variable speed using PI.

https://doi.org/10.1371/journal.pone.0312946.g009

$$\psi_{s\alpha} = \int \left(v_{s\alpha} - i_{s\alpha} R_s\right) dt = \int e_{s\alpha} dt \tag{3}$$

$$\psi_{s\beta} = \int \left(v_{s\beta} - i_{s\beta} R_s\right) dt = \int e_{s\beta} dt \tag{4}$$

Where $e_{s\alpha} = \left(v_{s\alpha} - i_{s\alpha} R_s\right)$ and $e_{s\beta} = \left(v_{s\beta} - i_{s\beta} R_s\right)$

The voltage at DC link, V_{dc} and the condition of the switches of inverter may be used to depict the stator voltages in the α-β axis as:

$$v_{s\alpha} = \frac{V_{dc}}{3}\left(2S_1 - S_3 - S_5\right) \tag{5}$$

$$v_{s\beta} = \frac{V_{dc}}{3}\left[\sqrt{3}\ \left(S_3 - S_5\right)\right] \tag{6}$$

where $\mathbf{S_1}$, $\mathbf{S_3}$ and $\mathbf{S_5}$ denote the locations of the switches on a voltage source inverter (VSI), with a value of 1 for an off state and a value of 0 for a on state. The stator currents are also described as:

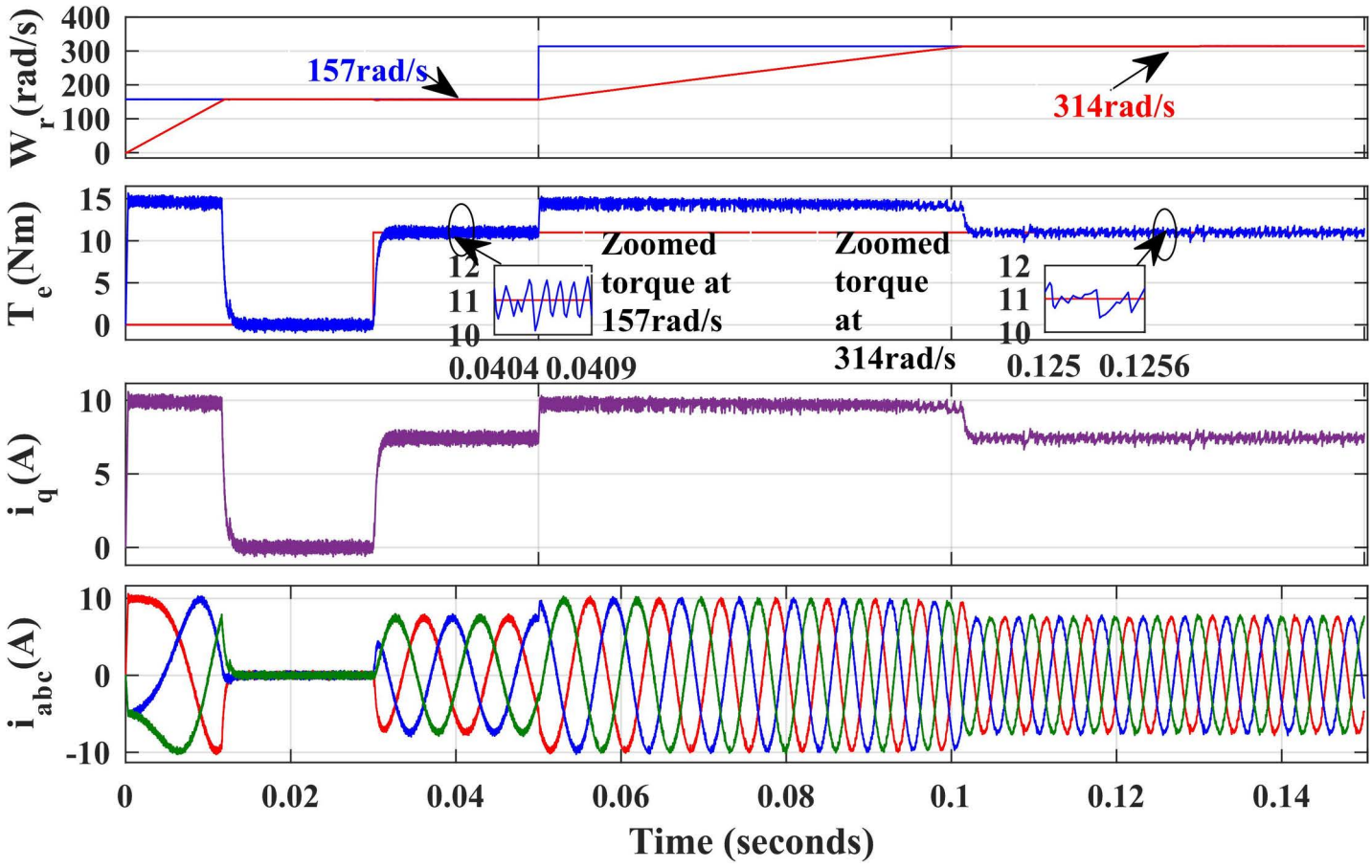


Fig 10. Simulation waveforms of W_r, T_e, i_q and i_{abc} of PMSM Drive at variable speed using IHC.

https://doi.org/10.1371/journal.pone.0312946.g010

$$i_{s\alpha} = i_a \tag{7}$$

$$i_{s\beta} = \frac{1}{\sqrt{3}}\left(i_b - i_c\right) \tag{8}$$

2. Adaptive flux estimator

Accurate flux estimation is required to achieve ripple free high dynamic performance of the PMSM drive with different control strategies such as FOC and DTC. Since the inverter is non-linear, the pure integrator used with the voltage model of the PMSM flux calculation introduces drift signal, sensors problem and many other electronics measuring instruments used while running the machine. Therefore, to overcome this problem a flux estimation technique was proposed in [25,26] for PMSM drives, but it has some limitations with it. In this method [27,28] a limiter was used to fix the flux at reference flux value of AC drive. Therefore, to achieve the variable flux levels an adaptive flux estimator is implemented in this paper as shown in Fig 1.

In this indefinite integral is substituted by a low pass filter added with the compensating feedback signal. Mathematically it can be stated as:

$$\tilde{\psi}_s = e_s \ \frac{1}{s+\omega_c} + Z_{comp} \ \frac{\omega_c}{s+\omega_c} = \tilde{\psi}_s = \psi_s^{\,1} + \psi_s^{\,2} \tag{9}$$

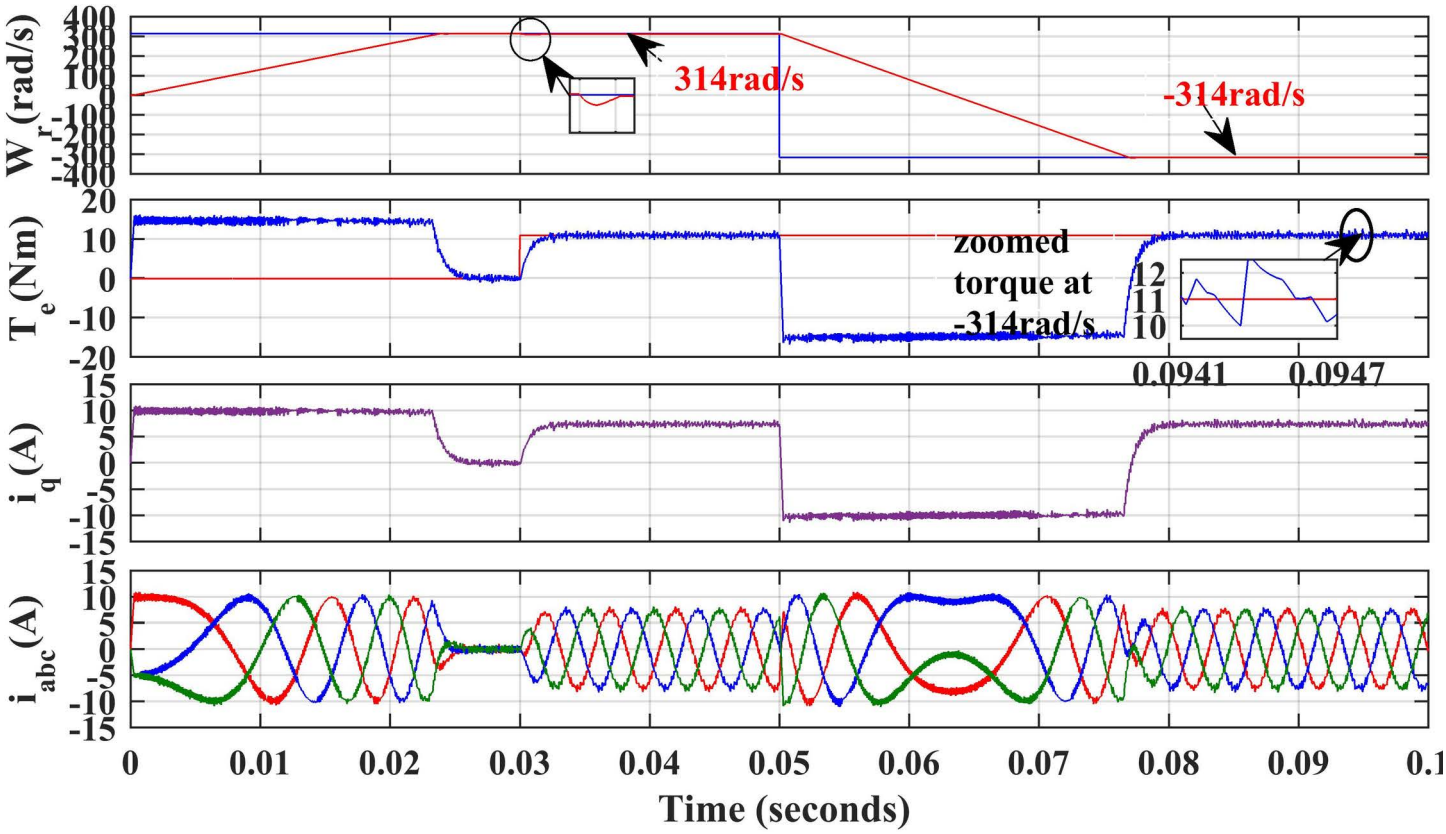


Fig 11. Simulation waveforms of W_r, T_e, i_q and i_{abc} of PMSM Drive in reverse rotation of speed using PI.

https://doi.org/10.1371/journal.pone.0312946.g011

Where, $\psi_s^1 = e_s \frac{1}{s+\omega_c}$ and $\psi_s^2 = Z_{comp} \frac{\omega_c}{s+\omega_c}$

First part of eqn. (9) is the integrator with back emf, and second part is compensating feedback flux signal to balance the output flux signal. In implemented adaptive flux estimator, motor flux and emf are in phase quadrature of each other. The two components $\Psi_{s\alpha}$ and $\Psi_{s\beta}$ of total flux Ψ_s comprised with two further components, i.e., feedback and feedforward as shown in Fig 1. Further any variations in these components lead to the phase angle deviations from 90 degrees between the flux and the emf. This deviation leads to the error which is given by the following equations:

$$|emf|.\ \cos\phi = \frac{\tilde{\psi}_{s\alpha} e_{s\alpha} + \tilde{\psi}_{s\beta} e_{s\beta}}{\left|\tilde{\psi}_s\right|} \tag{10}$$

Therefore, to eliminate this error, a detector is implemented with PI which is given by the following equation:

$$Z_{comp} = \left(k_p + \frac{k_i}{s}\right).\frac{\tilde{\psi}_{s\alpha} e_{s\alpha} + \tilde{\psi}_{s\beta} e_{s\beta}}{\left|\tilde{\psi}_s\right|} \tag{11}$$

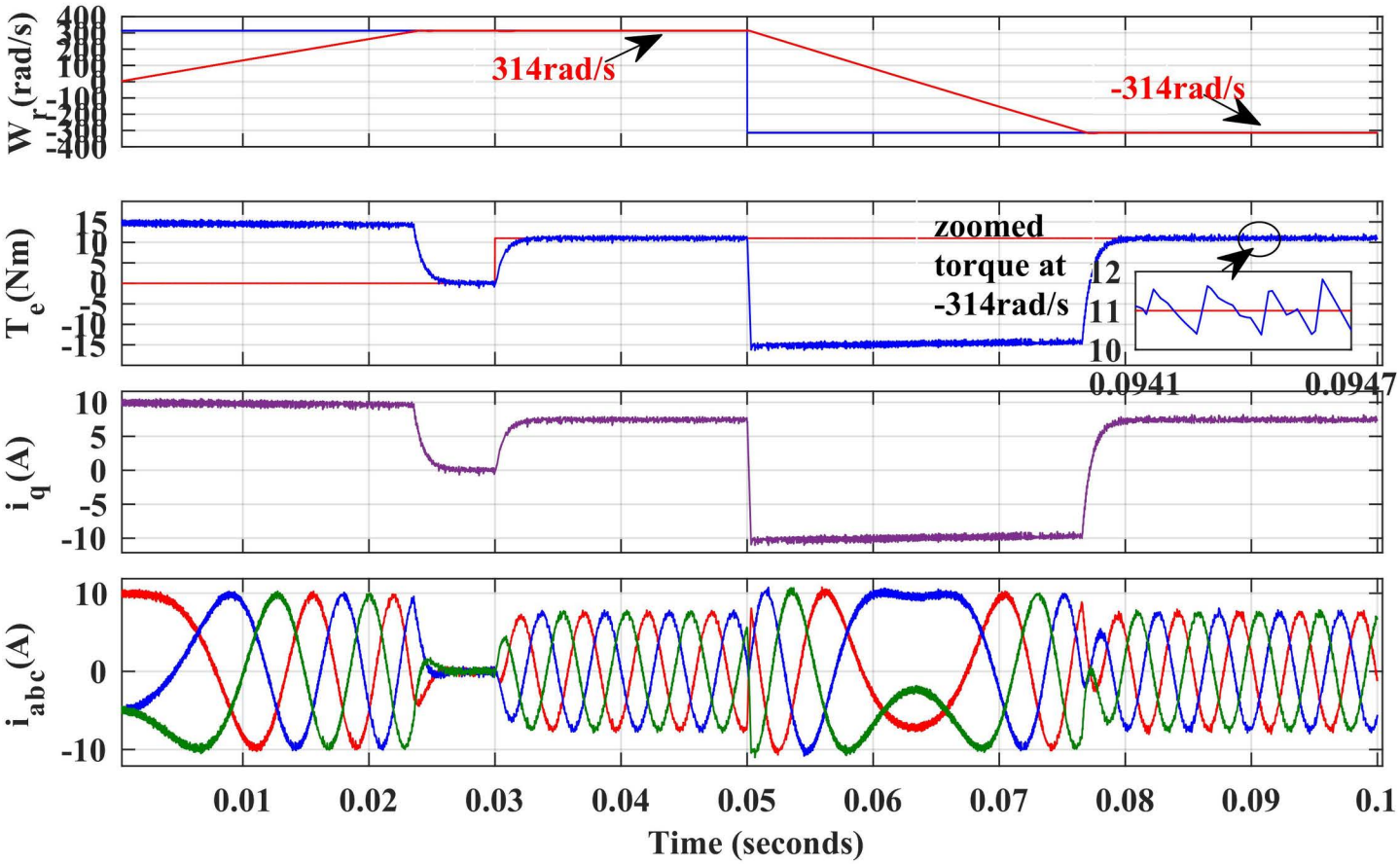


Fig 12. Simulation waveforms of W_r, T_e, i_q and i_{abc} of PMSM Drive at reverse rotation speed using IHC.

https://doi.org/10.1371/journal.pone.0312946.g012

Implementation of these methods for the operation of PMSM resulting into the lower speed, flux, and torque ripple. If the angle between the estimated flux and the and the emf is 90 degrees, then the output of the PI controller is zero.

3. Adaptive torque estimation

The estimated Torque, T_{est} is represented as given below:

$$T_{est} = \left(\frac{3}{2}\right)\left(\frac{P}{2}\right)\left(\tilde{\psi}_{s\alpha} i_{s\beta} - \tilde{\psi}_{s\beta} i_{s\alpha}\right) \tag{12}$$

'P' stands for the number of poles. Equations 1 to 12 are used to describe the torque estimator suggested here; it is seen in Fig 2.

4. PI speed and current controller

The $e_0(k)$ is output of PI controller and it is expressed as:

$$e_0(k) = K_p e(k) + K_i \int_0^k e(k) dt \tag{13}$$

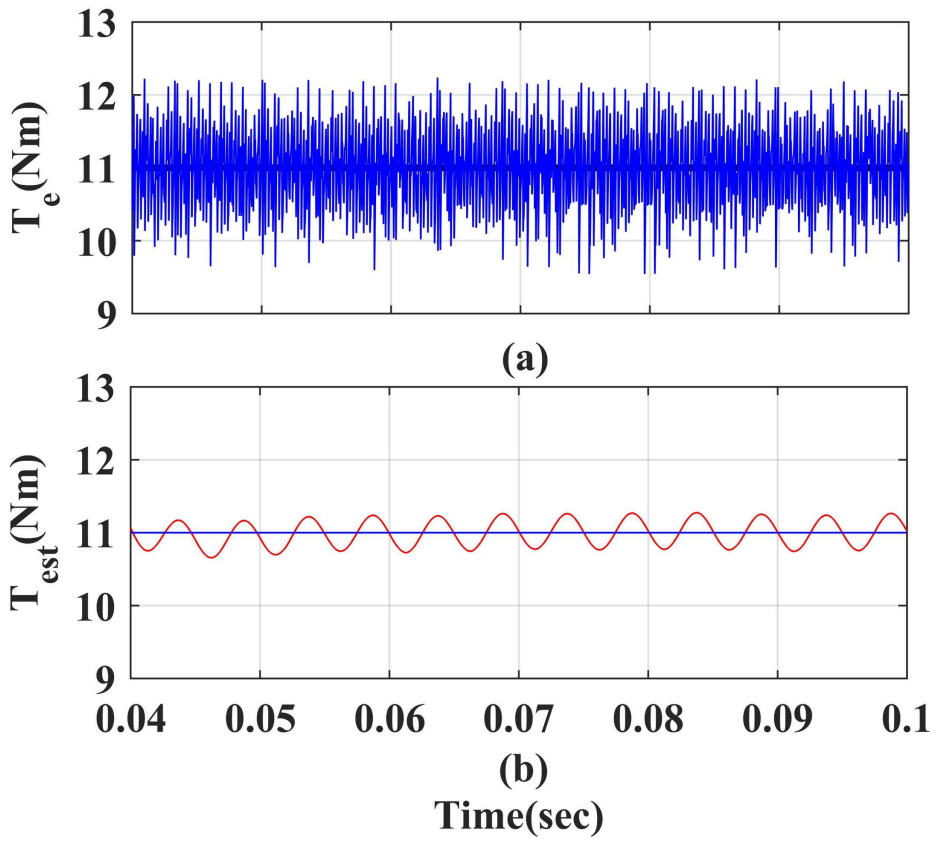


Fig 13. Measured and estimated values for the torque of the PMSM using the PI.

https://doi.org/10.1371/journal.pone.0312946.g013

Here, K_p and K_i are gains of PI, while e(k) is the error signal. By employing a PI controller, the expressions for the speed error $\omega_{r_{err}}$ that creates reference torque at the Kth instance, $T_e^*(K)$ are given.

$$\omega_{r_{err}} = \omega_r^* - \omega_r \tag{14}$$

$$T_e^*(K) = \left(K_{p\omega s} + \frac{K_{i\omega s}}{S}\right)\omega_{r_{err}}(K) \tag{15}$$

$$T_e^*(K) = \left(K_{p\omega s} + \frac{K_{i\omega s}}{S}\right)\omega_{r_{err}}(K) \tag{16}$$

Where ω_r^*, ω_r and $\omega_{r_{err}}(K)$ are the reference speed, actual speed and speed error at Kth instance respectively; $K_{p\omega s}$ and $K_{i\omega s}$ are gains of PI. The q-axis stator current, $i_q^*(K)$ is produced by the PI controller utilizing the error in torque, $T_{e_{err}}$ between the command torque, T_e^* and estimated torque, T_{est}.

$$T_{e_{err}} = T_e^* - T_{est} \tag{17}$$

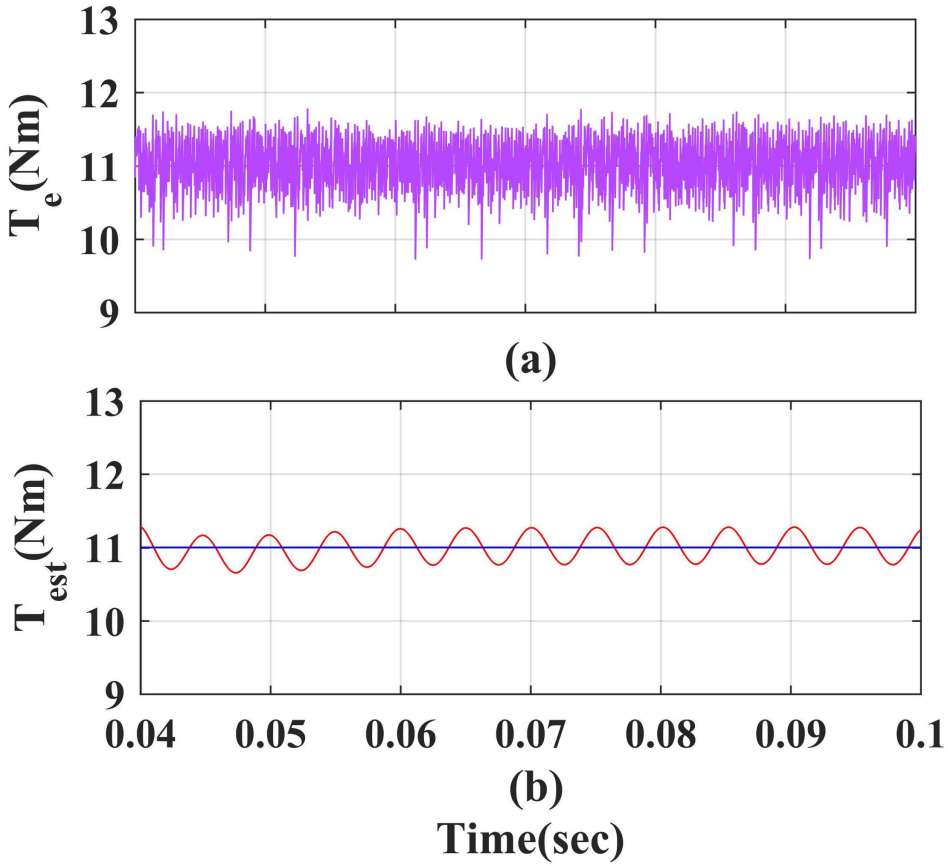


Fig 14. Measured and estimated values for the torque of the PMSM using the IHC.

https://doi.org/10.1371/journal.pone.0312946.g014

$$i_q^*(K) = \left(K_{pqs} + \frac{K_{iqs}}{S} \right) T_{e_{err}}(K) \tag{18}$$

Where K_{pqs} and K_{iqs} are defined as tuning parameter of PI controller respectively and $T_{e_{err}}(K)$ is the torque error at the Kth instance.

5. Design of intelligent hybrid controller

Soft computing approaches can be used to overcome the performance limits of traditional methods used to control the speed of PMSM drive, that rely on exact models [29,30]. The PMSM drive may experience difficulties with PI controller only to obtain the anticipated dynamic performance and when FLC is employed only then it may not be eliminating steady-state errors due to the absence of integrator. Therefore, a dual approach controller is designed which may overcome the demerits of both PI and FLC controllers. The proposed IHC compares the speed deviation with a certain fixed value. If the deviation is larger than the set value, the FLC works, and the PI controller functions as the set value is greater than the deviation. Fig 3. illustrates the controller proposed in this paper.

The IHC combines the FLC and PI controller through enhanced switching capability to utilize the PI controller during the PMSM drive's steady state and the FLC during the transient phase. It will depend on oscillations observed in the system that how will switch operate. Below is a description of the FLC algorithm created for PMSM drive speed control.

1. Determine I/O parameters.
2. Select MFs and state control rules.

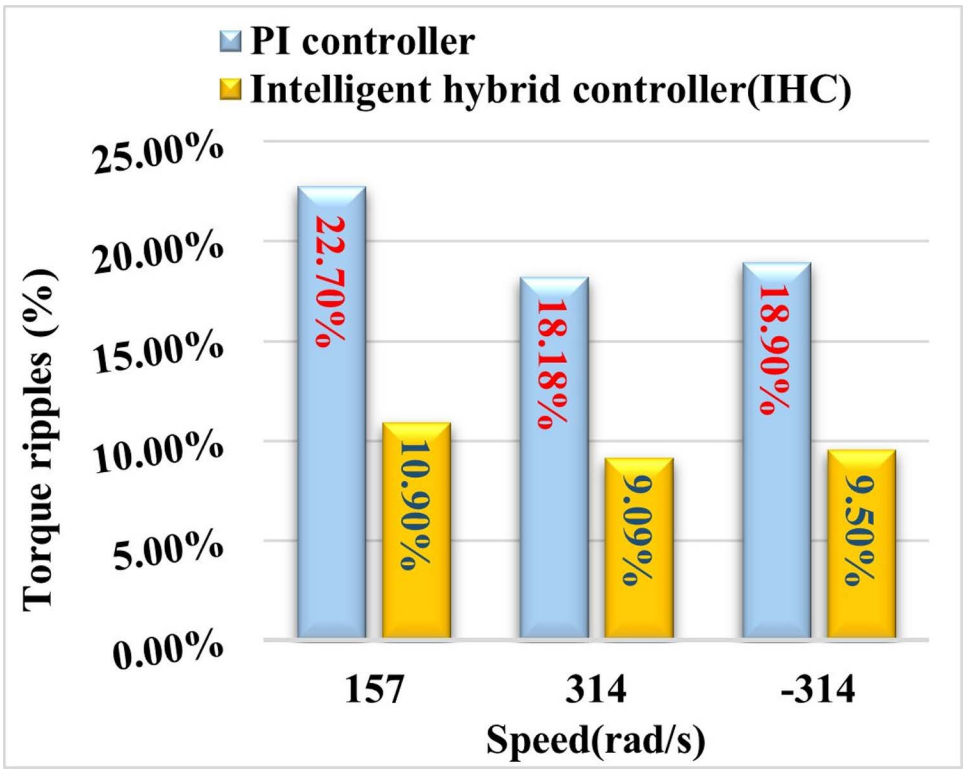


Fig 15. Torque ripples in PMSM with PI & IHC at various speeds.

https://doi.org/10.1371/journal.pone.0312946.g015

3. State probable inference with control rules and MFs.
4. Transform the fuzzy into crisp set.
5. To get the required performance, adjust the input and output gains properly.

Here input crisp variables are speed error, $\Delta\omega_r(k)$ and change in speed error, $\Delta e(k)$ which are described as:

$$\Delta\omega_r(k) = \omega_r^*(k) - \omega_r(k) \tag{19}$$

$$\Delta e(k) = \Delta\omega_r(k) - \Delta\omega_r(k-1) \tag{20}$$

Output is torque component, which is stated as

$$\Delta T_e^*(k) = \Delta T_e^*(k-1) + \mu.\Delta T_e^*(k) \tag{21}$$

where $\Delta T_e^*(k)$ is commanded value of torque at the kth sample time and μ is the gain. Fuzzyfication, rule execution, and defuzzyfication are the three main stages that make up the FLC's operation. Triangular membership functions (MFs) of I/O convert crisp variables into fuzzy variables in the first stage. The MFs for the two inputs, $\Delta e(k)$, $\Delta\omega_r(k)$ are shown in Fig 4 and Fig 5. For I/O variables, the IHC defines fifteen linguistic variables. The first input,

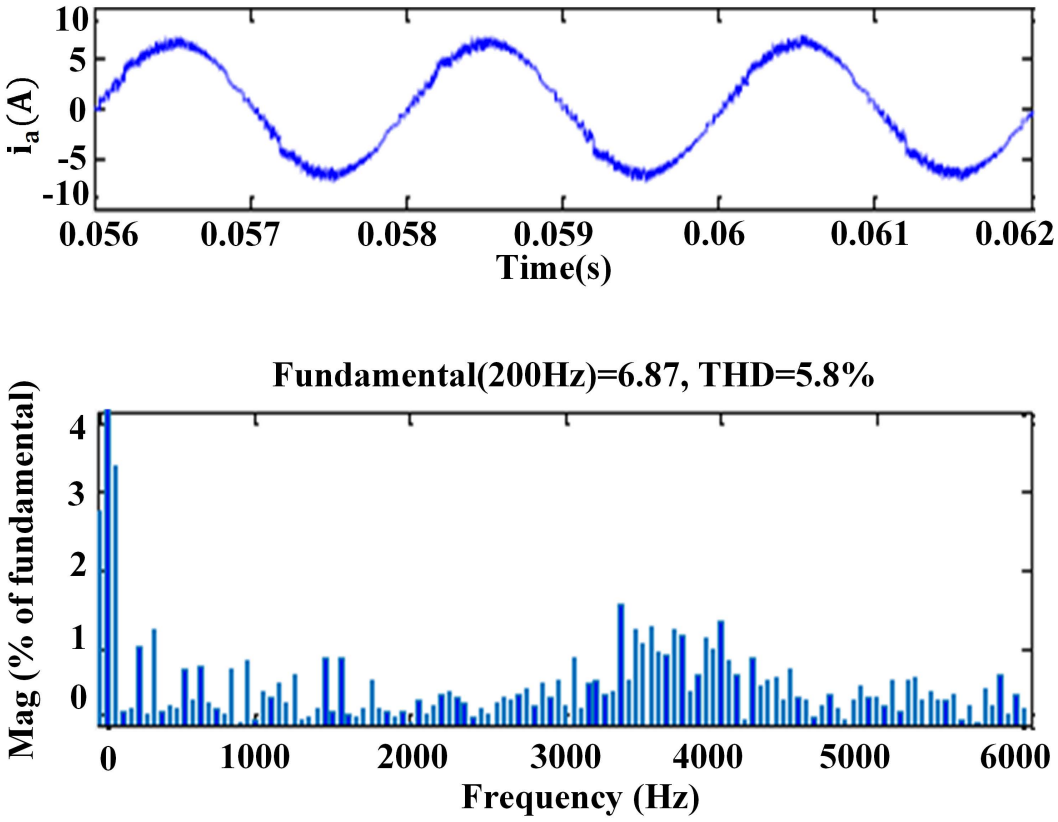


Fig 16. THD in motor current at 314r/sec with PI.

https://doi.org/10.1371/journal.pone.0312946.g016

e(k), has five linguistic variables: ***zo, nl, nm, nmh, nh.*** Likewise, the second input has ***zo, pl, pm, pmh, ph***. Thus, ***zo, pl, pm, pmh, and ph*** are the five output linguistic variables. The ***zo, nl, nm, nmh, nh, pl, pm, pmh, ph*** represents zero, negative low, negative medium, negative medium high, negative high, positive low, positive medium, positive medium high and positive high respectively.

Using fuzzy variables derived from expert knowledge, a set of control rules is constructed and put into action in the second stage. Fuzzy rules are stated as

IF $\Delta\omega_r(k)$ is 'a', $\Delta e(k)$ is 'b' THEN $\Delta T_e^*(k)$ is 'c'.

Here a, b and c are defined as fuzzy sets.

To produce computable results in fuzzy logic, fuzzy sets, and membership functions, fuzzy variables are transformed into crisp variables in the third phase, known as defuzzyfication.. Table 1 lists the fuzzy rules utilized in the IHC.

Proposed control scheme of PMSM drive

To regulate the space vectors of voltage, current, and flux, vector control uses a field-oriented technique [31,32]. High bandwidth inner loops are necessary in high performance drives in order to properly measure current, shorten the transient time, and make sure that VSI may be utilized as an amplifier [33–36]. Fig 6. demonstrates PMSM drive system with IHC.

The reference torque command is created by the IHC in the outer speed loop and contrasted with the estimated torque. The voltage at DC link, 3-phase currents (i_{abc}) and the switches positions (S1, S3 and S5) of VSI from the pulses generated by the current regulator

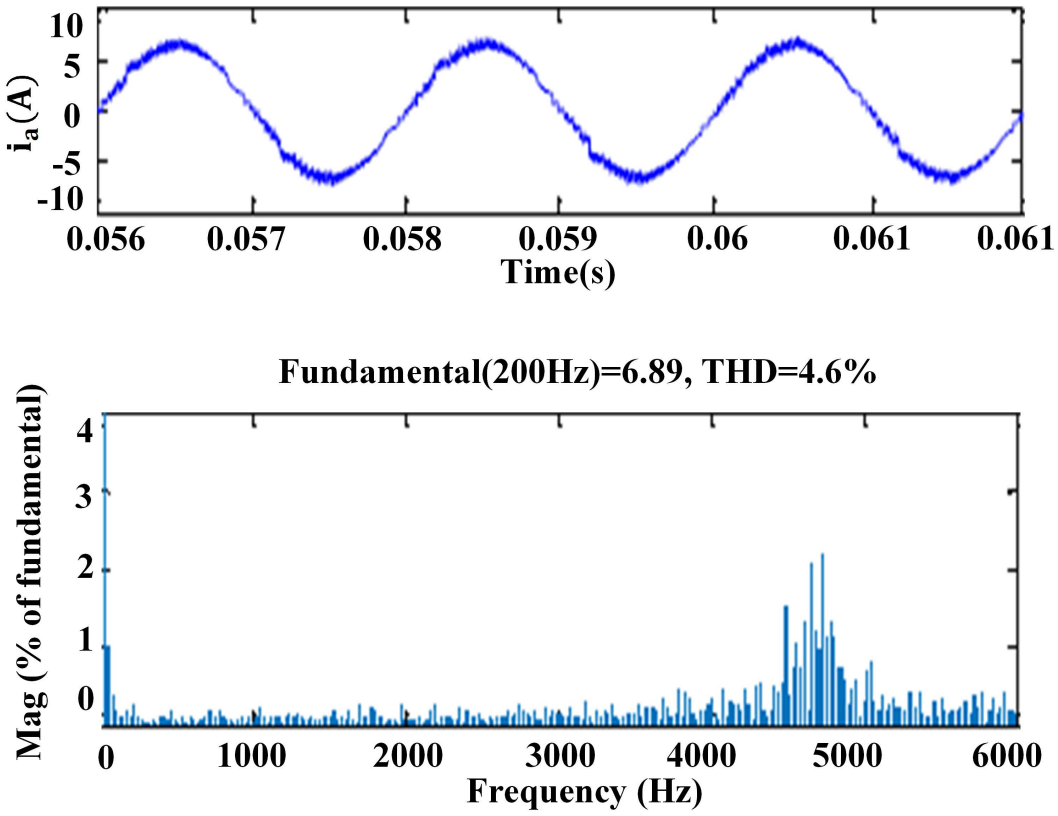


Fig 17. THD in motor current at 314r/sec with IHC.

https://doi.org/10.1371/journal.pone.0312946.g017

as stated before are used to determine the torque. The error signal obtained by comparing the estimated torque to the reference torque is given to the current controller, which produces i_q^*.

The 3-phase reference currents of stator, i_a^*, i_b^* and i_c^* are computed via Parks Transformation utilizing the rotor position from the encoder (θ), and quadrature and direct axis current, i_q^* & i_d^*.

$$i_a^* = i_q^* \cos\theta + i_d^* \sin\theta \tag{22}$$

$$i_b^* = i_q^* \cos\left(\theta - \frac{2\pi}{3}\right) + i_d^* \sin\left(\theta - \frac{2\pi}{3}\right) \tag{23}$$

$$i_c^* = i_q^* \cos\left(\theta + \frac{2\pi}{3}\right) + i_d^* \sin\left(\theta + \frac{2\pi}{3}\right) \tag{24}$$

Where $i_d^* = 0$ upto rated speed operation. Switching pulses are generated by comparing these reference currents with the triangular wave and fed to the inverter.

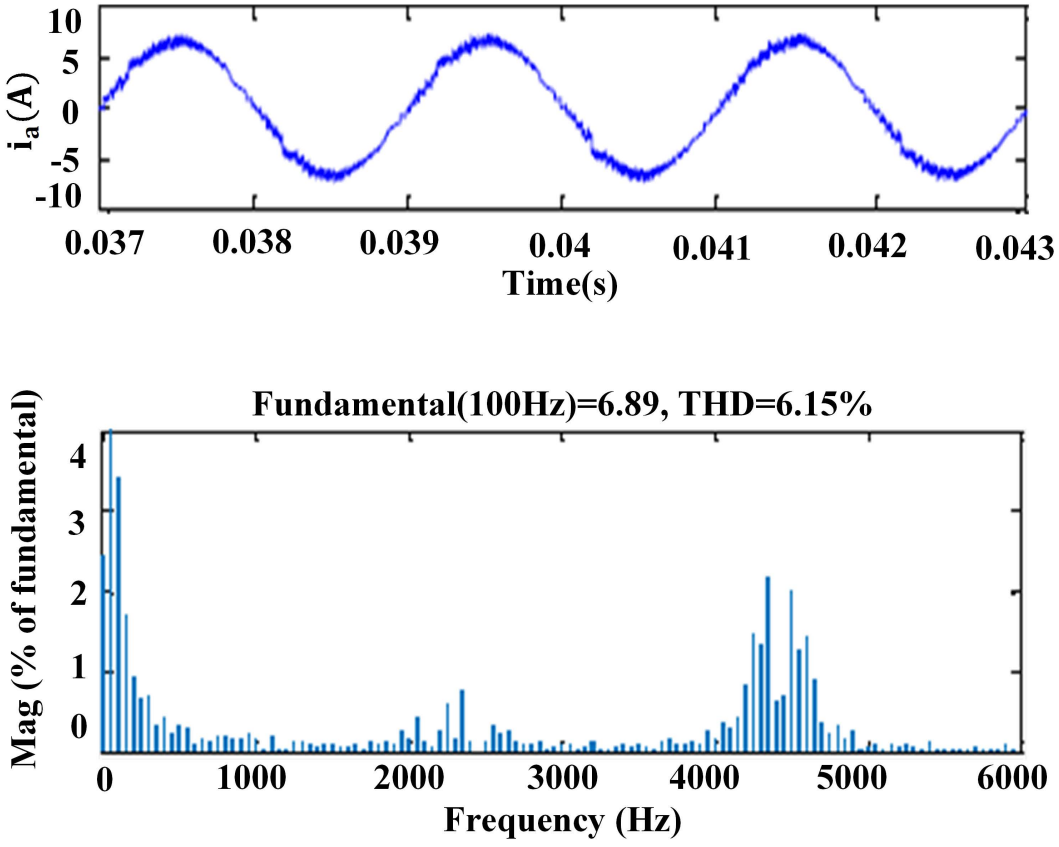


Fig 18. THD in motor current at 157r/sec with PI.

https://doi.org/10.1371/journal.pone.0312946.g018

Emulated results demonstration

In this section a simulation model of a 3.4-kW PMSM with the parameters listed in Table 2 is built in MATLAB/Simulink. The simulations are carried out to compare the proposed approach to the traditional PI control scheme and verify it at the sample time of 10μs.

1. Rated speed operation of PMSM drive

Figs 7 and 8 illustrate the simulated results of the PMSM drive's speed, torque, and stator currents when operating at 314 rad/sec, the rated speed with PI and IHC. To verify effectiveness of IHC the motor starts to rotate at speed of 314r/sec under zero load. A sudden fall is observed in speed response for very less time as load torque of 11Nm is applied at 0.03 seconds after that the drive recovers the reference speed again. It can be seen from the torque waveform in Fig 7 and Fig 8 that the IHC controller is more successful in reducing torque ripples than PI controller.

2. Drive performance with speed variation

Figs 9 and 10, illustrate the simulated results of both PI and IHC controller employed in PMSM drive at speed varying from 157r/sec to 314r/sec. At 157 r/sec, the motor is

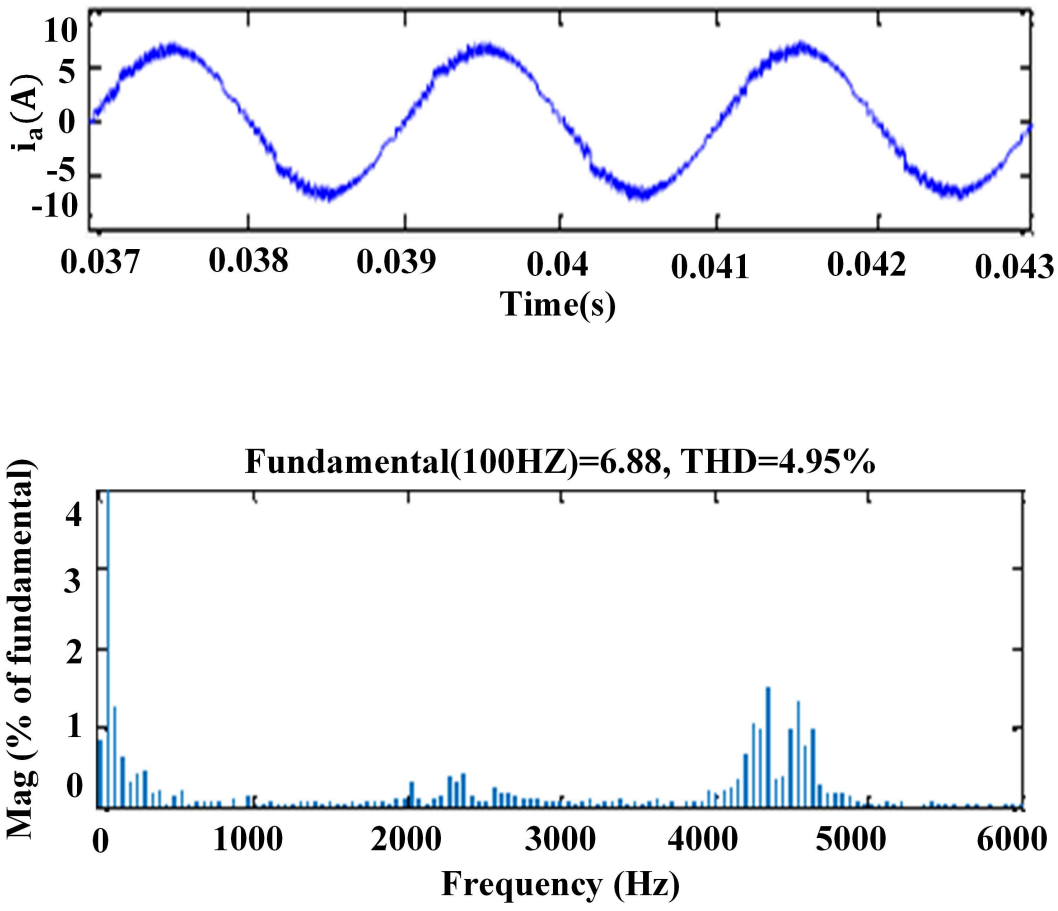


Fig 19. THD in motor current at 157r/sec with IHC.

https://doi.org/10.1371/journal.pone.0312946.g019

started at zero load torque. At 0.05s, step variation of speed is given. Initially the motor is running at no load and at 0.03 sec, 11Nm load is inserted. From the waveform given in Fig 9 and Fig 10, it is important to note that the suggested IHC outperforms the traditional PI in both steady-state and dynamic torque performance, with significantly reduction in ripples.

3. Drive performance in reverse rotation mode

Figs 11 and 12, illustrate the simulated results of both PI and IHC controller employed in PMSM drive with AFC at speed varying from 314r/sec to -314r/sec. Initially reference motor speed of rated value is given and at 0.05 sec step variation in speed 314r/sec to -314r/sec in reverse direction is applied to the motor. The motor speed and torque are observed to seamlessly track their set values. According to the torque waveform given in Fig 11 and Fig 12, there is a significant difference in torque ripples induced by both control methods. It can therefore be stated that proposed method can reduce the ripples in motor torque effectively over a conventional method.

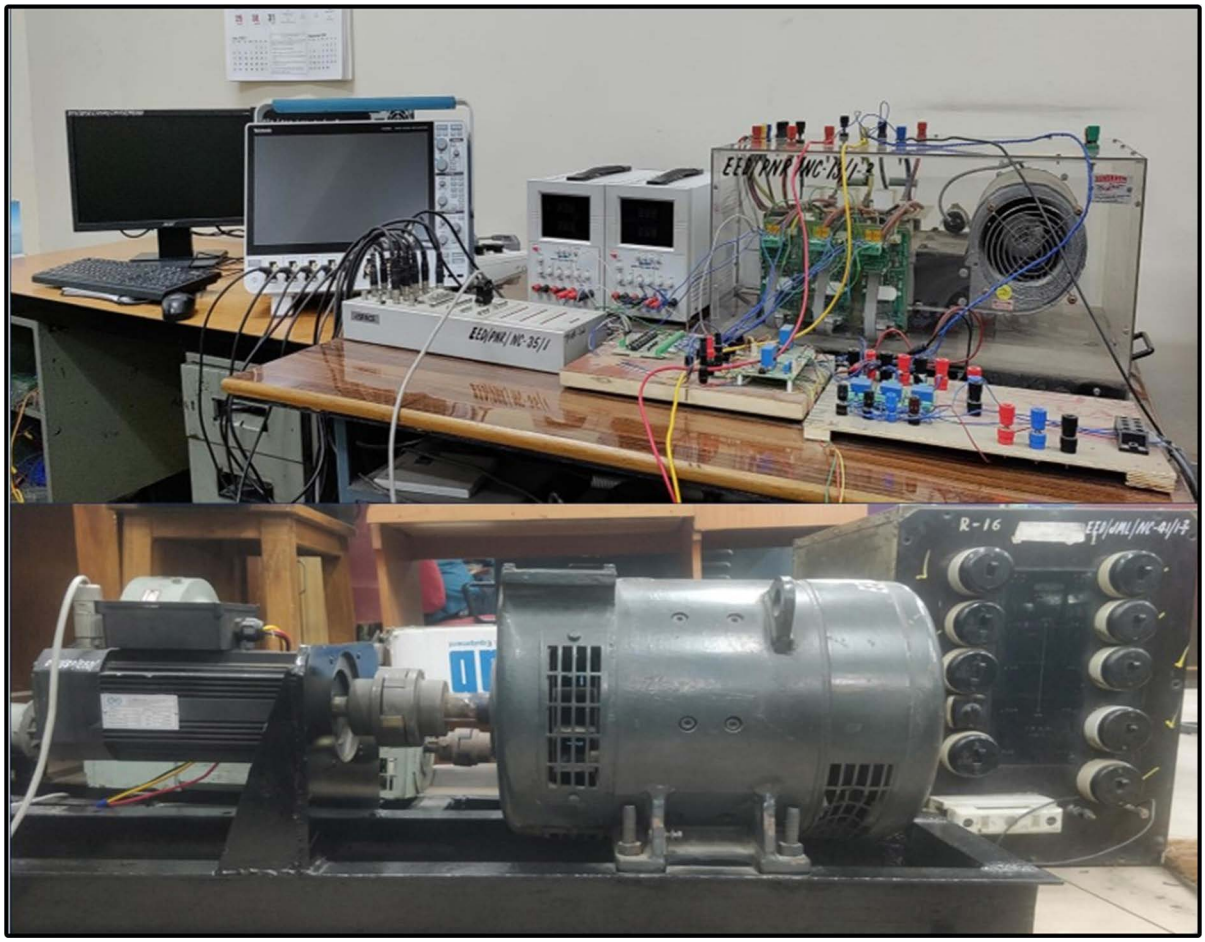


Fig 20. Experimental bench of PMSM drive.

https://doi.org/10.1371/journal.pone.0312946.g020

4. Steady state characteristics of estimated torque and measured torque

Fig 13 and Fig 14 show steady state measured value and estimated value of torque of the PMSM drive using PI and IHC with AFC. The measured value of torque (T_e) is 11Nm and estimate value of torque (T_{est}) is 10.9 Nm at rated speed. The average error between both torque values is observed to be 0.1Nm.

5. Calculation of torque ripple

Torque ripple is calculated using Equation 25 as given below:

$$\text{Torque ripple}(\%) = \frac{T_{max} - T_{min}}{T_e} \times 100 \tag{25}$$

Here T_{max}, T_{min} are maximum and minimum peak of toque which is observed from given torque waveforms. The percentage values of torque ripples running at various speeds with a PI and IHC in PMSM drive is shown in Fig 15. When employing IHC as a speed controller instead of a PI controller, ripples are seen to be minimized.

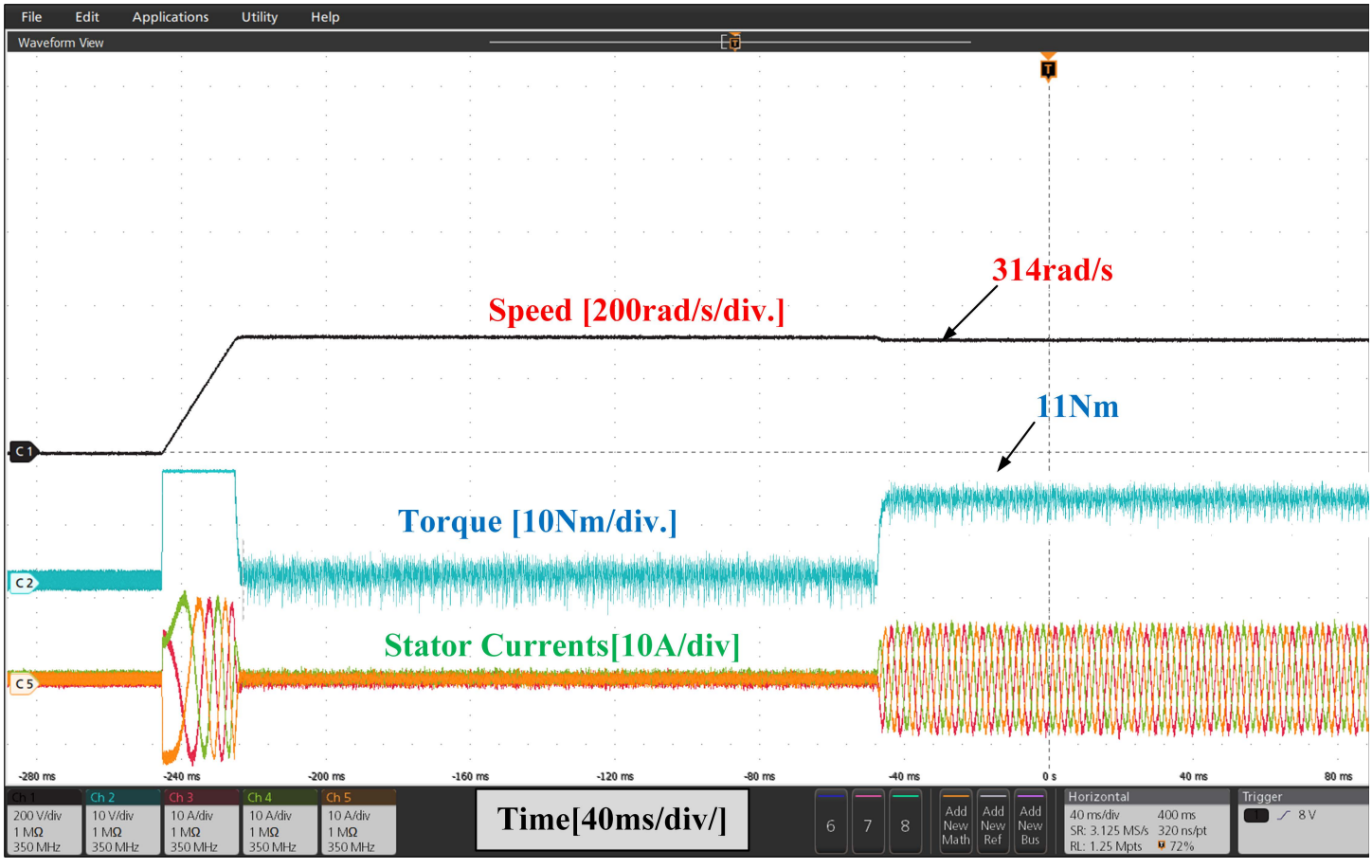


Fig 21. Experimental waveform of W_r, T_e and i_{abc} of motor running at 314r/sec using PI.

https://doi.org/10.1371/journal.pone.0312946.g021

6. THD analysis

The harmonic spectra of the stator current of PMSM with PI and IHC at 314r/sec is shown in Fig 16 and Fig 17. THD measured with PI controller is 5.8%, whereas THD measured with IHC is 4.6%. The harmonic spectra of the stator current for the drive with PI and IHC is shown in Fig 18 and Fig 19 while it is operating at half of its rated speed (157 rad/sec). THD is reported to be 4.95% with IHC and 6.15% with PI controller. The average value of switching frequency for all techniques is fixed constant at 7.5 kHz for simulation experiments. When THD is examined up to 6 kHz, it is found that the PMSM drive with Intelligent Hybrid Controller has lower THD in stator current than the PI controller.

Experimental results demonstration

An experimental setup is developed for the operation of PMSM drive shown in Fig 20. The ds1104 controller, VSI, current and voltage, driver gating circuit, and power sources are the basic components used in the experimental setup of PMSM drive. MATLAB/Simulink and dSPACE1104 are used to implement the different controllers for PMSM drives. Table 1 lists the PMSM parameters utilized for the experimental investigation. The VSI is controlled by

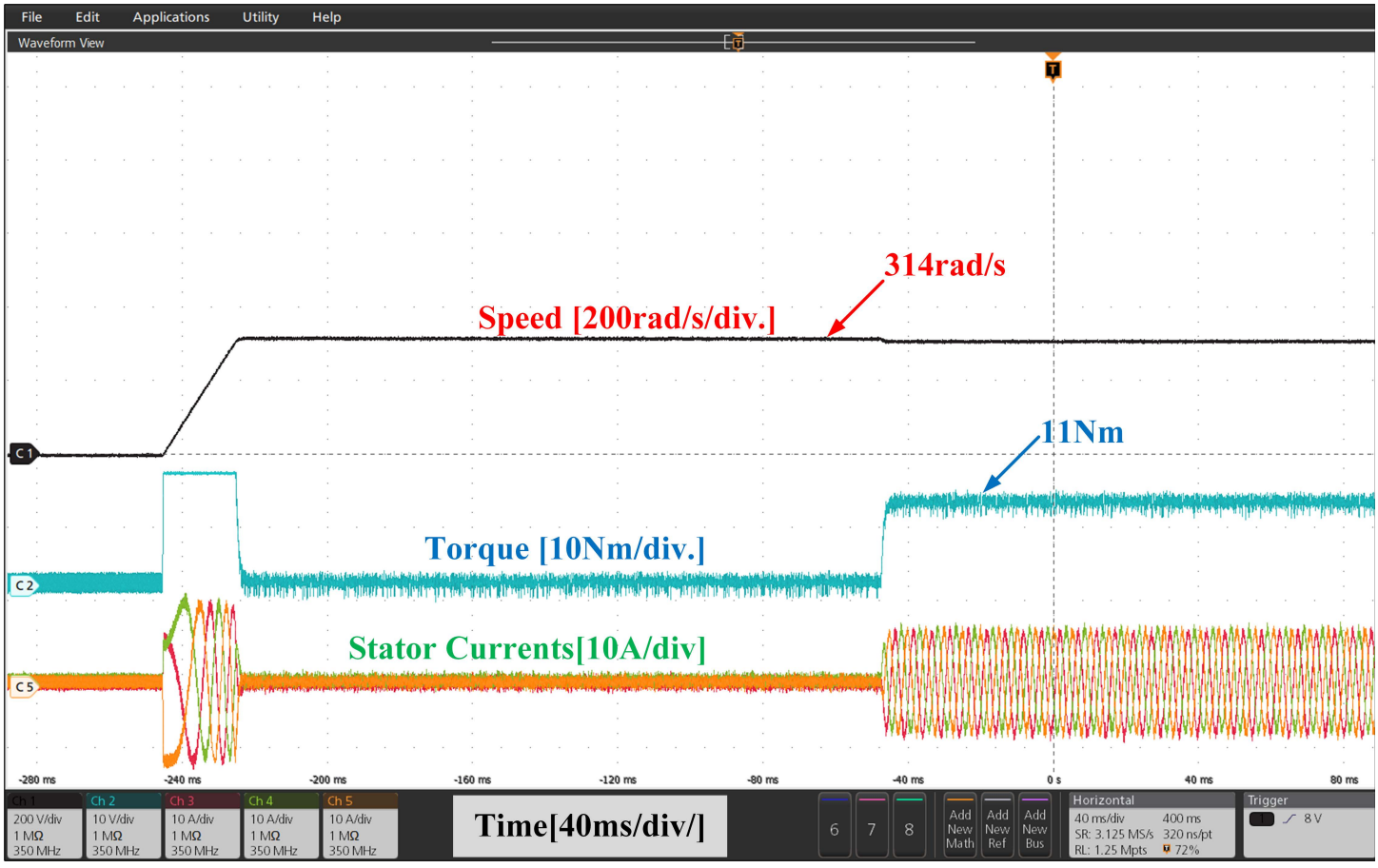


Fig 22. Experimental waveform of W_r, T_e and i_{abc} of motor running at 314r/sec using IHC.

https://doi.org/10.1371/journal.pone.0312946.g022

dS1104 controller and supplies the stator voltage for the PMSM. A DC generator is paired on the same shaft to provide electrical loading for the PMSM.

1. Experimental performance of PMSM drive

Fig 21, and Fig 22, show the experimental results of both PI and IHC employed in PMSM drive operated at 314r/sec. In starting the motor starts to rotate with full speed at zero load and easily trajectory the speed given in reference. When load of 11Nm is inserted at the shaft, a quick drop is perceived in speed curve and after few second the drive trajectories the reference speed effortlessly again. It can be seen from the torque waveform in Figs 7 and 8 that the IHC controller is more successful in reducing torque ripples than PI controller.

Fig 23 and Fig 24, illustrate the experimental results of both PI and IHC controller employed in PMSM drive at speed varying from 157r/sec to 314r/sec. At 157 r/sec, the motor is started at zero load torque. The motor is started to run at zero load condition at 157r/sec. A constant load of 11Nm is inserted to the shaft and it is perceived that the motor generates the net demanded torque without any interval. After motor reaches in steady state, a change in commanded speed from 157r/sec to 314r/sec is applied to motor. A rise in load torque is observed during the speed variation and it reaches again at rated value. The motor speed and

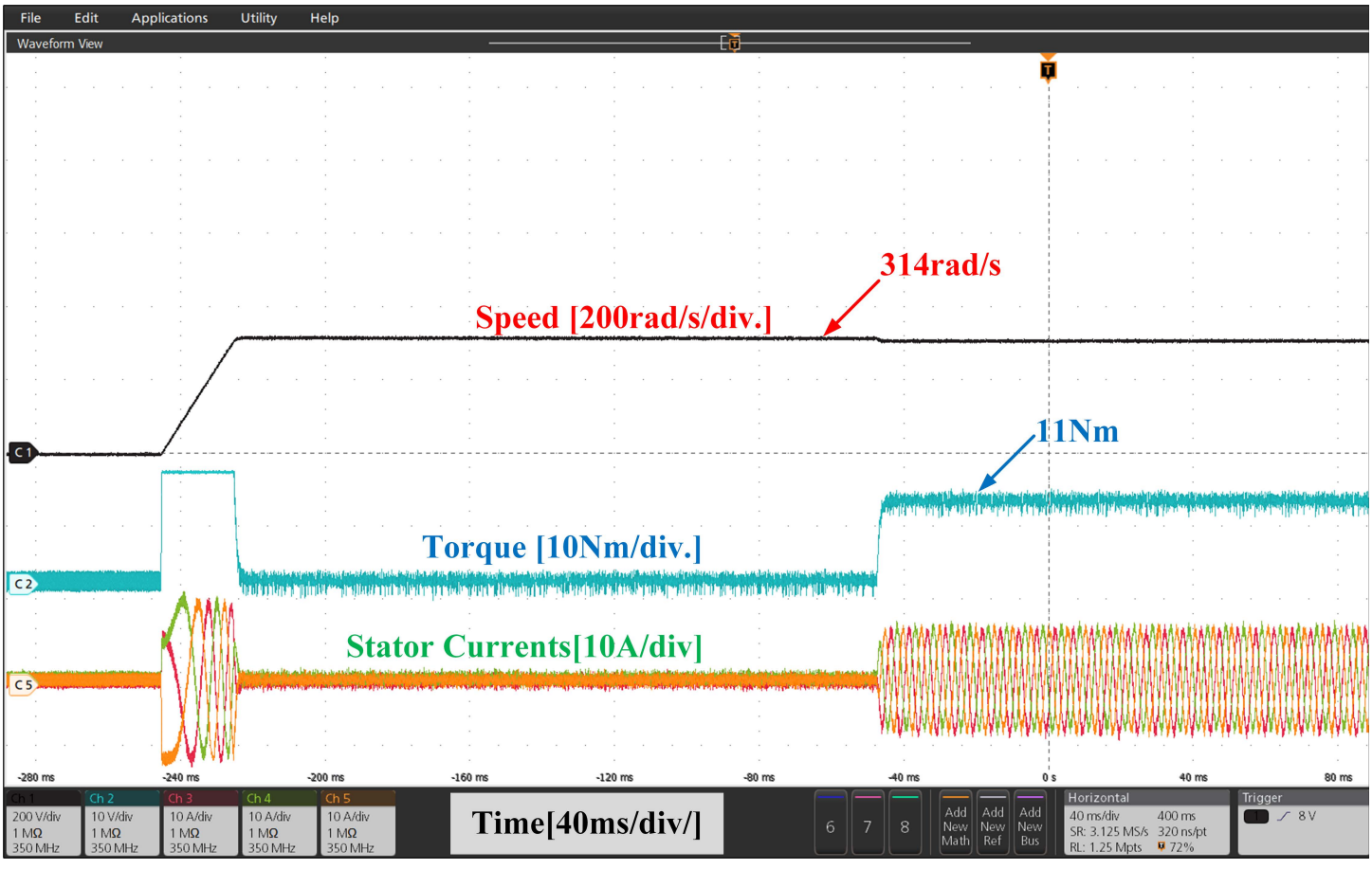


Fig 23. Experimental waveform of W_r, T_e and i_{abc} of motor with linear increase in speed using PI.

https://doi.org/10.1371/journal.pone.0312946.g023

torque are found to follow their set values. The IHC greatly reduces toque ripples, according to experimental analysis of the torque response of motor.

Fig. 25 and Fig. 26, illustrate the experimental results of both PI and IHC controller employed in PMSM drive with AFC at speed varying from 314r/sec to -314r/sec. Initially reference motor speed of rated value is given and after reaching at steady state condition a step variation in speed 314r/sec to -314r/sec in reverse direction is given to the motor. A drop in motor torque is witnessed during change in speed and it reaches again at rated value after few seconds. The motor speed and torque are observed to seamlessly track their set values. According to the torque waveform given in Figs 25 and 26, there is a significant difference in torque ripples induced by both control methods. It can therefore be stated that proposed method can reduce the ripples in motor torque effectively over a conventional method.

2. THD analysis

The harmonic spectra of the stator current of PMSM with PI and IHC at rated speed of 314r/sec is shown in Figs 27 and 28. THD measured with PI controller is 4.272%, whereas THD measured with IHC is 2.931%. The harmonic spectra of the stator current for the drive with PI and IHC is shown in Figs 29 and 30 while it is operating at half of its rated speed (157 rad/

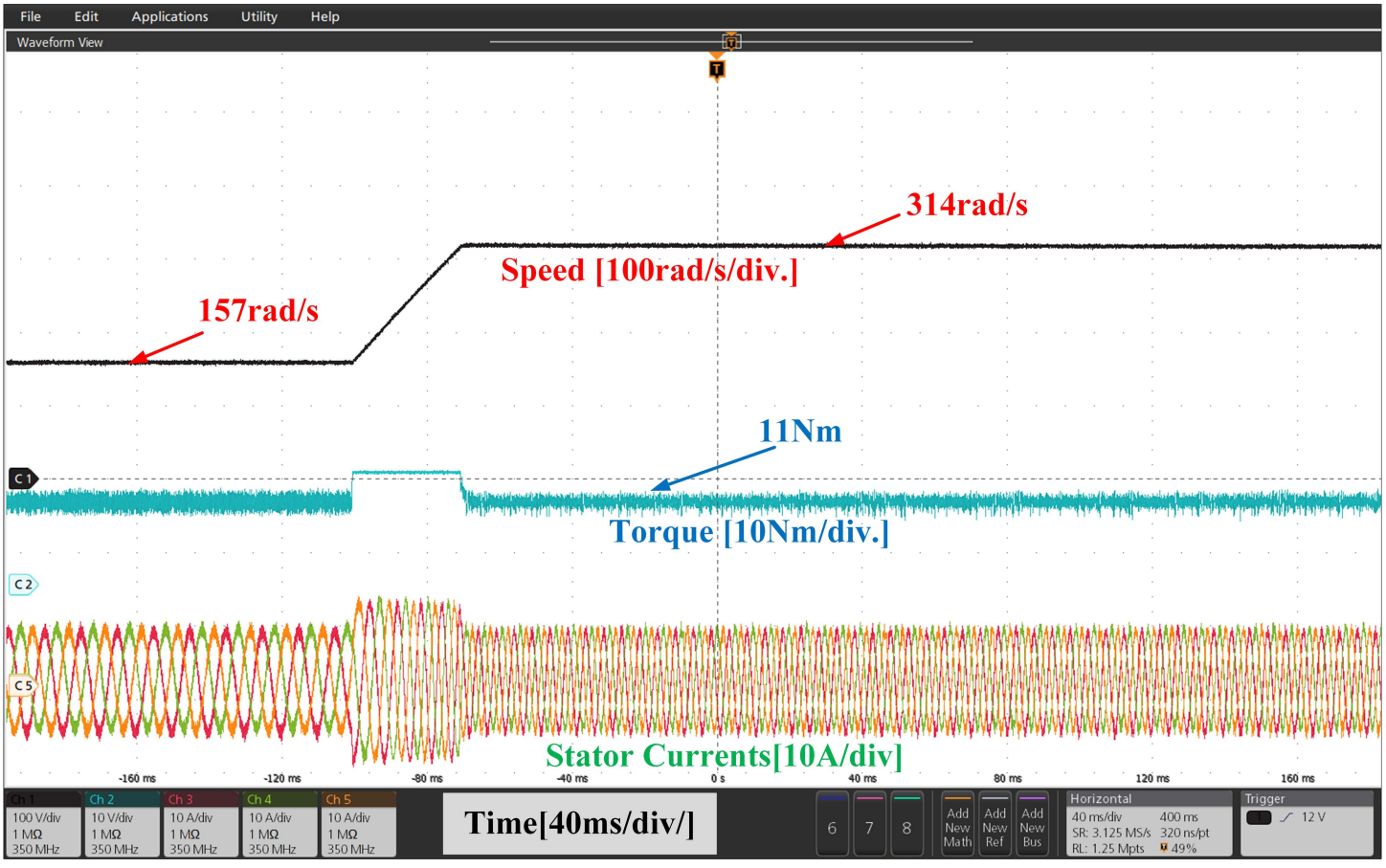


Fig 24. Experimental waveform of W_r, T_e and i_{abc} of motor with linear increase in speed using IHC.

https://doi.org/10.1371/journal.pone.0312946.g024

sec). THD is reported to be 1.848% with IHC and 3.522% with PI controller. When THD is examined by performing experiment, it is found that the PMSM drive with Intelligent Hybrid Controller has lower THD in stator current than the PI controller.

Conclusion

This study proposes an intelligent hybrid controller for Permanent Magnet Synchronous Motor (PMSM) drives, aiming to mitigate the problems arising from the independent operation of standard Proportional-Integral (PI) controller and Fuzzy Logic Controller (FLC). The proposed controller incorporates an adaptive flux and torque estimator. The controller under consideration is formulated by integrating a Proportional-Integral (PI) controller and a Fuzzy Logic Controller (FLC) in a cascaded manner, incorporating innovative switching functionalities. A unique methodology is proposed for the estimation of flux, with the aim of reducing disturbances in current, torque, and flux. In addition, comprehensive simulations and experimental findings have provided evidence that the suggested control methodology presents advantages such as decreased fluctuations in motor torque, lower total harmonic distortions in current, and enhanced overall drive performance in both steady-state and transient states across various operational scenarios, in contrast to the conventional proportional-integral (PI) controller. The torque ripples in motor are observed as 27.70%, 18.18% and

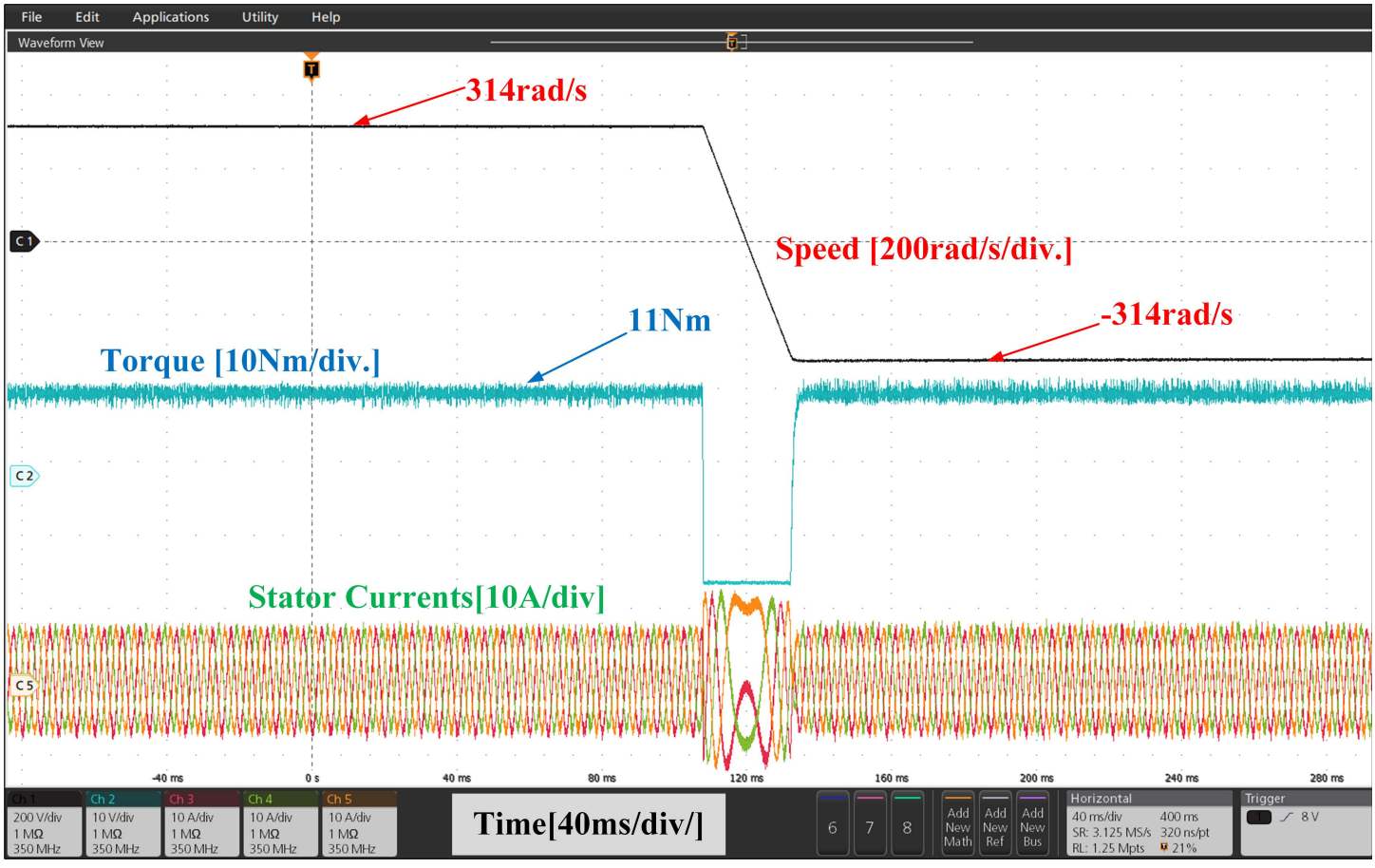


Fig 25. Experimental waveform of W_r, T_e and i_{abc} of motor with reverse rotation of speed using PI.

https://doi.org/10.1371/journal.pone.0312946.g025

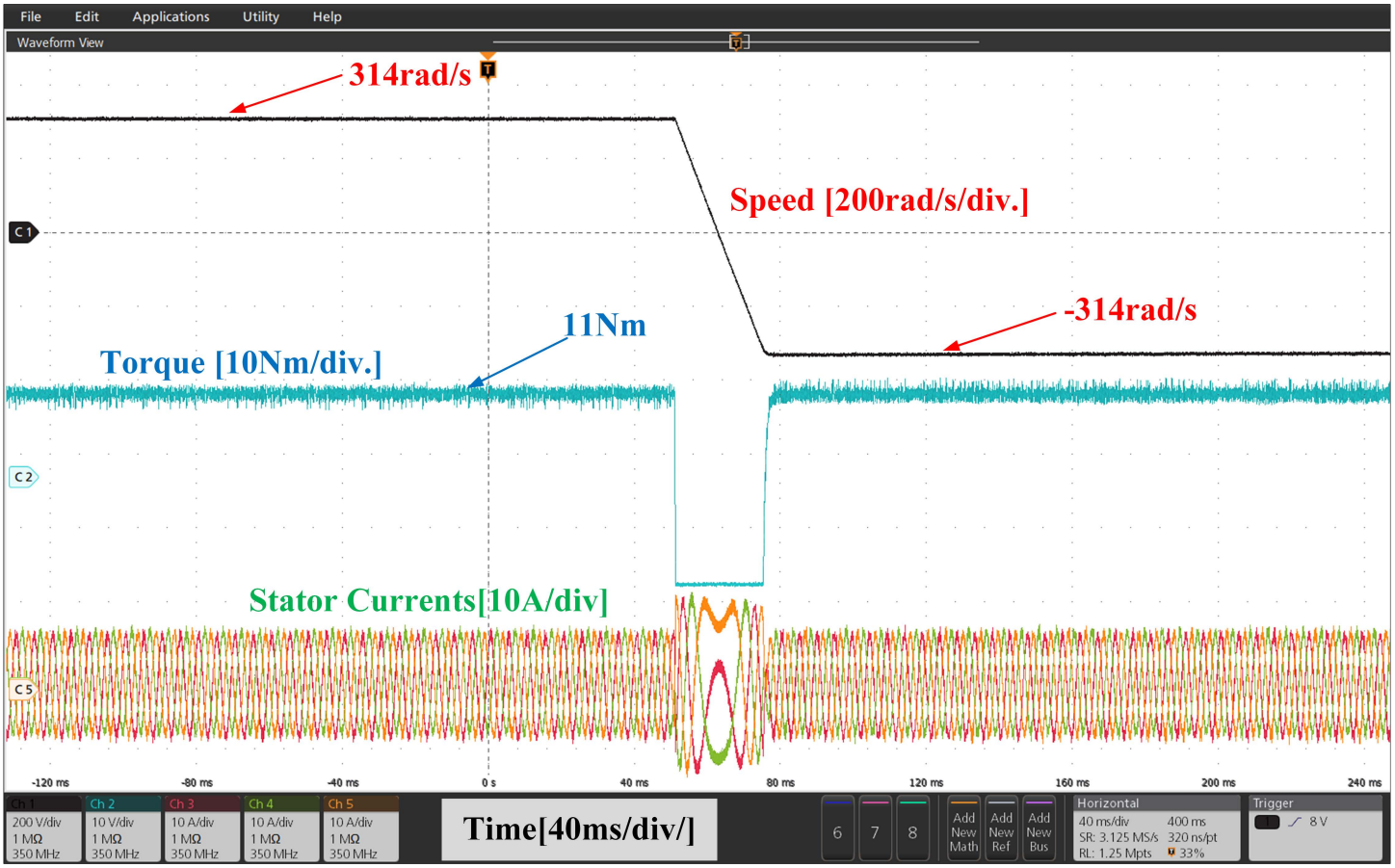


Fig 26. Experimental waveform of W_r, T_e and i_{abc} of motor with reverse rotation of speed using IHC.

https://doi.org/10.1371/journal.pone.0312946.g026

18.90% with PI controller and 10.90%, 9.09% and 9.50% with IHC controller at speed of 157 rad/s, 314 rad/s and -314 rad/s respectively. It shows that IHC can give better performance with reduced ripples in torque than PI controller. THD measured with PI controller is 5.8%, whereas THD measured with IHC is 4.6% at speed of 314 rad/s and 6.15% with PI controller, 4.95% with IHC at speed of 157 rad/s. Thus, IHC controller provides lesser THD in current than PI controller at different operating speed. In future work, the proposed method can be implemented for the electric vehicle application with some adaptive controller based on artificial intelligence.

Acknowledgments

This work was supported by the King Saud University, Saudi Arabia, Deanship of Scientific Research, Saudi Electricity Company Research Chair in Power System Reliability and Security.

The authors would like to express sincere gratitude to Intelligent Prognostic Private Limited Delhi, India for providing the necessary technical support to this research work.

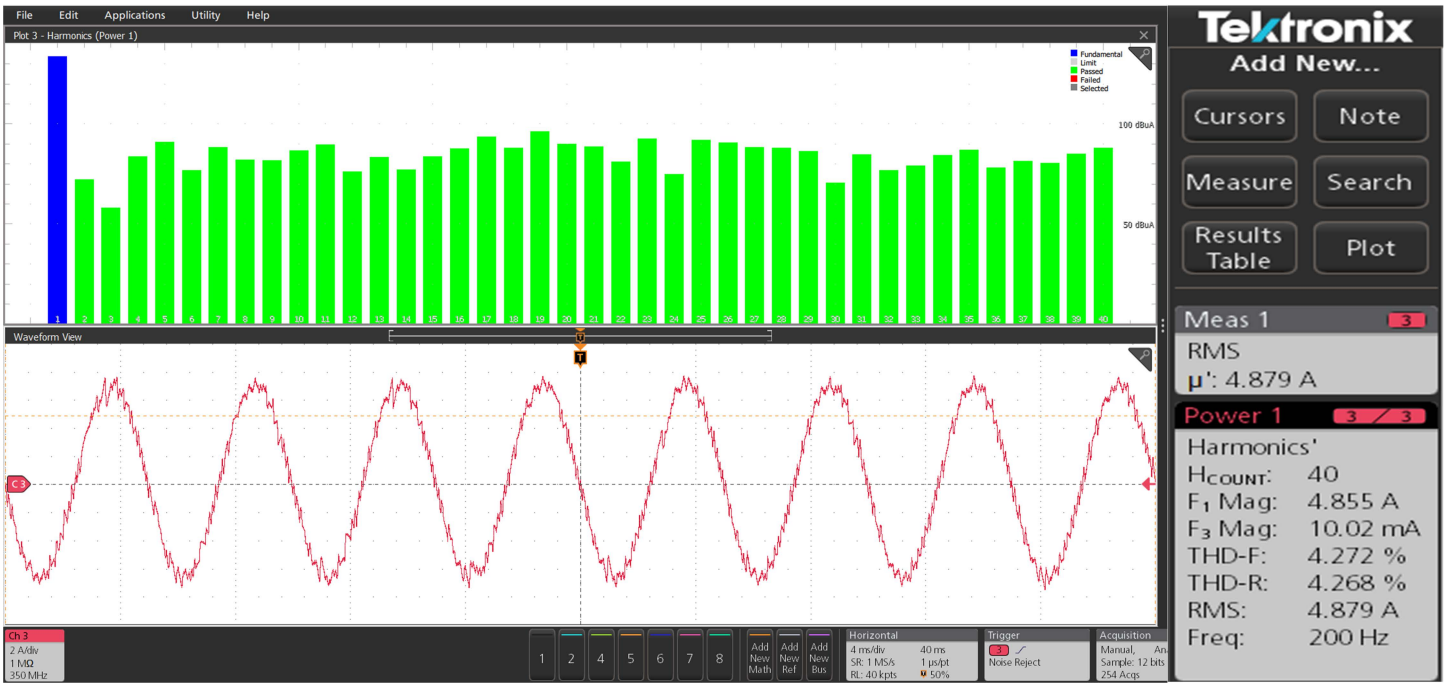


Fig 27. Current THD of PMSM at 314 r/sec with PI.

https://doi.org/10.1371/journal.pone.0312946.g027

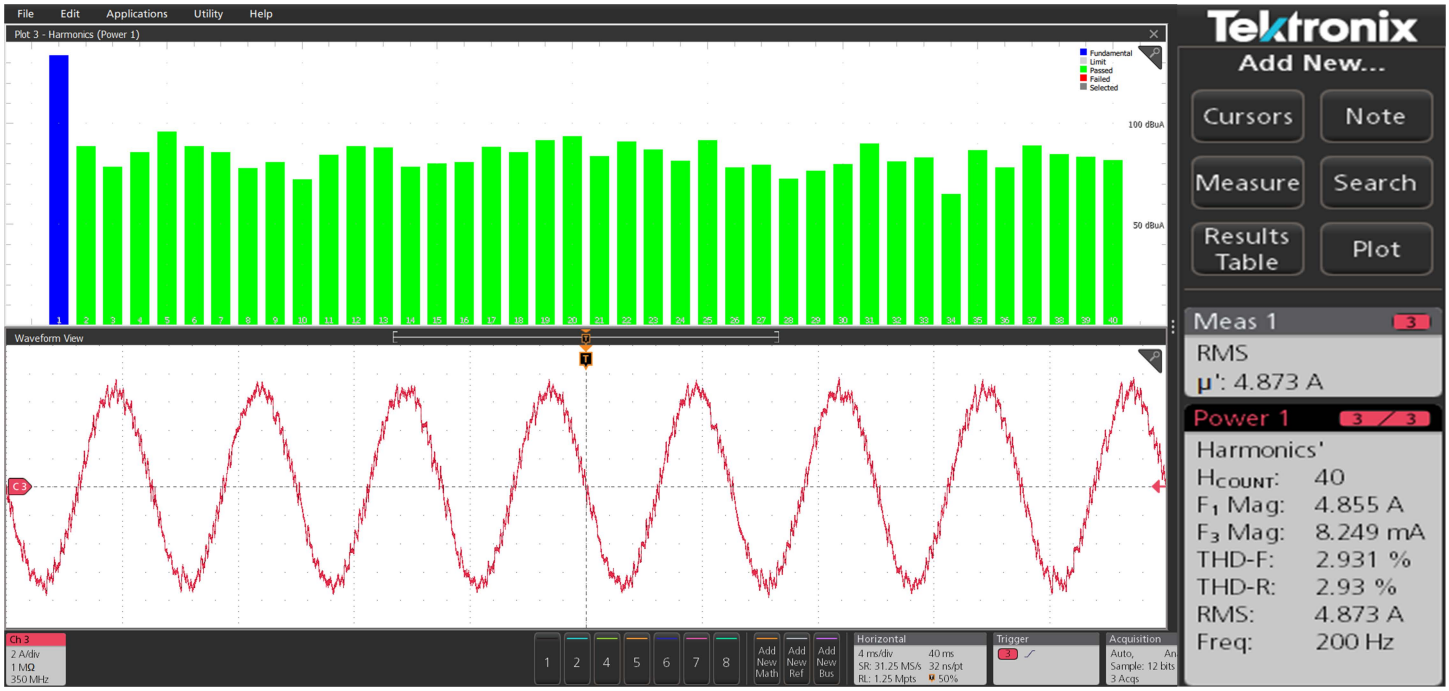


Fig 28. Current THD of PMSM at 314 r/sec with IHC.

https://doi.org/10.1371/journal.pone.0312946.g028

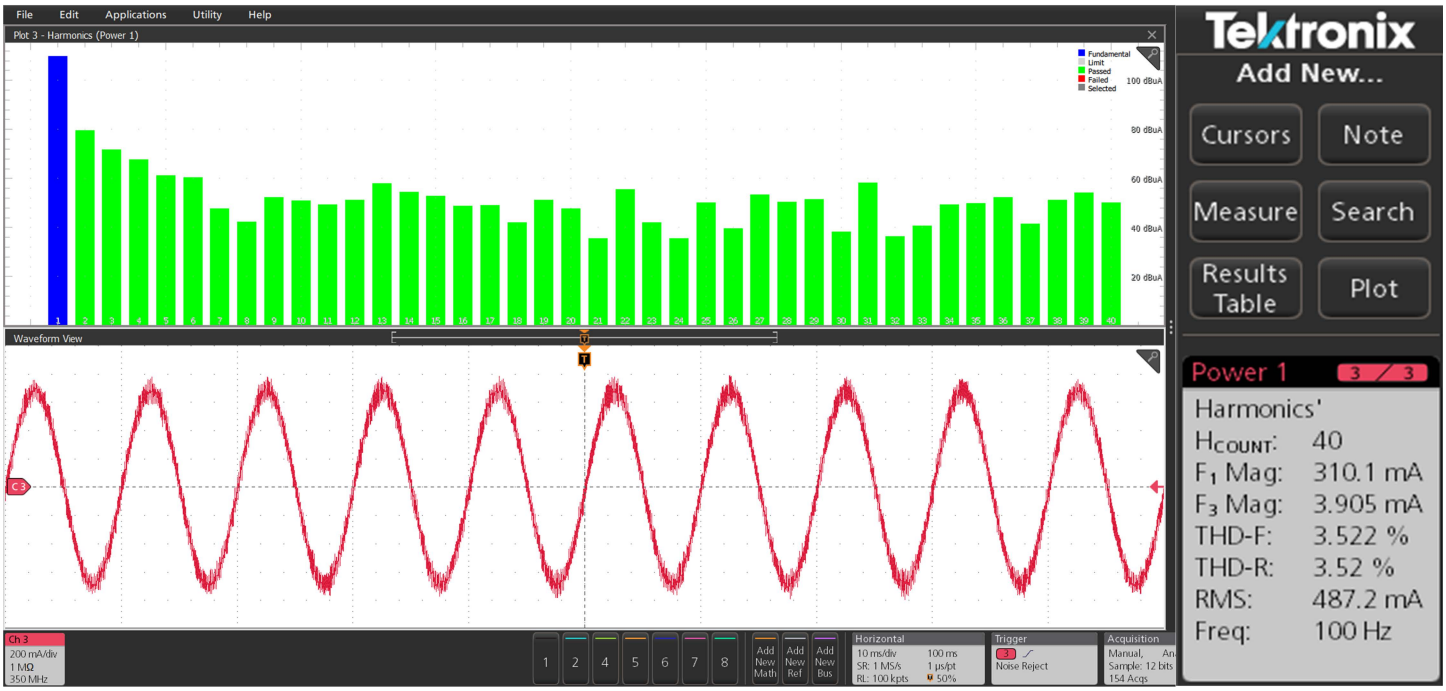


Fig 29. Current THD of PMSM at 157 r/sec with PI.

https://doi.org/10.1371/journal.pone.0312946.g029

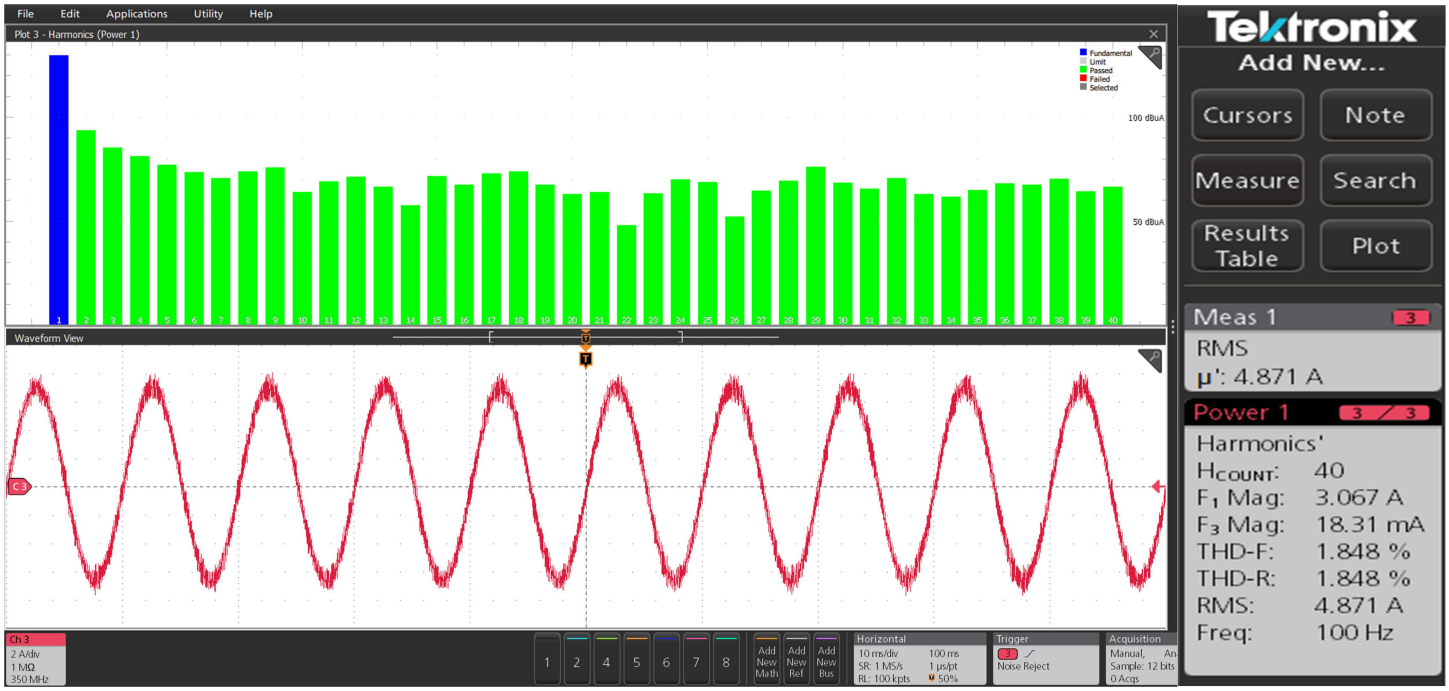


Fig 30. Current THD of PMSM at 157 r/sec with IHC.

https://doi.org/10.1371/journal.pone.0312946.g030

Author contributions

Conceptualization: Surya Kant, Mini Sreejeth, Madhusudan Singh, Ambrish Devanshu, Majed A. Alotaibi, Hasmat Malik, Fausto Pedro García Márquez.

Data curation: Surya Kant, Mini Sreejeth, Madhusudan Singh, Ambrish Devanshu, Majed A. Alotaibi, Hasmat Malik, Fausto Pedro García Márquez.

Formal analysis: Surya Kant, Mini Sreejeth, Madhusudan Singh, Ambrish Devanshu, Majed A. Alotaibi, Hasmat Malik, Fausto Pedro García Márquez.

Funding acquisition: Majed A. Alotaibi, Hasmat Malik.

Investigation: Surya Kant, Mini Sreejeth, Majed A. Alotaibi, Hasmat Malik, Fausto Pedro García Márquez.

Methodology: Surya Kant, Majed A. Alotaibi, Hasmat Malik, Fausto Pedro García Márquez.

Project administration: Majed A. Alotaibi, Hasmat Malik, Fausto Pedro García Márquez.

Resources: Surya Kant, Mini Sreejeth, Madhusudan Singh, Ambrish Devanshu, Majed A. Alotaibi, Hasmat Malik, Fausto Pedro García Márquez.

Software: Madhusudan Singh, Ambrish Devanshu, Majed A. Alotaibi, Hasmat Malik, Fausto Pedro García Márquez.

Supervision: Mini Sreejeth, Majed A. Alotaibi, Hasmat Malik.

Validation: Surya Kant, Ambrish Devanshu, Majed A. Alotaibi.

Visualization: Ambrish Devanshu, Majed A. Alotaibi.

Writing – original draft: Surya Kant, Madhusudan Singh, Majed A. Alotaibi, Fausto Pedro García Márquez.

Writing – review & editing: Madhusudan Singh, Ambrish Devanshu, Majed A. Alotaibi.

References

1. Harib KH, Khousa EA, Ismail A. Field oriented motion control of a 3 phase permanent magnet synchronous motor. In 2nd International Conference on Electric Power and Energy Conversion Systems (EPECS). Sharjah; 2011. p. 1–7.
2. Rahman MA, George GH, Radwan TS, Uddin MN. Performance of current controllers for VSI-fed IPMSM drive. IEEE Trans Ind Applicat. 2000;36(6):1531–8. https://doi.org/10.1109/28.887203
3. Pillay P, Krishnan R. Modeling, simulation, and analysis of permanent-magnet motor drives. part I. The permanent-magnet synchronous motor drive. IEEE Trans Ind Applicat. 1989;25(2):265–73. https://doi.org/10.1109/28.25541
4. Dhaoudi R, Mohan N. Analysis of current regulated voltage source inverter for permanent magnet synchronous motor drives in normal and extended speed ranges. IEEE Trans Energy Convers. 1990;5(1):137–44.
5. Alireza AK, Davood R, Jose R. Predictive control of permanent magnet synchronous motor with non-sinusoidal flux distribution for torque ripple minimization using the recursive least square identification method. IET Electric Power Applicat. 2017;11(5):847–56.
6. Türker U, Buyukkeles AFB. A robust predictive current controller for PMSM drives. IEEE Trans Ind Electron. 2016;63(6):3906–14.
7. Suryakant M, Sreejeth M, Singh AK. Minimization of torque ripples in PMSM drive using PI- resonant controller-based model predictive control. ElectrEng. 2022;105:207–19.
8. Qiao Z, Shi T, Wang Y, Yan Y, Xia C, He X. New sliding-mode observer for position sensorless control of permanent-magnet synchronous motor. IEEE Trans Ind Electron. 2013;60(2):710–9. https://doi.org/10.1109/tie.2012.2206359
9. Mani P, Rajan R, Shanmugam L, Joo YH. Adaptive fractional fuzzy integral sliding mode control for PMSM model. IEEE Trans Fuzzy Syst. 2019;27(8):1674–86. https://doi.org/10.1109/tfuzz.2018.2886169

10. Wai R-J, Chen M-W, Liu Y-K. Design of adaptive control and fuzzy neural network control for single-stage boost inverter. IEEE Trans Ind Electron. 2015;62(9):5434–45. https://doi.org/10.1109/tie.2015.2408571
11. Merzoug MS, Naceri F. Comparison of field-oriented control and direct torque control for permanent magnet synchronous motor. Int J Elect Comput Eng. 2008;2(9):1797–801.
12. Asgharpour-alamdari H, Alinejad-beromi YH, Yaghobi Y. A fuzzy-based speed controller for improvement of induction motor's drive performance. Iran J Fuzzy Syst. 2016;13(2):61–70.
13. Uddin MN, Rahman MA. Fuzzy logic based speed control of an IPM synchronous motor drive. Engineering Solutions for the Next Millennium. IEEE Canadian Conference on Electrical and Computer Engineering. Alberta, Canada: 1999. p. 1259–64.
14. Uddin MN, Chy MMI. A novel fuzzy-logic-controller-based torque and flux controls of IPM synchronous motor. IEEE Trans Ind Applicat. 2010;46:1220–9.
15. Sant A, Rajagopal KR. PM synchronous motor speed control using hybrid fuzzy-PI with novel switching functions. IEEE Trans Magnet. 2009;45(10):4672–5.
16. Sant A, Rajagopal KR, Sheth NK. Permanent magnet synchronous motor drive using hybrid PI speed controller with inherent and non-inherent switching functions. IEEE Trans Magnet. 2011;47(10):4088–91. https://doi.org/insert-doi-here
17. Zerikat M, Chekroun S. Design and implementation of a hybrid fuzzy controller for a high-performance induction motor. In Proceeding of World Academy of Science, Engineering and Technology. Int J Comput Inf Eng. 2007;20:263–9.
18. Rubaai A, Ricketts D, Kankam MD. Experimental verification of a hybrid fuzzy control strategy for a high-performance brushless DC drive system. IEEE Trans on Ind Applicat. 2001;37(2):503–12. https://doi.org/10.1109/28.913715
19. Singh B, Singh BP, Dwivedi S. DSP based implementation of hybrid fuzzy PI speed controller for direct torque controlled permanent magnet synchronous motor drive. International J Emerg Elect Power Syst. 2007;8(2):1–22. https://doi.org/10.2202/1553-779x.1246
20. Nakao N, Akatsu K. Suppressing pulsating torques: torque ripple control for synchronous motors. IEEE Ind Appl Mag. 2014;20(6):33–44. https://doi.org/10.1109/mias.2013.2288383
21. Truong P, Flieller D, Nguyen NK, Mercklé J, Sturtzer G. Torque ripple minimization in non-sinusoidal synchronous reluctance motors based on artificial neural networks. Elect Power Syst Res. 2016;140:37–45.
22. Flieller DN, Nguyen K, Wira P, Sturtzer G, Abdeslam D, Merckle J. A self-learning solution for torque ripple reduction for nonsinusoidal permanent-magnet motor drives based on artificial neural networks. IEEE Trans Ind Electron. 2014;61(2):655–66. https://doi.org/10.1109/TIE.2013.2281234
23. Feng G, Lai C, Kar NC. A closed-loop fuzzy-logic-based current controller for pmsm torque ripple minimization using the magnitude of speed harmonic as the feedback control signal. IEEE Trans Ind Electron. 2017;64(4):2642–53. https://doi.org/10.1109/tie.2016.2631524
24. Xiao G, Wang X, Tu W, Tang L, Zhang H, Yang K. A sensor less control of surface-mount permanent magnet synchronous motor based on rotor flux estimation. 20th International Conference on Electrical Machines and Systems (ICEMS). Sydney, NSW: 2017. p. 1–4.
25. Devanshu A, Singh M, Kumar N. An improved nonlinear flux observer based sensorless FOC IM drive with adaptive predictive current control. IEEE Trans Power Electron. 2020;35(1):652–66.
26. Devanshu A, Singh M, Kumar N. Nonlinear flux observer-based feedback linearisation control of IM drives with ANN speed and flux controller. Int J Electron. 2021;108(1):139–61.
27. Kamel HM, Hasanien HM, Ibrahim HE. Speed control of permanent magnet synchronous motor using fuzzy logic controller. Proceedings of the IEEE International Conference on Electric Machines and Drives. Miami, Florida, USA.: 2009. p. 1587–91.
28. Chaoui H, Sicard P. Adaptive Fuzzy Logic Control of Permanent Magnet Synchronous Machines With Nonlinear Friction. IEEE Trans Ind Electron. 2012;59(2):1123–33. https://doi.org/10.1109/tie.2011.2148678
29. Yu JS, Kim SH, Lee BK, Won CY, Hur J. Fuzzy-logic-based vector control scheme for permanent-magnet synchronous motors in elevator drive applications. IEEE Trans Ind Electron. 2007;54(4):2190–200. https://doi.org/10.1109/tie.2007.894692
30. Krishnan R, Beutler AJ. Performance and design of an axial field permanent magnet synchronous motor servo drive. Proceedings of IEEE IAS Annual Meeting. 1985. p. 634–40.
31. Lajoie-Mazenc M, Villanueva C, Hector J. Study and implementation of hysteresis controlled inverter on a permanent magnet synchronous machine. IEEE Trans on Ind Applicat. 1985;IA-21(2):408–13. https://doi.org/10.1109/tia.1985.349662

32. Aner M, Benaifa E, Nowicki A. A PMSM drive design with inverter-stage soft-switching hysteresis current control and space vector modulation for two-level operation of a very sparse matrix converter. IEEE Electrical Power & Energy Conference (EPEC). Montreal, QC; 2009. p. 1–8.

33. Suryakant M, Sreejeth M, Singh M. Performance Analysis of PMSM Drive using Hysteresis Current Controller and PWM Current Controller. Proceedings of the 2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS). Bhopal, India; 2018. p. 1–5.

34. Suryakant M, Sreejeth M, Singh M. Sensor-less control of PMSM drive with BEMF based MRAC algorithm. Proceedings of the 2019 International Symposium on Advanced Electrical and Communication Technologies (ISAECT). Rome, Italy; 2019. p. 1–6.

35. Panda SK, Xu J-X, Qian W. Review of torque ripple minimization in PM synchronous motor drives. 2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century. Pittsburgh, PA; 2008. p. 1–6.

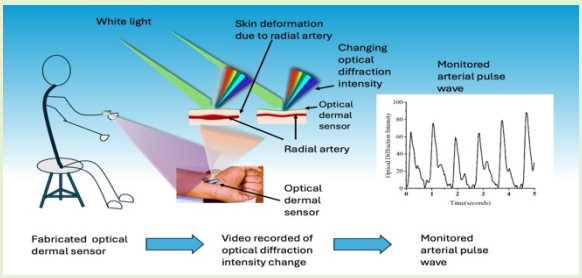
36. Shukla S, Sreejeth M, Singh M. Minimization of ripples in stator current and torque of PMSM drive using advanced predictive current controller based on deadbeat control theory. J Power Electron. 2021;21:142–52.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2025.3560222

Diffractive optical dermal sensor for non-invasive arterial pulse wave monitoring

Mayank Pandey , *Sang-Woo Seo

***Abstract*— The arterial pulse wave is a crucial indicator for assessing cardiovascular health. Wearable pulse detection technology has the potential to transform daily healthcare monitoring that helps to identify cardiovascular diseases. However, the technology faces challenges due to its limited flexibility, low sensitivity, and the need for cost-effective manufacturing to enable widespread implementation. To address these issues, we introduce an innovative, non-invasive sensor made of a deformable optical grating sensor patch. This sensor allows for real-time acquisition of arterial pulse waveforms at the radial artery site. It operates by detecting strain-induced changes in the diffraction pattern that occur due to fluctuations in blood pressure. When illuminated with light, the optical grating produces a diffraction pattern that shifts in response to pulsatile blood flow. These spectral shifts are then processed and analyzed to measure changes in blood pressure, enabling real-time tracking of dynamic arterial pulse waves. With its lightweight and flexible design, along with a low-cost manufacturing approach, this sensor has the potential to be used in wearable health monitoring systems.**



***Index Terms*— Arterial pulse wave monitoring, Cardiovascular health monitoring, Optical sensor, Radial artery, Wearable pulse wave detection**

I. Introduction

THE arterial pulse wave (APW) is a vital component of human physiological signaling, carrying abundant health information that can help identify cardiovascular diseases (CVDs) [1], [2], [3]. CVDs such as ischemic heart disease and stroke, are among the leading global causes of mortality, responsible for approximately one-third of all deaths annually [4]. By 2030, this number is projected to rise to 23.6 million[5]. Despite the high mortality rates associated with CVDs[6], many of these conditions - including arteriosclerosis [7], diabetes [8], myocardial infarction[9], coronary heart disease[10], and hypertension[11] - can be prevented and treated through early diagnosis and continuous monitoring of physiological signals. While physical symptoms of CVDs are often subtle in the early stages, they can be detected through small changes in APW patterns [12]. APWs reflect the interaction between the driving force of the heart and the impedance of blood flow, serving as reliable indicators of cardiovascular health[13], [14]. These waveforms can be measured at various arterial sites, each with unique characteristics.

Manuscript received Feb 7 2025. “This work was supported by the National Science Foundation under Grant (NSF-2324052).”

Mayank Pandey is with the Department of Electrical Engineering City College of New York, New York, 10031, NY, USA (e-mail: pmayank000@citymail.cuny.edu)

*Sang-Woo Seo is with the Department of Electrical Engineering, City College of New York, New York, 10031, NY, USA (e-mail: swseo@ccny.cuny.edu)

Recent advancements in wearable devices equipped with flexible sensors and Internet of Things technologies have enabled near-real-time monitoring of APWs. Common non-invasive techniques for measuring APWs include electrocardiography (ECG), photoplethysmography (PPG), and remote photoplethysmogrpahy (rPPG)[15], [16], [17]. However, these methods have their limitations: ECG requires forming a circuit through the heart, complicating the miniaturization of devices, while PPG is energy-intensive due to its reliance on light-emitting diodes and photodetectors. rPPG allows remote sensing but its accuracy can be affected by variations in illumination, skin tone, and other measurement conditions. Skin-contact sensors, such as piezoresistive, piezoelectric, triboelectric, capacitive, or microelectromechanical systems (MEMS), are gaining attention for transducing APWs [18], [19], [20], [21]. These sensors can effectively capture APWs, but the reliability of the signal heavily depends on the sensor's position. The radial artery, which is typically 3-4 mm in diameter and located beneath the skin, poses challenges for precise sensor placement. Even slight misplacement can compromise signal accuracy, making the process dependent on the accuracy of the placement. Although various technologies exist for measuring APWs, they face challenges related to accuracy, portability, motion sensitivity, environmental interference, and power efficiency. Emerging wearable technologies are addressing these limitations, but further research is necessary to achieve an optimal balance between precision, comfort, and practicality. Continued advancements will pave the way for reliable and

XXXX-XXXX © XXXX IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:39:50 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2025.3560222

widespread use of APW monitoring systems in continuous health monitoring.

In this paper, we present a diffractive optical patch designed to detect subtle skin deformations associated with APWs. While diffractive optical detection schemes have been used for various applications, such as force measurements[22], monitoring local pressure in microfluidic devices [23] and acousto-optic sensors[24], there are few reports focused on monitoring APWs for human health. The sensor features a thin and flexible grating structure that can be applied to the radial arterial area, allowing it to visualize subtle skin deformations through variations in light intensity. The operation of the sensor is based on detecting strain-induced changes in the grating's diffraction condition caused by fluctuations in blood pressure. When illuminated with white light, the grating produces a diffraction pattern that shifts in wavelength as the pulsatile blood flow alters the diffraction conditions. These spectral shifts are analyzed to quantify changes in blood pressure waves, providing a reliable measurement of physiological conditions. This innovative design facilitates accurate detection of APWs without the need for bulky or complex sensing systems. Its lightweight and flexible structure make it particularly suitable for wearable health monitoring devices. The sensor's capability to monitor dynamic blood pressure changes in real time highlights its potential for integration into advanced healthcare systems, enabling continuous, non-invasive tracking of cardiovascular health. Moreover, the sensor can be manufactured using a cost-effective and scalable approach. These advantages promise to significantly enhance cardiovascular health monitoring by offering a reliable, accurate, and economical solution for tracking critical health metrics.

II. Design and Fabrication

Figure 1(a) presents a schematic design of the proposed sensor, which serves as a simple but sensitive device for detecting arterial pulse waves (APWs) on a target skin surface. The optical grating sensor is made from flexible polydimethylsiloxane (PDMS) elastomer. It features a deformable reflection-type grating structure coated with a metal layer to enhance its optical diffraction. The elastic properties of PDMS allow the grating to adopt a dynamic and deformable structure that responds to subtle changes in arterial pressure. Since it covers the broad skin area of radial artery, dynamic optical intensity change can be easily visualized without dedicated sensor positioning. Additionally, the unique optical sensing method operates without physical wiring, enabling remote detection of APWs and improving its application in non-invasive cardiovascular monitoring. Figure 1(b) presents image of the arterial pulse sensor on the wrist, indicating the region with vivid optical signal quality for APW measurement. Pressure waves are generated by the rhythmic contractions and expansions of the heart muscles, which propagate through the arterial vessels during the systolic and diastolic phases. These pressure waves produce the pulse observed in arteries. In the systolic phase, the arterial diameter expands (known as vasodilation) due to increased blood flow, while in the diastolic phase, the diameter reduces (referred to as vasoconstriction). These cyclic changes in arterial diameter are transmitted through soft tissues and muscles, causing surface skin pulsations that can be detected to assess the dynamics of the heart and arterial walls.

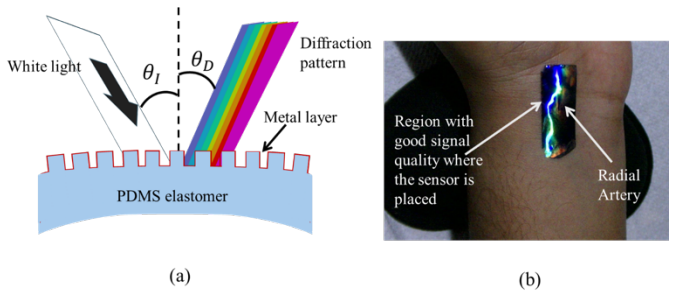


Fig. 1. (a) Schematic design of the proposed diffractive optical dermal sensor. (b) An illustration of a radial artery region where the sensor is positioned to measure APWs.

The detection mechanism of the proposed grating sensor is based on the principle of light diffraction. When incident white light interacts with the periodic structure of the grating grooves, it disperses into beams of various wavelengths. Each groove acts as a source of scattered light that interferes with light from the adjacent grooves. This interference produces alternating light and dark bands, which are characterized by constructive and destructive interference. Constructive interference occurs when light waves from adjacent grooves are in phase, resulting in bright spots that correspond to the diffraction orders. The diffraction pattern created by the grating is governed by the grating equation:

$$m\lambda = d(\sin\theta_I + \sin\theta_D)$$

Where m is the diffraction order, d is the grating period, λ is the wavelength, θ_I Is the incident angle measured from the normal of the grating, and θ_D is the diffraction angle measured from the normal of the grating. This equation shows that the diffraction angle θ_D varies depending on the wavelength λ of the light and grating period d [22], [25]. When the grating structure on the skin surface deforms in response to pulsatile blood flow, the diffraction conditions of the light are altered. As illustrated in Figure 1(b), the PDMS grating sensor is placed on the skin surface of the wrist. The arterial pressure waves periodically stretch the grating sensor, altering the diffraction conditions and causing fluctuations in the reflected light. Since the sensor is patched on a relatively broad area of the skin, the most noticeable color changes can be visualized on the specific spot area where the skin surface is highly modulated. Capturing these light spot shifts with a camera enables real-time monitoring of APWs.

Figure 2(a) illustrates the fabrication process of the PDMS grating sensor, highlighting its cost-effective production. The master mold for the PDMS grating sensor was simply generated using a low-cost DVD optical disk, which consists of two polycarbonate (PC) layers that were separated manually. Subsequently, scotch tape was employed to remove any remaining layers from the PC-grating structure. Finally, the dye layer was eliminated by immersing the structure in isopropanol and allowing it to dry in ambient air. The master mold was subsequently cut into several rectangular shapes to prepare the grating (Step 1). To prepare the PDMS grating sensor, a 10:1 weight ratio of the PDMS base (Sylgard 184 silicone elastomer) and the cross-linker was prepared and thoroughly mixed. The

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:39:50 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2025.3560222

PDMS mixture was then degassed in a vacuum chamber to remove air bubbles. Sensors with two different thicknesses were prepared by spinning the PDMS mixture onto the master mold at different spin speeds (Step 2 and 3). After this, the master mold with the PDMS film was placed in a vacuum chamber again to remove any residual air bubbles. It was then transferred to an oven and heated at 60°C for 4 hours to ensure complete curing of the PDMS film. Next, isopropyl alcohol was utilized as a lubricant for the careful detachment of the PDMS grating from the master mold (Step 4). The thickness of each PDMS grating film was then measured using the laser displacement system (Keyence LC-2400A, USA), which indicated a thickness of 250 µm for 300 RPM spin speed and 100 µm for 900 RPM spin speed. Finally, layers of 10nm chromium (Cr) and 50nm copper (Cu) were thermally deposited onto the fabricated grating surface (Step 5 and 6). This deposition significantly enhanced the reflective-diffraction from the PDMS grating film. Figure 2(b) shows an atomic force microscope (AFM) image of the master mold, which illustrates the grating pattern to be replicated on the PDMS film. The AFM image demonstrates the surface morphology and alignment of the fabricated grooves. The grating period and the modulation depth of the master mold were measured at 760 nm and 185 nm, respectively. Figure 2(c) presents a microscope image of the PDMS grating, with a scale bar indicating 10 µm. This confirms that the PDMS grating was successfully fabricated with the intended grating period and modulation depth, closely matching the master mold.

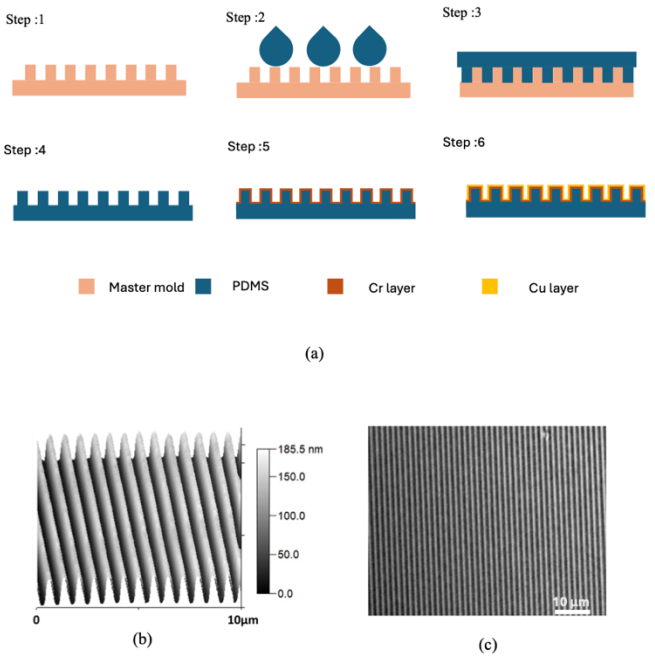


Fig. 2. (a) The sensor fabrication procedure includes Steps: 1 (Master mold), 2 (Dropping PDMS mixer on master mold), 3 (Spin-coating of PDMS mixer), 4 (Optical grating sensor after curing), and 5 and 6 (Optical grating sensor after metal deposition). (b) An AFM image of a master mold. (c) An optical microscope image of a PDMS sensor.

III. Result and Discussion

Figure 3(a) shows the structure of the PDMS grating sensor immediately after its detachment from the master mold. Figure 3(b) shows the sensor following the sequential deposition of Cr and Cu layers, which significantly improved the diffraction efficiency of the grating. Cr layer served as an adhesion layer of the subsequent Cu layer to the PDMS layer. Figure 3(d) presents a PDMS grating sensor after being placed on the person's wrist.

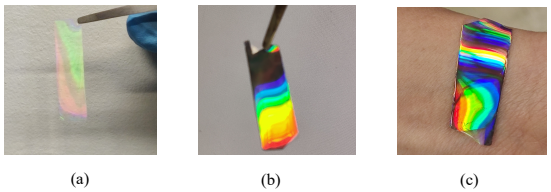


Fig.3. (a) An example of a fabricated sensor before metallization. (b) An example of a fabricated sensor after metallization. (c) An example of a fabricated sensor placed on the radial artery of a person's wrist.

Figure 4 presents a schematic representation of the experimental setup used to measure APWs on a person's wrist. A digital camera with a 10x optical zoom and the capability to record high-speed video at 200 frames per second was connected to a personal computer. The camera was then aligned to focus on the PDMS grating sensor positioned on the person's wrist. A white light source was positioned on the other side of the hand, enabling the observation of APWs with vivid optical reflection, which is illustrated in a supplemental video. This configuration effectively positioned the PDMS grating sensor on the wrist without the need for precise alignment, allowing for accurate measurements of changes in diffraction patterns caused by surface deformation related to blood pressure fluctuations. After recording a video, it was processed for data analysis using ImageJ software to display light intensity changes in multiple spot areas where vivid diffraction changes were observed. While slight movements of the subject may influence the measurement, these movements occur at a much lower frequency than typical APW signals and can be effectively subtracted from the data.

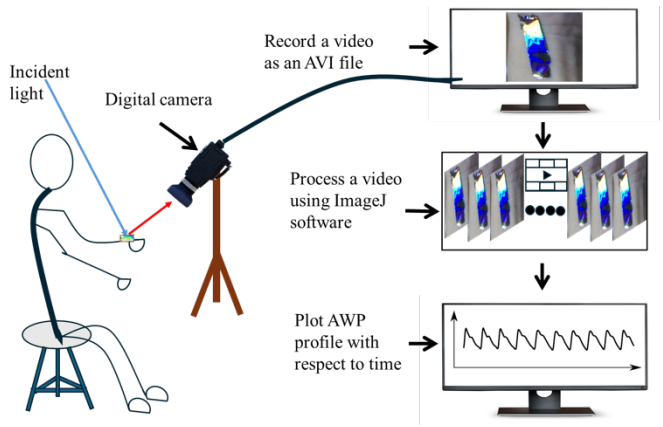


Fig. 4. Schematic diagram of a measurement setup.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:39:50 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2025.3560222

Figures 5(a) and 5(b) illustrate the two distinct wrist configurations under which APWs were monitored. It was observed that the sensor's performance was sensitive to how it was positioned on the skin surface of the wrist. Table 1 presents the measurement parameters for eight experimental sets that we have followed to measure APWs using PDMS grating sensors. These measurements aimed to evaluate the sensor's performance and analyze variations in APWs across different wrist configurations, providing valuable insights into its operational reliability. The APWs response was measured for each wrist configuration 1 and 2, and the recordings were captured for 30 seconds to observe the sensor's response for all four experimental sets. The study varied several key factors, including the sensor's thickness and the skin's surface condition. Wet state represents that the sensor is mounted with a thin layer of ultrasound gel on the skin surface. The dry state represents that the sensor is mounted without any additional layer. The natural elastic properties of PDMS make the sensor to adhere on the skin surface. These factors were tested to evaluate their impact on the detection capability.

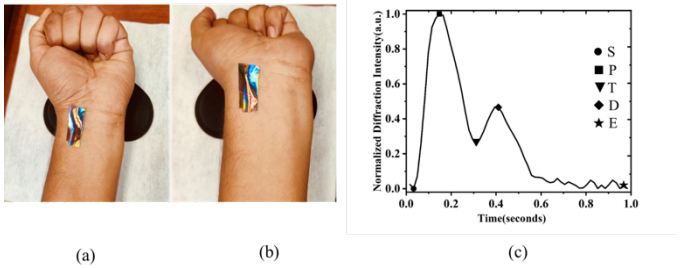


Fig. 5. (a) A sensor positioned on the radial artery indicated as a wrist configuration 1. (b) A sensor positioned on the radial artery indicated as a wrist configuration 2. (c) A measured result of APWs from a thin PDMS grating sensor including five key points: S-systolic phase onset, P-systolic phase peak (systolic pressure), T-dicrotic notch, D-diastolic phase peak, and E-diastolic phase endpoint (diastolic pressure).

Sensor Configuration	Thickness (μm)	Skin of Person's wrist	Wrist configuration
1	Thick 250 μm	Dry	1
2	Thick 250 μm	Dry	2
3	Thick 250 μm	Wet	1
4	Thick 250 μm	Wet	2
5	Thin 100 μm	Dry	1
6	Thin 100 μm	Dry	2
7	Thin 100 μm	Wet	1
8	Thin 100 μm	Wet	2

Table 1. The sensor measurement conditions.

Figure 5(c) shows an example of APW measurement using a thin PDMS grating sensor. The result consists of five key points: S-systolic phase onset, where arterial pressure begins to rise as the heart contracts; P-systolic phase peak (systolic pressure), the maximum pressure during heart contraction; T-dicrotic notch, indicating aortic valve closure and the transition to diastole; D-diastolic phase peak, reflecting stabilized pressure in early diastole; and E-diastolic phase endpoint (diastolic pressure), the lowest arterial pressure during heart relaxation[26]. The measured individual APWs represent similar artery pulse waves with early and late systolic peaks with dicrotic notches observed in other published papers[27], [28], [29], [30].

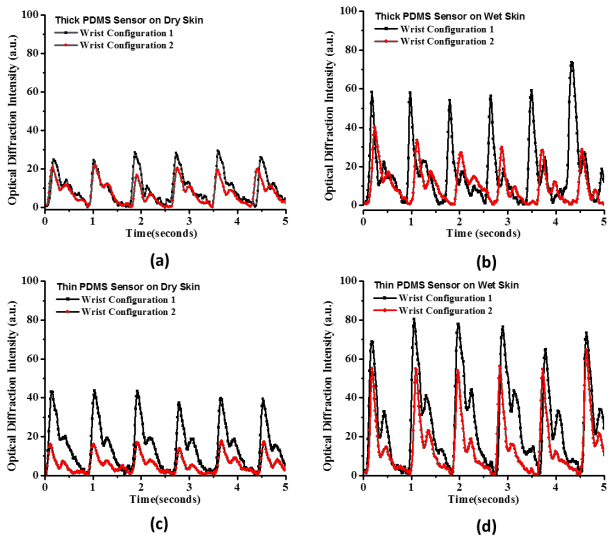


Fig. 6. Measured APWs over a 5-second duration for (a) A thick PDMS sensor placed on dry skin for wrist configuration 1 and 2, (b) A thick PDMS sensor placed on wet skin for wrist configuration 1 and 2, (c) A thin PDMS sensor placed on dry skin for wrist configuration 1 and 2, (d) A thin PDMS sensor placed on wet skin for wrist configuration 1 and 2.

Figure 6 shows the measured responses of APWs over a 5-second duration, using different wrist and sensor configurations. All measurements demonstrate the capability of the sensor to detect APWs successfully, including the five key waveform points. For example, an analysis of the systolic peaks in Figure 6(d) estimates a heart rate of around 67.4 beats per minute, which falls within the normal resting range of adults. However, the sensitivity of the sensor was affected by the measurement parameters. In general, when the wrist is in a stretched position, as seen in configuration 2, the sensor signal is lower. This can be attributed to reduced skin deformation in that configuration, which results in less deformation of the sensor itself.

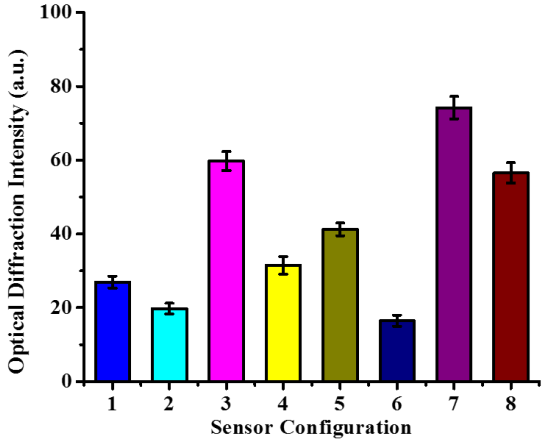


Fig. 7. Summarized measurement results for various sensor configurations.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:39:50 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2025.3560222

Figure 7 shows the summarized diffraction intensity for various sensor configurations. The data were obtained from five different regions of interest (ROIs) that exhibited the most pronounced intensity changes for each measurement configuration. The average peak intensity and its standard deviation from the ROIs were extracted. Overall, thinner sensors showed greater changes in diffraction intensity compared to thicker ones, indicating that thinner sensors better followed skin deformation. Additionally, overall sensitivity improved when a thin layer of ultrasound gel was applied to the skin surface before attaching the sensor. It was also confirmed that the sensor could be used multiple times without significant signal loss when the skin was treated with the gel. However, when the sensor was used repeatedly on dry skin, the adhesive strength weakened, which occasionally affected the signal quality. For instance, under wet skin conditions, the diffraction intensity maintained within a 10 % variation over more than 10 cycles of use, provided that the grating surface of the sensor remained undamaged.

IV. Conclusion

A PDMS-based optical grating sensor was developed for non-invasive monitoring of APWs over the radial artery. The sensor was fabricated using a cost-effective soft lithography technique, incorporating a flexible PDMS elastomer with a metal coating to enhance the reflective diffraction pattern. The elasticity of the PDMS film enables the detection of subtle dynamic skin deformations caused by arterial pressure changes. Covering a relatively broad area over the radial artery, the sensor facilitates easy visualization of dynamic optical intensity changes without requiring precise positioning. Various wrist and sensor configurations have demonstrated the sensor's ability to detect APWs successfully, with optimal configurations identified for maximum sensitivity and repeated usage. Considering its low-cost fabrication process, high sensitivity, and ability to operate without complex demodulation schemes and precise positioning, the proposed sensor has significant potential for widespread adaption in general health monitoring systems for evaluating cardiovascular health.

Acknowledgment

This work was supported by National Science Foundation grant (NSF-2324052).

References

[1] T. Kanda, E. Nakamura, T. Moritani, and Y. Yamori, "Arterial pulse wave velocity and risk factors for peripheral vascular disease," European Journal of Applied Physiology, vol. 82, no. 1–2, pp. 1–7, 2000, doi: 10.1007/s004210050644.

[2] Li B, Gao H, Li X, Liu Y, Wang M. "Correlation between Brachial-Ankle Pulse Wave Velocity and Arterial Compliance and Cardiovascular Risk Factors in Elderly Patients with Arteriosclerosis," Nature Hypertension Res Vol. 29, pp. 309–314, 2006. doi: 10.1291/hypres.29.309.

[3] Qiao, Li-jie & Qi, Zhen & Tu, Li-ping & Zhang, Yu-hang & Zhu, Li-ping & Xu, Jia-tuo & Zhang, Zhi-feng, "The Association of Radial Artery Pulse Wave Variables with the Pulse Wave Velocity and Echocardiographic Parameters in Hypertension," Evidence-Based Complementary and Alternative Medicine, vol. 2018, no. 1, 2018, doi: 10.1155/2018/5291759.

[4] M. Ezzati, Z. Obermeyer, I. Tzoulaki, B. M. Mayosi, P. Elliott, and D. A. Leon, "Contributions of risk factors and medical care to cardiovascular mortality trends," Nature Reviews Cardiology, vol. 12, no. 9, pp. 508–530, 2015, doi: 10.1038/nrcardio.2015.82.

[5] S. Chen, J. Qi, S. Fan, Z. Qiao, J. C. Yeo, and C. T. Lim, "Flexible Wearable Sensors for Cardiovascular Health Monitoring," Advanced Healthcare Materials, vol. 10, no. 17, 2021, doi: 10.1002/adhm.202100116.

[6] Carnethon MR, Pu J, Howard G, et al.., "Cardiovascular Health in African Americans: A Scientific Statement From the American Heart Association," Circulation, vol. 136, no. 21, 2017, doi: 10.1161/CIR.0000000000000534.

[7] P.-C. Hsu, H.-T. Wu, and C.-K. Sun, "Assessment of Subtle Changes in Diabetes-Associated Arteriosclerosis using Photoplethysmographic Pulse Wave from Index Finger," Journal of Medical Systems, vol. 42, no. 3, p. 43, 2018, doi: 10.1007/s10916-018-0901-1.

[8] Knowler, W. C., Fowler, S. E., Hamman, R. F., Christophi, C. A., Hoffman, H. J., Brenneman, A. T., Brown-Friday, J. O., Goldberg, R., Venditti, E., & Nathan, D. M et al., "10-year follow-up of diabetes incidence and weight loss in the Diabetes Prevention Program Outcomes Study," The Lancet, vol. 374, no. 9702, pp. 1677–1686, 2009, doi: 10.1016/S0140-6736(09)61457-4.

[9] M. F. Oliver and L. H. Opie, "Management of acute myocardial infarction," The Lancet, vol. 383, no. 9915, pp. 409–410, 2014, doi: 10.1016/S0140-6736(14)60147-1.

[10] M. Houston, "The role of noninvasive cardiovascular testing, applied clinical nutrition and nutritional supplements in the prevention and treatment of coronary heart disease," Therapeutic Advances in Cardiovascular Disease, vol. 12, no. 3, pp. 85–108, 2018, doi: 10.1177/1753944717743920.

[11] Olsen MH, Angell SY, Asma S, et al., "A call to action and a lifecourse strategy to address the global burden of raised blood pressure on current and future generations: the Lancet Commission on hypertension," The Lancet, vol. 388, no. 10060, pp. 2665–2712, 2016, doi: 10.1016/S0140-6736(16)31134-5.

[12] Wang C, Li X, Hu H, et al., "Monitoring of the central blood pressure waveform via a conformal ultrasonic device," Nature Biomedical Engineering, vol. 2, no. 9, pp. 687–695, 2018, doi: 10.1038/s41551-018-0287-x.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:39:50 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2025.3560222

[13] H. L. Kim, "Arterial stiffness and hypertension," Clinical Hypertension, vol. 29, no. 1., 2023. doi: 10.1186/s40885-023-00258-1.

[14] N. Pilz, V. Heinz, T. Ax, L. Fesseler, A. Patzak, and T. L. Bothe, "Pulse wave velocity: Methodology, clinical applications, and interplay with heart rate variability," Reviews in Cardiovascular Medicine, vol. 25, no. 7., 2024. doi: 10.31083/j.rcm2507266.

[15] R. Gilgen-Ammann, T. Schweizer, and T. Wyss, "RR interval signal quality of a heart rate monitor and an ECG Holter at rest and during exercise," European Journal of Applied Physiology, vol. 119, no. 7, pp. 1525–1532, 2019, doi: 10.1007/s00421-019-04142-5.

[16] Fine J, Branan KL, Rodriguez AJ, et al., "Sources of Inaccuracy in Photoplethysmography for Continuous Cardiovascular Monitoring," Biosensors, vol. 11, no. 4, p. 126, 2021, doi: 10.3390/bios11040126.

[17] P. Pirzada, A. Wilde, and D. Harris-Birtill, "Remote Photoplethysmography for Heart Rate and Blood Oxygenation Measurement: A Review," IEEE Sensors Journal, vol. 24, no. 15, pp. 23436–23453, 2024, doi: 10.1109/JSEN.2024.3405414.

[18] Lv Y, Min L, Niu F, et al., "Conductive Nanomaterials with Different Dimensions for Flexible Piezoresistive Sensors: From Selectivity to Applications," Advanced Materials Technologies, vol. 8, no. 11, 2023, doi: 10.1002/admt.202201886.

[19] H. H. Jang, J. S. Park, and B. Choi, "Flexible piezoresistive pulse sensor using biomimetic PDMS mold replicated negatively from shark skin and PEDOT:PSS thin film," Sensors and Actuators, A: Physical, vol. 286, pp. 107–114, 2019, doi: 10.1016/j.sna.2018.12.015.

[20] J. S. Meena, S. bin Choi, S.-B. Jung, and J.-W. Kim, "Electronic textiles: New age of wearable technology for healthcare and fitness solutions," Materials Today Bio, vol. 19, p. 100565, 2023, doi: 10.1016/j.mtbio.2023.100565.

[21] S. Wang, J. Xiao, H. Liu, and L. Zhang, "Silk nanofibrous iontronic sensors for accurate blood pressure monitoring," Chemical Engineering Journal, vol. 453, p. 139815, 2023, doi: 10.1016/j.cej.2022.139815.

[22] C. Jin, C. Ma, Z. Yang, and H. Lin, "A Force Measurement Method Based on Flexible PDMS Grating," Applied Sciences, vol. 10, no. 7, p. 2296, Mar. 2020, doi: 10.3390/app10072296.

[23] K. Hosokawa, K. Hanada, and R. Maeda, "A polydimethylsiloxane (PDMS) deformable diffraction grating for monitoring of local pressure in microfluidic devices" 200 J. Micromech. Microeng vol. 12 pp. 1, 2001. doi: 10.1088/0960-1317/12/1/301

[24] A. Jou Xie, F. Song, and S.-W. Seo, "An acousto-optic sensor based on resonance grating waveguide structure," Sensors and Actuators A: Physical, vol. 216, pp. 364–369, 2014, doi: 10.1016/j.sna.2014.06.012.

[25] S. C. Truxal, Y.-C. Tung, and K. Kurabayashi, "A PDMS-on-Silicon Deformable Grating for Spectral Differentiation of Mixed Wavelength Signals," in TRANSDUCERS 2007 - International Solid-State Sensors, Actuators and Microsystems Conference, IEEE, 2007, pp. 1087–1090. doi: 10.1109/SENSOR.2007.4300323.

[26] R. Pal, A. Rudas, S. Kim, J. N. Chiang, A. Barney, and M. Cannesson, "An algorithm to detect dicrotic notch in arterial blood pressure and photoplethysmography waveforms using the iterative envelope mean method," Computer Methods and Programs in Biomedicine, vol. 254, 2024, doi: 10.1016/j.cmpb.2024.108283.

[27] C. C. Tyan, S. H. Liu, J. Y. Chen, J. J. Chen, and W. M. Liang, "A novel noninvasive measurement technique for analyzing the pressure pulse waveform of the radial artery," IEEE Transactions on Biomedical Engineering, vol. 55, no. 1, pp. 288–297, 2008, doi: 10.1109/TBME.2007.910681.

[28] S. Pant, S. Umesh, and S. Asokan, "A Novel Approach to Acquire the Arterial Pulse by Finger Plethysmography Using Fiber Bragg Grating Sensor," IEEE Sensors Journal, vol. 20, no. 11, pp. 5921–5928, 2020, doi: 10.1109/JSEN.2020.2973342.

[29] R. B. Gowda, P. Sharan, Saara K, M. Braim, and A. N. Alodhayb, "An <scp>FBG</scp> -based optical pressure sensor for the measurement of radial artery pulse pressure," Journal of Biophotonics, vol. 17, no. 7, 2024, doi: 10.1002/jbio.202400083.

[30] X. Fan et al., "Alignment-Free Liquid-Capsule Pressure Sensor for Cardiovascular Monitoring," Advanced Functional Materials, vol. 28, no. 44, 2018, doi: 10.1002/adfm.201805045.

Mayank Pandey received his B.Sc. degree in Physics from the University of Delhi, Delhi, India, in 2016. He received his M.Sc. degree in Physics from Visvesvaraya National Institute of Technology (VNIT), Nagpur, India, in 2019. He worked as a Teaching Assistant at Delhi Technological University, Delhi, India, from 2020-2021. Currently, he is pursuing his Ph.D. in Electrical Engineering at The City College of New York. His research focuses on micro-actuators and sensors, integrated micro/nano systems, electrochemical sensors, conducting polymers, and micro-fabrication and processing.

Sang-Woo Seo received the B.S. degree in electrical engineering from Ajou University in 1997, the M.S. degree from Gwangju Institute of Science and Technology, South Korea, in 1999, and the Ph.D. degree in electrical engineering from Georgia Institute of Technology, USA, in 2003. From

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:39:50 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2025.3560222

2004 to 2007, he was a Research Engineer with the Group at Georgia Tech and Duke University, to develop thin film photonic devices and integration methods. He then joined as a Research Scientist with the University of California, Davis, where he worked on developing InP-based chip-scale optical code-division multiplex access and arbitrary waveform generator systems. In 2007, he joined the Department of Electrical Engineering, The City College of New York, NY, USA, where he is currently a Professor. His research interests include heterogeneous optoelectronic device integration, integrated microfluidic/photonic sensors, integrated THz systems, bioMEMs, and optical micromechanical systems.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:39:50 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. All rights reserved, including rights for text and data mining and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

AIP Conference Proceedings

RESEARCH ARTICLE | APRIL 07 2025

Diverse machine learning approaches for early prediction of heart attack

Meena Rawat; Dhruv Shakti Singh; Rahul Boadh

AIP Conf. Proc. 3283, 040037 (2025)
https://doi.org/10.1063/5.0265501

View Online Export Citation

Articles You May Be Interested In

Heart disease prediction using ensemble of k-nearest neighbour, random forest and logistic regression method

AIP Conf. Proc. (March 2024)

Digital transformation and machine learning to empower smart healthcare

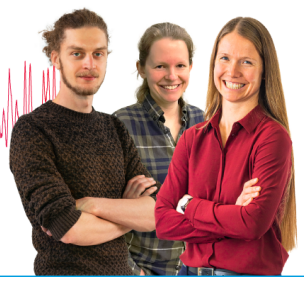
AIP Conf. Proc. (February 2023)

15 April 2025 04:45:48

Webinar From Noise to Knowledge

May 13th – Register now

Zurich Instruments Universität Konstanz



AIP Publishing

Diverse Machine Learning Approaches for Early Prediction of Heart Attack

Meena Rawat[1, a)], Dhruv Shakti Singh[2, b)] and Rahul Boadh[3, c)]

[1]Department of Applied Science, Delhi Technological University, Delhi-110042, India
[2]School of Information & Communication Studies, University College Dublin, Dublin, Ireland
[3]Department of Mathematics, Shyam Lal College, University of Delhi, New Delhi, India

a) rawatmeenasu@gmail.com
b) dhruv.rawat@ucdconnect.ie
c) **Corresponding author:** rboadh.iitkgp@gmail.com

Abstract. A significant evolution has occurred in the realm of medical sciences, attributed to advancements in computing capabilities and refining techniques, particularly in the detection of human heart diseases. Presently, heart diseases rank among the most perilous ailments worldwide, posing severe threats to human life. Prompt and precise identification of these conditions is pivotal in averting heart failure at its nascent stages, thereby enhancing patient survival rates. In addressing these challenges, machine learning algorithms emerge as efficacious and dependable tools for detecting and categorizing individuals afflicted with heart diseases versus those who remain healthy. Leveraging a variety of machine learning algorithms, we embarked on a comprehensive study aimed at identifying and predicting heart diseases, utilizing a heart disease dataset to evaluate performance metrics like sensitivity, specificity, F1-measure, and classification accuracy. Employing five classifiers including Random Forest (RF), Support Vector Machine (SVM), Naïve Bayes (NB), Logistic Regression (LR), and Decision Tree (DT), we subjected the dataset to hyperparameter tuning to enhance classifier efficacy. The proposed decision support system, based on machine learning, is poised to aid doctors in diagnosing heart patients with enhanced efficiency.

15 April 2025 04:45:48

1. INTRODUCTION

A heart attack, medically known as a myocardial infarction, represents a critical cardiovascular event marked by the abrupt disruption of blood flow to a segment of the heart muscle [1]. This interruption primarily stems from the obstruction or constriction of coronary arteries, vital conduits responsible for delivering oxygenated blood to the heart. With heart attacks serving as a primary contributor to mortality on a global scale, their prevalence significantly impacts the burden of disease across [2]. Understanding the epidemiology, risk factors, symptoms, and preventive measures associated with heart attacks is crucial for the effective management and prevention of this life-threatening condition[3].

Heart attacks have a substantial impact on public health, affecting millions of individuals globally each year. As per data from the World Health Organization (WHO), cardiovascular diseases, encompassing heart attacks, contribute to roughly 17.9 million fatalities annually, positioning them as the foremost cause of death globally. Within the United States, figures from the Centers for Disease Control and Prevention (CDC) reveal a staggering occurrence rate, with an individual suffering a heart attack approximately every 40 seconds. This equates to an estimated 805,000 Americans experiencing a heart attack annually. Furthermore, heart attacks contribute to significant healthcare costs, hospitalizations, and long-term disability, placing a considerable economic burden on healthcare systems and societies.

International Conference on Advances in Pure & Applied Mathematics (ICAPAM)
AIP Conf. Proc. 3283, 040037-1–040037-12; https://doi.org/10.1063/5.0265501
Published under an exclusive license by AIP Publishing. 978-0-7354-5161-2/$30.00

Numerous risk factors play pivotal roles in the onset of heart attacks, comprising both modifiable and non-modifiable elements. Modifiable factors encompass lifestyle choices and behaviors that individuals can alter to reduce their risk, such as smoking, unhealthy diet, physical inactivity, and excessive alcohol consumption [4]. Non-modifiable factors, on the other hand, are inherent characteristics or conditions that individuals cannot change, including age, gender, family history of heart disease, and genetic predispositions. Understanding and addressing these risk factors are crucial steps in mitigating the likelihood of experiencing a heart attack. Men face a greater risk of heart attacks than women, although women's risk escalates following menopause [5]. Additionally, individuals with a family history of heart disease are more susceptible to experiencing heart attacks at a younger age. Modifiable risk factors encompass lifestyle behaviors and underlying medical conditions[6]. These include smoking, unhealthy diet, physical inactivity, obesity, hypertension, high cholesterol, diabetes, and chronic stress. Addressing these risk factors through lifestyle modifications and medical management is crucial for reducing the risk of heart attacks and improving overall cardiovascular health.

Heart attacks present with a variety of symptoms, ranging from mild to severe, and may differ among individuals. The predominant symptom is chest pain or discomfort, frequently characterized as pressure, tightness, squeezing, or aching in the chest [7]. Other symptoms may include shortness of breath, nausea, vomiting, indigestion, sweating, lightheadedness, palpitations, and pain or discomfort in the arms, back, neck, jaw, or stomach[8]. However, some people, particularly women, older adults, and individuals with diabetes, may experience atypical symptoms or no symptoms at all, known as silent heart attacks. Early recognition of symptoms and prompt medical intervention are crucial for preventing complications and improving outcomes[9].

Recently, Machine learning (ML) has surfaced as a promising tool. for early prediction and diagnosis of heart attacks[10]. ML algorithms possess the capability to scrutinize extensive sets of patient data, incorporating demographic details, medical records, laboratory findings, imaging results, and lifestyle variables. This comprehensive analysis enables the identification of patterns and facilitates the prediction of heart attack probabilities [11]. By leveraging predictive modeling techniques, ML algorithms can assist healthcare providers in identifying high-risk individuals, enabling targeted interventions and preventive measures to reduce the incidence of heart attacks[12].

Furthermore, ML-based risk prediction tools can be integrated into electronic health records (EHRs) and clinical decision support systems, enabling healthcare providers to assess patients' cardiovascular risk more efficiently and tailor preventive strategies accordingly[13]. By harnessing the power of ML, healthcare systems can enhance risk stratification, optimize resource allocation, and improve patient outcomes in the management of heart attacks and other cardiovascular diseases. ANFIS, FIS and mathematical modeling have been used for different type of work in health and other industries in previous studies [14-28].

2. MATERIAL AND METHODS

The subsequent subsections provide a succinct overview of the research materials and methodologies employed in this study.

15 April 2025 04:45:48

2.1 Dataset: The dataset utilized in this study originates from 1988 and comprises four distinct databases: Cleveland, Hungary, Switzerland, and Long Beach V[29].

TABLE 1: Description of Heart Diseases dataset

SN	Feature Name	Type	Description	Range of values
1	Age	Continuous	Age in Years	$29 \leq Age \leq 77$
2	Sex	Discrete	Male and Female	Male = 1 Female = 0
3	Chest Pain Type	Discrete	Atypical Angina Typical Angina Asymptomatic Nonanginal Pain	Atypical Angina = 1 Typical Angina = Asymptomatic = 3 Nonanginal pain = 4
4	Resting Blood Pressure	Continuous	In mm Hg admitted at the hospital	94mm – 200mm
5	Serum Cholesterol	Continuous	In mg/dl	120 mg/dl -564 mg/dl
6	Fasting Blood Sugar	Discrete	Fasting Blood Sugar $>$ 120 or not	True = 1 False = 0
7	Resting Electrocardiographic Results	Discrete	normal or having ST-T or hypertrophy	Normal = 0; having ST-T = 1; hypertrophy = 2
8	Maximum Heart Rate Achieved	Continuous	-	71-202
9	Exercise Induced Angina	Discrete	Yes or No	Yes = 1; No = 0
10	ST depression induced by exercise relative to rest	Continuous	-	0 – 6.2
11	Slope of the Peak Exercise ST Segment	Discrete	For upslope, Stable and downslope	For upslope = 1; Stable = 2 and downslope = 3
12	Number of Major Vessels (0-3) Colored by Fluoroscopy	Discrete	-	Four classes 0,1,2,3
13	Thallium Scan	Discrete	Normal, fixed defect and reversible defect	Normal = 3; fixed defect = 6 and reversible defect =7
14	Heart Diseases	Discrete	Binary	True = 1, False = 0

Comprising a total of 1025 records, each entry encompasses 76 attributes, inclusive of the predicted attribute. However, the emphasis of this work is on using a subset of the dataset that includes 14 features. The target output label represents two classes, distinguishing between heart patients and normal subjects. Consequently, the extracted dataset forms a 1025 × 13 features matrix. A detailed description of the 1025 instances, each with 13 features, is provided in Table 1.

2.2 Proposed Methodology: The proposed system has been designed with the primary objective of accurately classifying individuals into two categories: those afflicted with heart disease and those who are deemed healthy. The heart disease diagnosis system incorporated a range of machine learning classifiers including LR, K-NN, SVM, RF, and NB. Following this, the model underwent

15 April 2025 04:45:48

rigorous validation and performance assessment using established metrics. The methodology of the proposed system was methodically structured into five discernible stages: (1) preprocessing the dataset, (2) dividing the dataset into training and testing subsets, (3) employing cross-validation techniques, (4) utilizing machine learning classifiers, and (5) assessing classifier performance. The comprehensive approach adopted for heart disease diagnosis is encapsulated in the framework depicted in Fig. 1, signifying a structured and thorough process.

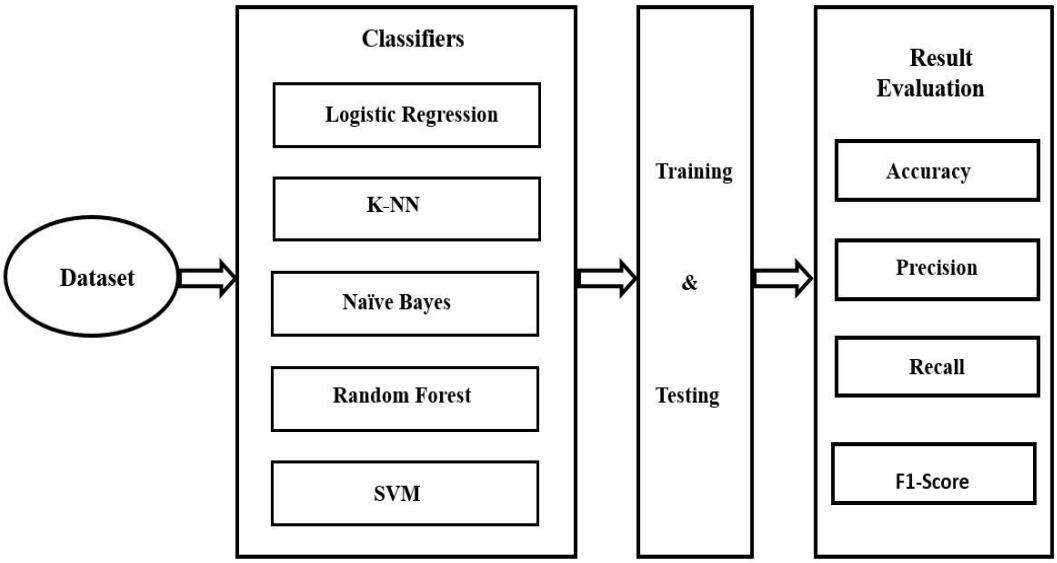


FIG.1: Proposed Methodology for the diagnosis of Heart-Diseases

2.2.1 Data preprocessing: Data preprocessing plays a crucial role in ensuring the efficient representation of data and facilitating the effective training and testing of ML classifiers. To achieve this, various preprocessing techniques, including the removal of missing values, standard scaling, and Min-Max scaling, were applied to the dataset [30]. The standard scalar technique was employed to standardize features by ensuring that each feature has a mean of 0 and variance of 1, thereby bringing all features to the same scale. Similarly, Min-Max scaling was utilized to shift the data so that all features fall within the range of 0 and 1. While missing values were observed to be absent in the dataset used for predicting heart diseases, any rows containing missing values were simply deleted from the dataset. Apart from this, all other steps of data preprocessing techniques were meticulously executed in this research endeavor.

2.2.2 Machine Learning Classifiers: ML Classifiers play a pivotal role in categorizing individuals into heart patients and healthy individuals. These classifiers leverage various classification algorithms to make predictions based on input data[31]. In this paper, several popular classification algorithms and their theoretical underpinnings are succinctly elucidated.

Logistic Regression: LR operates by modeling the relationship between one or more predictor variables and a binary outcome[32]. It employs a logistic function to depict the probability of the binary outcome. This function, commonly referred to as the sigmoid function, is defined as:

15 April 2025 04:45:48

$$h_\theta(x) = \frac{1}{1 + e^{-\theta^t x}} ;$$

Where $h_\theta(x)$ is the predicted probability of the positive class for input x,
θ is the vector of coefficients (parameters) to be estimated,
x is the vector of predictor variables,
$\theta^t x$ is the dot product of θ and x
e is the base of the natural logarithm.
During training, LR aims to estimate the parameters θ that best fit the data. This is achieved by maximizing the likelihood function, which measures the likelihood of the observed data given the parameters θ. Mathematically, it can be expressed as:

$$l(\theta) = \prod_{i=1}^{n} h_\theta\left(x^{(i)}\right)^{y^{(i)}} (1 - h_\theta(x^{(i)}))^{1-y^{(i)}}$$

Where $l(\theta)$ is the likelihood function,
$x^{(i)}$ is the ith observation (input data),
$y^{(i)}$ is the ith (binary outcome),
$h_\theta(x^{(i)})$ is the predicted probability for the ith observation.
The parameters θ are estimated by minimizing the negative log-likelihood function (or maximizing the log-likelihood function). This optimization task is commonly tackled through iterative optimization algorithms like gradient descent or Newton'os method. Once the parameters θ are estimated, Logistic Regression can make predictions on new data by applying the logistic function to calculate the predicted probabilities. A decision threshold (usually 0.5) is then applied to these probabilities to classify the observations into the positive or negative class.
Finally, the performance of the LR model is assessed through diverse metrics to assess its effectiveness in binary classification tasks, providing interpretable results and insights into the relationship between predictor variables and outcomes.

Random Forest: A RF comprises a multitude of decision trees, each trained on a random subset of both data and features. Predictions from the Random Forest are derived by consolidating the forecasts of individual trees, usually employing a majority voting mechanism for classification purposes [33]. So, the classifier employs an ensemble of decision trees to make predictions, utilizing the following mathematical equation:

$$\hat{Y} = \frac{1}{N} \left(\sum_{i=1}^{N} T_i(X)\right);$$

Where $\hat{Y}$ represents the predicted class for a given input X,
N represents the number of decision trees in the RF ensemble,
$T_i(X)$ stands for the prediction of the ith decision tree on input X.
Each decision tree is constructed based on a random sampling of the data and features, and its prediction is obtained by the majority vote of the nodes within the tree. The ultimate forecast of the Random Forest is achieved by consolidating the predictions made by all decision trees within the ensemble. By employing this ensemble methodology, the Random Forest classifier attains

15 April 2025 04:45:48

resilience and precision in its prognostications, thereby rendering it appropriate for a diverse range of classification undertakings.

Support Vector Machine: In order to group data points into discrete groups and maximize the margin between them while minimizing classification errors, the SVM classifier determines the best hyperplane [34]. It achieves this by finding the decision boundary that best separates the support vectors of each class.

$$f(x) = sign\left(\sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + b\right)$$

Where $f(x)$ represents the decision function for classifying new data point x,
α_i are the Lagrange multipliers determined during the process of training,
y_i is the class label of the training data point x_i
$K(x_i, x)$ is the kernel function, which measures the resemblance between x_i and x,
B is the bias term.
In linear SVM, the kernel function $K(x_i, x)$ simplifies to the dot product of the input vectors x_i and x. However, for non-linear SVM, diverse kernel functions like polynomial or radial basis functions are utilized to transform the input data into a higher-dimensional space, enabling a linear hyperplane to segregate the classes. Throughout training, SVM endeavors to maximize the margin, defined as the distance between the hyperplane and the nearest data points (support vectors) of each class. This optimization task typically involves solving a convex quadratic programming problem using optimization techniques. Once the optimal hyperplane is established, SVM can classify new data points by assessing the decision function f(x), assigning them to the class based on the function output's sign. Through this mathematical framework, SVM achieves efficient binary classification by identifying the hyperplane that effectively separates the data points of distinct classes while maximizing the margin.

K-Nearest Neighbor: For tasks involving regression as well as classification, the non-parametric k-NN classifier is employed [35]. To function, a data point is allocated to the majority class among its k nearest neighbors in the feature space. Mathematically, the process can be represented as follows:

$$\hat{y}(x) = mode\ (y_{i1}, y_{i2}, y_{i3}, \ldots, y_{ik})$$

where $\hat{y}(x)$ denotes the predicted class for the data point.
$(y_{i1}, y_{i2}, y_{i3}, \ldots, y_{ik})$ Represents the class labels of the k nearest neighbors of x.
Mode calculates the most frequent class label among the k neighbors.
In simple terms, k-NN determines the class of a data point by considering the classes of its nearest neighbors. The value of k is a hyperparameter that needs to be specified before training. Larger values of k result in smoother decision boundaries but may lead to misclassification of data points near the boundary. Conversely, smaller values of k result in more complex decision boundaries but may be sensitive to noise in the data.

15 April 2025 04:45:48

Gaussian Naive Bayes: Based on the Bayes theorem, the NB classifier asserts that characteristics, given the class label, are conditionally independent [36]. Mathematically, the classification decision for a new data point x can be expressed as:

$$\hat{y} = \underbrace{\arg\max P(y \mid x)}_{y} = \underbrace{\arg max}_{y} \frac{P(x|y)\,P(y)}{P(x)}$$

In this equation $\hat{y}$ represents the predicted class for the new data point x.
P (y | x) represents the posterior probability of class y given the features x.
P (x | y) represents the likelihood of observing features x given class y.
P(y) represents the prior probability of class y.
P(x) represents the probability of observing features x (often considered as a constant for all classes and thus ignored in the classification decision).

2.2.3 Training and Testing: The dataset is partitioned into training and testing data sets. The training set comprises instances with known outputs, which the model utilizes to learn and later generalize its predictions to unseen data. In this study, training is conducted using 70% of the dataset. Accordingly, 717 cases out of 1025 instances in the dataset are selected as the training set, while the remaining 308 instances are allocated for testing purposes, representing 30% of the total data. To optimize model performance, a technique known as parameter tuning is employed, ensuring that each model operates at its highest potential.

3. RESULT AND DISCUSSION

In this study, we implemented the k-fold cross-validation (CV) method alongside four distinct performance evaluation metrics. The k-fold cross-validation technique involves dividing the dataset into k equal-sized segments, where k - 1 groups serve as training data for the classifiers, and the remaining portion is reserved for performance evaluation in each iteration. This validation procedure is iterated k times, with the classifier's performance gauged based on the outcomes of these k iterations. Various values of k can be chosen for CV, and for our experiment, we selected k = 10 due to its favorable performance. Throughout the 10-fold CV process, 90% of the data were utilized for training, while the remaining 10% were earmarked for testing. This iterative process was repeated 10 times for each fold, with instances randomly allocated to training and testing groups before each cycle. Subsequently, following the 10-fold process, the averages of all performance metrics were computed to comprehensively evaluate the classifier's overall performance.

3.1 Performance Evaluated by the Classifiers: Table 2 presents the performance metrics of different classifiers. The Random Forest classifier achieves an accuracy of 98.05%, with perfect precision and a high recall of 95.97%, resulting in an F1-score of 97.94%.
Similarly, the K-NN classifier demonstrates superior performance, with an accuracy of 99.02%, perfect precision, a high recall of 97.99%, and an impressive F1-score of 98.98%. In contrast, both Logistic Regression and Gaussian classifiers exhibit lower accuracy, precision, and recall values. Logistic Regression achieves an accuracy of 81.16%, with a precision of 77.24%, recall of 86.57%, and an F1-score of 81.64%. The Gaussian classifier shows comparable results, with an accuracy of 81.49%, precision of 76.44%, recall of 89.26%, and an F1-score of 82.35%.

15 April 2025 04:45:48

TABLE 2: presents the performance metrics of different classifiers

	Accuracy	Precision	Recall	F1-Score
Random Forest	**98.05**	**100**	**95.97**	**97.94**
K-NN	**99.02**	**100**	**97.99**	**98.98**
Logistic Regression	**81.16**	**77.24**	**86.57**	**81.64**
Gaussian Naïve Bayes	**81.49**	**76.44**	**89.26**	**82.35**
SVM	**97.08**	**100**	**93.96**	**96.89**

While comparing the results with the statistical graph that are shown in Fig. 2, we observe the following points:

1. Accuracy and Precision: Random Forest and K-NN classifiers demonstrate superior accuracy and precision, achieving near-perfect scores of over 98%. SVM also shows high accuracy and perfect precision, while Logistic Regression and Gaussian classifiers exhibit lower scores, indicating less reliable classification.
2. Recall: K-NN classifier achieves the highest recall, indicating its ability to effectively capture positive instances, closely followed by Random Forest. While SVM also performs well in recall, Logistic Regression and Gaussian classifiers have comparatively lower recall rates, suggesting they may miss some positive instances.
3. F1-Score: K-NN classifier achieves the highest F1-score, reflecting a balanced performance in terms of precision and recall. Random Forest also demonstrates a strong F1-score, indicating a good balance between precision and recall. On the other hand, Logistic Regression and Gaussian classifiers exhibit lower F1-scores, suggesting a trade-off between precision and recall.
4. Robustness: Random Forest and K-NN classifiers appear to be more robust, as they consistently perform well across all metrics. SVM also shows robust performance, particularly in accuracy and precision. However, Logistic Regression and Gaussian classifiers display more variability in performance, with lower scores in some metrics.
5. Model Complexity: Random Forest and K-NN classifiers are relatively simple models yet achieve high performance, making them suitable choices for various classification tasks. SVM, while more complex, also demonstrates strong performance. In contrast, Logistic Regression and Gaussian classifiers are simpler models but show comparatively weaker performance, indicating limitations in capturing complex relationships in the data.

15 April 2025 04:45:48

4. CONCLUSION

Heart disease remains a significant public health challenge, contributing to substantial morbidity, mortality, and healthcare costs worldwide. Understanding the epidemiology, risk factors, symptoms, and preventive measures associated with heart disease is essential for reducing the burden of this disease and improving cardiovascular health. Machine learning offers exciting opportunities for early prediction and diagnosis of heart attacks, empowering healthcare providers to identify high-risk individuals and implement targeted interventions to prevent adverse cardiovascular events. By leveraging ML algorithms and predictive modeling techniques, healthcare systems can enhance risk stratification, optimize resource allocation, and improve patient outcomes in the management of heart attacks and other cardiovascular diseases. Through continued research, innovation, and collaboration, ML-based approaches have the potential to revolutionize the prevention, diagnosis, and treatment of heart attacks, ultimately saving lives and improving the quality of care for individuals at risk.

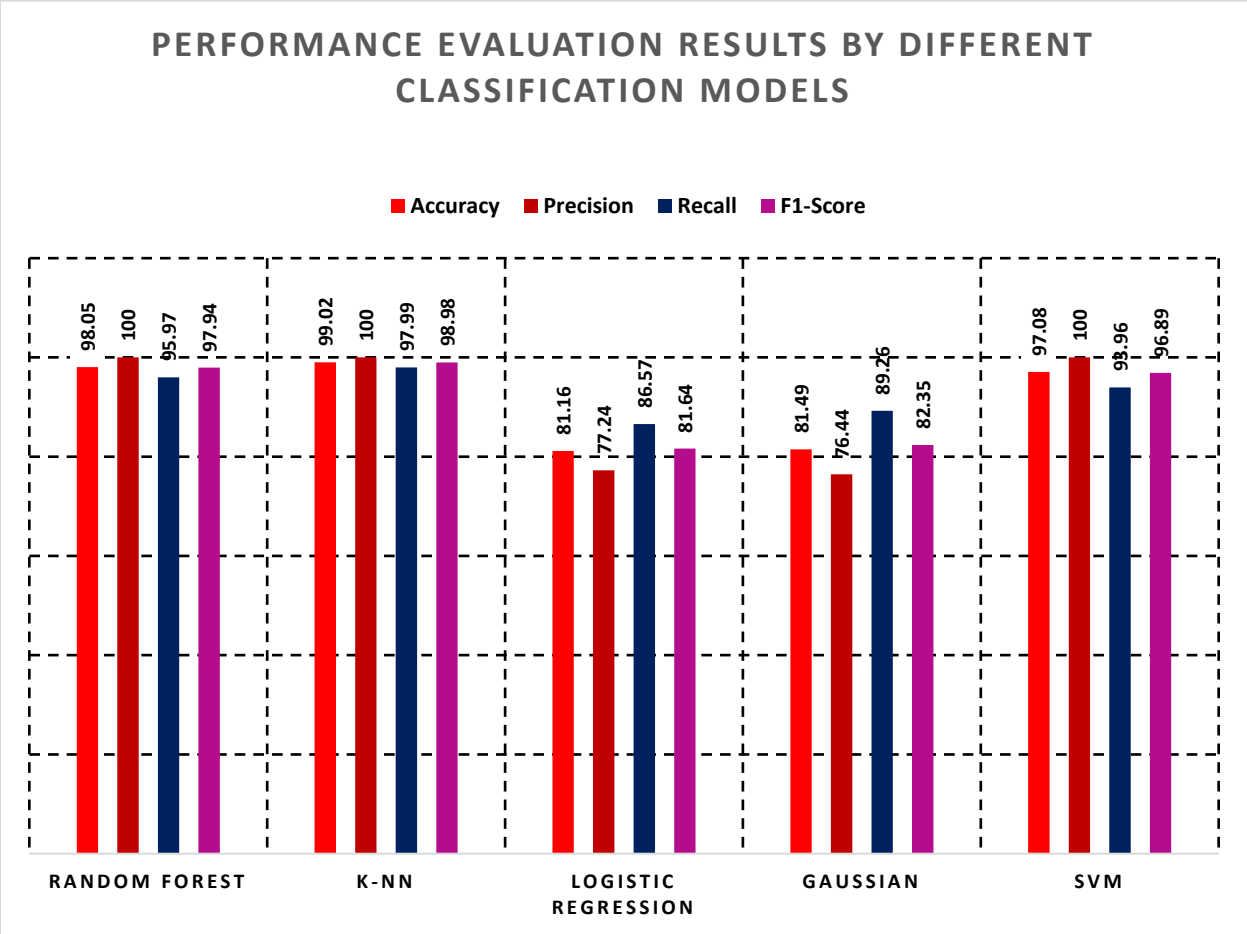


FIG.2: Performance Evaluation Results by different classification Models

The comprehensive evaluation of various classification algorithms provides invaluable insights into the diagnosis of heart diseases. With high accuracy, precision, recall, and F1-scores, the Random Forest and K-NN classifiers emerge as robust contenders, showcasing their efficacy in accurately identifying individuals with heart conditions. Their remarkable performance

15 April 2025 04:45:48

underscores their potential as reliable tools in clinical settings for early detection and intervention. Furthermore, the solid performance of the SVM classifier reaffirms its utility, particularly in scenarios where precision and accuracy are paramount. However, the comparatively lower scores of the Logistic Regression and Gaussian classifiers suggest potential limitations in their ability to discern subtle patterns within the data. Nonetheless, these findings collectively emphasize the significance of leveraging machine learning algorithms for enhancing the diagnostic process of heart diseases, ultimately contributing to improved patient outcomes and healthcare management.

REFERENCES

[1] Md. I. Hossain *et al.*, "Heart disease prediction using distinct artificial intelligence techniques: performance analysis and comparison," *Iran J Comput Sci*, vol. 6, no. 4, pp. 397–417, Dec. 2023, doi: 10.1007/s42044-023-00148-7.

[2] A. U. Haq, J. P. Li, M. H. Memon, S. Nazir, and R. Sun, "A Hybrid Intelligent System Framework for the Prediction of Heart Disease Using Machine Learning Algorithms," *Mobile Information Systems*, vol. 2018, p. e3860146, Dec. 2018, doi: 10.1155/2018/3860146.

[3] L. Yahaya, N. D. Oye, and E. J. Garba, "A comprehensive review on heart disease prediction using data mining and machine learning techniques," *American Journal of Artificial Intelligence*, vol. 4, no. 1, pp. 20–29, 2020.

[4] P. E. Rubini, C. A. Subasini, A. V. Katharine, V. Kumaresan, S. G. Kumar, and T. M. Nithya, "A cardiovascular disease prediction using machine learning algorithms," *Annals of the Romanian Society for Cell Biology*, pp. 904–912, 2021.

[5] L. Chandrika and K. Madhavi, "A hybrid framework for heart disease prediction using machine learning algorithms," in *E3S Web of Conferences*, EDP Sciences, 2021, p. 01043. Accessed: Jan. 31, 2024. [Online]. Available: https://www.e3s-conferences.org/articles/e3sconf/abs/2021/85/e3sconf_icmed2021_01043/e3sconf_icmed2021_01043.html

[6] M. P. Behera, A. Sarangi, D. Mishra, and S. K. Sarangi, "A Hybrid Machine Learning algorithm for Heart and Liver Disease Prediction Using Modified Particle Swarm Optimization with Support Vector Machine," *Procedia Computer Science*, vol. 218, pp. 818–827, Jan. 2023, doi: 10.1016/j.procs.2023.01.062.

[7] Y. J. Chauhan, "Cardiovascular disease prediction using classification algorithms of machine learning," *International Journal of Science and Research*, vol. 9, no. 5, pp. 194–200, 2020.

[8] S. N. Pasha, D. Ramesh, S. Mohmmad, A. Harshavardhan, and Shabana, "Cardiovascular disease prediction using deep learning techniques," *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 981, no. 2, p. 022006, Dec. 2020, doi: 10.1088/1757-899X/981/2/022006.

[9] H. Jindal, S. Agrawal, R. Khera, R. Jain, and P. Nagrath, "Heart disease prediction using machine learning algorithms," in *IOP conference series: materials science and engineering*, IOP Publishing, 2021, p. 012072. Accessed: Jan. 31, 2024. [Online]. Available: https://iopscience.iop.org/article/10.1088/1757-899X/1022/1/012072/meta

[10] S. Mohan, C. Thirumalai, and G. Srivastava, "Effective heart disease prediction using hybrid machine learning techniques," *IEEE access*, vol. 7, pp. 81542–81554, 2019.

[11] P. Anbuselvan, "Heart disease prediction using machine learning techniques," *Int. J. Eng. Res. Technol*, vol. 9, pp. 515–518, 2020.

[12] V. Sharma, S. Yadav, and M. Gupta, "Heart disease prediction using machine learning techniques," in *2020 2nd international conference on advances in computing, communication*

15 April 2025 04:45:48

control and networking (ICACCCN), IEEE, 2020, pp. 177–181. Accessed: Jan. 31, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9362842/

[13] D. Shah, S. Patel, and S. K. Bharti, "Heart Disease Prediction using Machine Learning Techniques," *SN COMPUT. SCI.*, vol. 1, no. 6, p. 345, Nov. 2020, doi: 10.1007/s42979-020-00365-y.

[14] Kalra, G., Kishor,K., Kumar, A., Rajoria, Y. K., Yadav, A., and Boadh, R. (2022a). FIS-Based Prediction and Estimation of Health Insurance for Workers in the Manufacturing Sector, Journal of Pharmaceutical Negative Results. Vol. 13 (S8), pp. 2236-2245. DOI: 10.47750/pnr.2022.13.S08.33118.

[15] Bhagat, R. K., Yadav, A., Rajoria, Y. K., Raj, S. and Boadh, (2022). Study of Fuzzy and Artificial Neural Network (ANN) BasedTechniques to Diagnose Heart Disease, Journal of Pharmaceutical Negative Results. Vol. 13 (S5), pp. 1023-1029. Doi: 10.47750/pnr.2022.13. S05.161.19.

[16] Poswal, P., Chauhan, A., Aarya, D. D., Boadh, R., Rajoria, Y. K. and Gaiola, S. U. (2022a). Optimal Strategy for Remanufacturing System of Sustainable Products with Trade credit under Uncertain Scenario, Materials Today: Proceedings. https://doi.org/10.1016/j.matpr.2022.08.303.20.

[17] Boadh, R., Yadav, S. N., Tiwari, A., Rajoria, Y. K. and Singh, J. (2022a). Application of fuzzy inference system (FIS) for assessment and predication of compressive asset of concrete containing fly ash, Materials Today: Proceedings. https://doi.org/10.1016/j.matpr.2022.08.160.21.

[18] Poswal P, Chauhan A, Boadh R and Rajoria Y.K (2022b). A review on fuzzy economic order quantity model under shortage, AIP Proceeding, 2481, 040023(1-13) 10.1063/5.0103757.22.

[19] Poswal, P., Chauhan, A., Rajoria, Y. K., Boadh, R. and Goel, A. (2022d). Fuzzy optimization model of two parameter weibull deteriorating rate with quadratic demand and variable holding cost under allowable shortages, Yugoslav Journal of Operations Research, Vol.32 (4).25.

[20] Boadh, R. Aarya, D. D., Dahiya, M., Rathee, S., Kumar, A., Jain, S., Rajoria, Y. R., and Rani, S. (2022b). Study and prediction of prostate cancer using fuzzy inference system, Materials Today: Proceedings, vol. 56 (1), pp 157-164. doi:https://doi.org/10.1016/j.matpr.2022.01.040.26.

[21] Boadh, R. Grover,R., Dahiya, M., Kumar, A., Rathee, R., Rajoria, Y. R., Rawat, M., and Rani, S. (2022c) Study of fuzzy expert system for the diagnosis of various types of cancer, Materials Today: Proceedings, vol. 56 (1), pp 298-307. doi: https://doi.org/10.1016/j.matpr.2022.01.16127.

[22] Boadh, R., Chaudhary, K., Dahiya, M., Dogra, N., Rathee, S., Kumar, A., Rajoria, Y. K. (2022d). Analysis and investigation of fuzzy expert system for predicting the child anaemia, Materials Today: Proceedings, vol. 56 (1), pp 231-236. doi:https://doi.org/10.1016/j.matpr.2022.01.094.28.

[23] Karla, G., Rajoria, Y. K.. Boadh, R., Rajendra, P., Khatak, N., and Kumar, A. (2022c). Study offuzzy expert systems towards prediction and detection of fraud case in health care insurance, Materials Today: Proceedings, vol. 56 (1), pp 477-480. doi: https://doi.org/10.1016/j.matpr.2022.02.157.29.

[24] Poswal, P., Chouhan, A., Boadh, R., Rajoria, Y. K., Kumar, A. and Khatak, N. (2022e). Investigation and analysis of fuzzy EOQ model for price sensitive and stock dependent

15 April 2025 04:45:48

demand under shortages, Materials Today: Proceedings, vol. 56 (1), pp 542-548 doi:https://doi.org/10.1016/j.matpr.2022.02.273.30.
[25] Aarya, D., Rajoria, Y. K., Gupta, N., Singh.Y, & Rathee, R, Boadh, R. and Kumar, A. (2022). Selling price, time dependent demand and variable holding cost inventory model with two storage facilities. Materials Today: Proceedings. 56(1), pp. 245-251, doi: 10.1016/j.matpr.2022.01.11138.
[26] Poswal, P., Chauhan, A., Rajoria, Y.K., Boadh, R., Goel, A. (2022). Fuzzy optimization model of two parameter weibull deteriorating rate with quadratic demand and variable holding cost under allowable shortages, Yugoslav Journal of Operations Research, 32(4), pp. 453-479. DOI: 10.2298/YJOR220115021P44.
[27] Tiwari, A., Rajoria, Y. K,, Boadh, R. (2022). Comparative Study of Maximization Assignment Model by Existing Method and Newly Proposed Methods. International journal of Operational Research, (in press). DOI: 10.1504/IJOR.2022.1004664445.
[28] Kaur, K., Jai Shankar Prasad, Deo Datta Aarya, Hare Krishna, Vishal B. Tambe, Yogendra Kumar Rajoria, Rahul Boadh. (2022). Implementation of an Adaptive Artificial Neural Network with Fuzzy Expert System for Diagnoses the Breast and Prostate Cancer: A Hybrid Technique. Journal of Pharmaceutical Negative Results, Vol. 13 (08), pp. 3806–3812. https://doi.org/10.47750/pnr.2022.13.S08.47
[29] "Heart Disease Dataset." Accessed: Feb. 28, 2024. [Online]. Available: https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset
[30] E. Eessaar, "The database normalization theory and the theory of normalized systems: finding a common ground," *Baltic Journal of Modern Computing*, vol. 4, no. 1, p. 5, 2016.
[31] "Study of machine learning algorithms for special disease prediction using principal of component analysis | IEEE Conference Publication | IEEE Xplore." Accessed: Jan. 31, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7955260
[32] I. H. Sarker, "Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions," *SN COMPUT. SCI.*, vol. 2, no. 6, p. 420, Nov. 2021, doi: 10.1007/s42979-021-00815-1.
[33] B. H. Menze *et al.*, "A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data," *BMC bioinformatics*, vol. 10, no. 1, pp. 1–16, 2009.
[34] Dr. H. Fraihat, "A Comparison of ANFIS, SVM and KNN models for the face recognition in unconstrained environment," *IJATCSE*, vol. 9, no. 4, pp. 6518–6526, Aug. 2020, doi: 10.30534/ijatcse/2020/339942020.
[35] Y. Kumar, S. Gupta, R. Singla, and Y.-C. Hu, "A Systematic Review of Artificial Intelligence Techniques in Cancer Prediction and Diagnosis," *Arch Computat Methods Eng*, vol. 29, no. 4, pp. 2043–2070, Jun. 2022, doi: 10.1007/s11831-021-09648-w.
[36] M. Marimuthu, M. Abinaya, K. S. Hariesh, K. Madhankumar, and V. Pavithra, "A review on heart disease prediction using machine learning and data analytics approach," *International Journal of Computer Applications*, vol. 181, no. 18, pp. 20–25, 2018.

15 April 2025 04:45:48

The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/1463-5771.htm

Benchmarking:
An International
Journal

Does ESG impact corporate performance in environmentally sensitive and non-sensitive industries? A dynamic panel approach

Prachi Jain and Priya Malhotra
University School of Management and Entrepreneurship (USME), Delhi Technological University, Delhi, India

Received 7 July 2024
Revised 13 January 2025
26 February 2025
Accepted 28 February 2025

Abstract

Purpose – This study explores how aggregate ESG disclosure, along with its three individual dimensions – environmental, social and corporate governance – affects corporate performance by categorizing corporations into environmental sensitivity industries (SI), characterized by business activities with significant socioenvironmental landscapes and more likely to damage social and environmental factors and non-environmental sensitivity industries (NSI) in India.

Design/methodology/approach – This research employs a system-generalized method of moments (GMM) technique to analyze 605 Indian firms listed on the National Stock Exchange (NSE) from 2015 to 2023. The methodology ensures robust and endogeneity-resistant results, reliably assessing the relationship between the variables.

Findings – The study indicates that ESG factors influence corporate outcomes differently across industries, and simply following sustainability – is somewhat naïve. In SI industries, ESG efforts show an insignificant association and implementing effective systems for corporate governance reduces profitability due to challenges and scrutiny from stakeholders. Conversely, NSI corporations benefit significantly from ESG practices, enhancing long-term value and investor appeal.

Originality/value – The research offers novelty and contributes to the existing literature by providing robust empirical evidence on the diverse impact of ESG (and its dimensions). It underscores the necessity for concerned companies, especially in emerging economies, to customize sustainable strategies that align with their unique operational contexts. The findings provide valuable insights for managers seeking to effectively manage ESG implementation and its financial effects, enhancing strategic approaches to sustainability in different industry contexts.

Keywords ESG disclosure, Environmental disclosure, Social disclosure, Corporate governance disclosure, Corporate performance, Environmental sensitivity, System-generalized method of moments (GMM), Indian firms

Paper type Research paper

1. Introduction

In recent years, sustainability has become integral to corporate strategy, emphasizing long-term viability and financial market resilience (Aras *et al.*, 2018). This shift stems from corporations' concerns about economic stability, investor confidence, and competitive advantage. Thus, sustainability has emerged as a critical factor for organizational success and growth, focusing on waste management, carbon footprint reduction, and operational efficiency (Eccles and Serafeim, 2013). This transition is fueled by growing stakeholder awareness (Maji and Lohia, 2023), prompting firms to consider environmental and social responsibility alongside profit maximization (Arianpoor *et al.*, 2023; Khandelwal *et al.*, 2023). As social and climate change gain prominence (Ge *et al.*, 2016; Quang *et al.*, 2022), they converge environmental stewardship, social responsibility, and corporate governance

JEL Classification — M14, L25, Q01

Conflict of interest: The authors declare that they have no financial or non-financial conflicts of interest to disclose regarding this research article.

Funding: This study has not received any external funding or financial support.

Benchmarking: An International Journal
© Emerald Publishing Limited
e-ISSN: 1758-4094
p-ISSN: 1463-5771
DOI 10.1108/BIJ-12-2024-1087

performance into (ESG) principles. The ESG word, first introduced by the United Nations in 2004 through the "Who Cares Wins (n.d.)" reports, guides investors in integrating environmental, social, and governance factors into their investment decisions (United Nations, 2020).

With the growing interest, ESG reporting has become a pivotal framework for assessing corporate sustainability (Kolk and Perego, 2010; Zhou, 2022). Effective ESG management mitigates environmental degradation, social controversies, and governance scandals risk, thereby averting financial and reputational harm (Eccles *et al*., 2014). Moreover, robust ESG performance can lead to cost savings through enhanced efficiencies, reputation, financing, profitability, and lower capital costs. These advantages, however, are contingent on market conditions, regulations, and firm actions (Busch *et al*., 2016; Eccles *et al*., 2014). Hence, the influence of ESG on the performance of corporations can be complex, influenced by the company's operations and commitment to adopting eco-friendly practices (Rahman *et al*., 2023).

The relationship between ESG and corporate performance is intricate and context-dependent, yielding mixed results across past studies (Alareeni and Hamdan, 2020; Jyoti and Khanna, 2021; Rahman *et al*., 2023; Savio *et al*., 2023; Zhou *et al*., 2022). Some research reflects that ESG benefits (Rahman *et al*., 2023), while others report insignificant or negative (Almulhim and Aljughaiman, 2023; Brammer *et al*., 2005; Jain *et al*., 2019). Discrepancies arise from data selection, sample size, industries, timing, methodology, and study context (i.e. developed vs. developing countries listed or following a particular index) (Busch *et al*., 2016). Furthermore, ESG approaches and performance metrics further complicate the findings, especially in emerging markets (Orlitzky *et al*., 2003; Orlitzky and Shen, 2013). These inconsistencies underscore the need for more rigorous research on this domain (Khan, 2022; Lin *et al*., 2019).

Due to significant challenges, industries with high environmental impact, such as oil, chemicals, and mining, face increasing regulatory pressure to improve their environmental performance (Pisani *et al*., 2017). Indian firms are no exception, including those involved in environmental disasters over the past decades; notable examples include the Bhopal disaster in 1984, the world's worst industrial accident caused by methyl isocyanate gas exposure, and the 2020 Assam gas leak, which resulted in a blown out, methane leakage, subsequent fires, fatalities, and environmental damage. These incidents have heightened scrutiny of corporate sustainability. Hence, Indian firms now prioritize ESG disclosure to enhance credibility and secure stakeholder support (Braam and Peeters, 2018; Dong *et al*., 2014; Tadros and Magnan, 2019). However, such reporting is often met with skepticism, as stakeholders question the credibility of environmental claims (Cho and Patten, 2007). To address this, many companies seek external assurance from an independent third party to enhance their ecological disclosure credibility and safeguard their image and reputation by procuring voluntary external assurance services (Gillet-Monjarret, 2015; Cho *et al*., 2014).

This study categorizes firms based on environmental sensitivity. Following the definition by Halkos and Skouloudis (2016), environmentally sensitive industries comprise sectors whose business activities significantly impact the socio-environmental landscape. In line with prior studies, Gracia and Siregar (2021), Miralles-Quirós *et al*. (2018), and Yoon *et al*. (2018), we classify industries within the energy, material, and utilities sectors as environmentally sensitive industries, following the Global Industry Classification Standard (GICS). All other industries are categorized as NSI industries. Unlike prior studies that primarily focused on developed economies (Ademi and Klungseth, 2022; Hamrouni *et al*., 2019), this study diverges from and extends the literature by investigating the relationship between ESG disclosure and corporate performance in SI and NSI companies in India, a rapidly growing economy. We examine how ESG disclosures affect operating, financial, and market performance throughout 2015–2023, encompassing Indian listed firms. This study determines whether there is any difference in industries' profitability due to sustainability, highlighting the varying effects on corporate performance in SI and NSI (Arian *et al*., 2023;

Omar and Zallom, 2016). Moreover, we incorporate individual ESG dimensions — environmental, social, and corporate governance —offering nuanced insights into their distinct impacts. At this end, the following research question will be explored. How does the extent of ESG disclosure impact corporate performance (i.e. operating, financial, and Tobin's Q) in the environmentally sensitive and non-sensitive industries of Indian companies? Is there any difference in the association between ESG (and its dimensions) and corporate performance based on the environmental sensitivity of the corporation?

Notably, we focused on India for several reasons. India became one of the world's top five economies in terms of Gross domestic product (GDP) [1] and most prominent in terms of demographic footprints, overtaking China [2]. It emerged as a leading participant in sustainability discourse, driven by its unique institutional environment, cultural diversity, and evolving regulatory framework. A notable shift towards sustainable practices is evident, as Indian corporations are progressively embracing Corporate Social Responsibility (CSR) initiatives and releasing comprehensive "Business Responsibilities and Sustainability Reports (BRSR)" in 2021 as a result of regulatory requirements such as the Companies Act of 2013 and supervision from the Securities and Exchange Board of India (SEBI). With mounting scrutiny from regulators and investors, the volume of ESG research has seen significant growth, particularly in exploring the interaction of ESG factors and profitability, particularly in the Indian context (Narula *et al*., 2024; Nenavani *et al*., 2024; Sharma *et al*., 2019; Suprabha *et al*., 2024; Veeravel *et al*., 2024a, b; Yadav and Jain, 2023).

Furthermore, India's rapid economic growth and technological advancement have also intensified environmental challenges (Malik and Kashiramka, 2024). Additionally, the country's vulnerability to climate change (manifested in rising sea levels and extreme weather events) emphasizes the critical need for corporations to adopt robust environmental practices that mitigate risks and contribute to a sustainable future. Hence, Indian firms' environmental commitments are pivotal in global climate change efforts (Lu and Cheng, 2023; Zhao *et al*., 2018). Addressing social issues like poverty, inequality, and access to basic needs aligns corporate strategies with societal goals. Strong governance practices enhance accountability, attract investment, and build trust in the evolving market. Furthermore, increasing public awareness, stakeholder activism, and shifting consumer preferences have intensified the demand for ESG adoption. Moreover, India's prominence as a top investment destination in Europe and the US, driven by efforts to reduce reliance on uncertain economies, further amplifies the global significance of its sustainability practices (Narula, 2012; Palmate *et al*., 2021; Rao *et al*., 2023). India is also at the forefront of development among emerging nations [3], offering significant value as a model that can be replicated in other emerging countries. The country's evolving sustainability landscape, regulatory advancements, and corporate strategies for ESG provide actionable insights for nations with similar socioeconomics and institutional contexts.

This study makes several contributions to the existing literature. First, it addresses a critical gap by focusing on the association between ESG reporting and corporate performance in SI and NSI companies. It also conducts an industry-focused study for India, providing implications for other emerging markets. Second, it explores how environmental sensitivity influences ESG's impact on operational, financial, and market performance, providing an in-depth understanding of the orientation diversity and sensitivity in ESG practices. Third, it considers individual ESG dimensions (E, S, and G) beyond overall ESG impact, offering different perspectives. Fourth, the study's rigorous methodology uses a robust econometric approach, specifically system-GMM, to control for time effects and potential endogeneity issues, thereby enhancing the credibility and generalizability of the findings. Fifth, we employed a sample of Indian listed companies rather than relying on any specific representative index, enhancing its generalizability across- small, medium, and large enterprises. This broad sampling approach ensures that the findings are equally replicable to various companies in the securities basket, providing more comprehensive insights to managers and investors. Finally, the empirical results assist in designing legal and regulatory

frameworks for enhancing sustainability in respective industries. As a result, it suggests that ESG factors influence corporate outcomes differently across both industry types, and simply following sustainability- is somewhat naïve. Across listed businesses in SI, there is an insignificant association with strengthening the credibility of their non-financial reporting and implementing effective systems for corporate governance, reducing profitability as sensitive firms face more challenges and scrutiny from stakeholders. Meanwhile, in NSI companies, ESG practices positively and significantly enhance performance, highlighting their importance for long-term value creation and investor recognition. The varying effects of environmental, social, and governance (ESG) factors underscore the need for companies to develop customized sustainable strategies to see both SI and NSI corporations succeed economically.

2. Theoretical background and literature review

The relationship between ESG and corporate performance has garnered significant attention, with numerous meta-analyses and literature surveys examining the dynamic (Brooks and Oikonomou, 2018; Friede *et al*., 2015; Hang *et al*., 2019). For instance, Peloza (2009) and Friede *et al*. (2015) report that most research indicates positive results, attributing this to competitive advantage (Danso *et al*., 2019) by enhancing reputation (Zimon *et al*., 2022) and consumer loyalty (Orlitzky *et al*., 2003). This, in turn, is expected to boost profits (Freeman, 1984). Furthermore, Chen *et al*. (2023a, b) found that ESG substantially impacts profitability in large-scale companies and high-risk scenarios. Conversely, some studies highlight the costs and time required for sustainable practices, which may reduce profitability (Jaeger and Upadhyay, 2020; Cerciello *et al*., 2023). Others, like Franco *et al*. (2020), discover a U-shaped, indicating the complexity of ESG's impact.

2.1 Aggregating ESG and corporate performance

Despite the extensive studies, there still needs to be a consensus on the impact of ESG disclosure on firm profitability (Khandelwal *et al*., 2023; Lu and Cheng, 2023). For instance, Salehi and Arianpoor (2021) reveal a mutual relationship between financial and non-financial sustainability performance, emphasizing the need for a comprehensive approach to assessing sustainability performance across various components. Variability in empirical findings from diverse characteristics of ESG data and methodologies employed in its collection and analysis (Busch and Hoffmann, 2011; Chatterji *et al*., 2009; Stefan and Paul, 2008). Studies reveal that firms disclose their ESG commitment to meet legitimacy criteria and improve performance (Lee and Raschke, 2023). Contrary to these observations, agency cost theory posits that sustainability disclosures adversely impact the firm value due to associated costs (Giannopoulos *et al*., 2022; Hussaini *et al*., 2021; Lahouel *et al*., 2021). Nevertheless, these practices may yield positive effects in the long term (Behl *et al*., 2022).

Stakeholder theory claims that fostering positive relationships with internal and external stakeholders through corporate disclosure enhances performance, thereby ensuring long-term business sustainability (Freeman, 2010). Also, ESG practices assist in bolstering efficiency and value (Mohammad and Wasiuzzaman, 2021). Chelawat and Trivedi (2016) assess that ESG positively influences not only market value but also the operating and finances of the firms. A strong ESG score enhances reputation (Pham and Tran, 2020), lower debt cost (Gigante and Manglaviti, 2022), and improves capital access at a lower rate (Ramirez *et al*., 2022; Wong *et al*., 2021). These benefits are observed in developed economies (Gartia and Panda, 2024) due to stringent regulations and effective ESG communication. Recently, emerging economies have also moved towards mandatory regulations. For instance, Indian regulators have mandated ESG disclosures for India's top 1,000 listed companies. Complying with these requirements necessitated substantial investments, meaning firms must have adequate financial resources to support and enhance their ESG disclosures.

In India, the impact of ESG disclosures on corporate performance yields mixed results. For instance, a 2024 study by Veeravel *et al.* (2024a, b) using GMM methodology reveals that better ESG disclosures lead to superior financial performance. Anklesaria-Dalal and Thaker (2019), Chelawat and Trivedi (2016), and Sinha Ray and Goel (2023) similarly note that improved accounting and market-based outcomes for firms with strong ESG performance. However, Bansal *et al.* (2022) find benefits only when sustainability reporting exceeds a certain threshold. Gupta *et al.* (2021) emphasize the significant role of environmental sustainability across performance dimensions. In contrast, Jha and Rangarajan (2020) reported no substantial link, indicating that firms with varying sustainability measures performance are similar in economic measures. In another study, Mishra *et al.* (2024) show a significant negative of ESG disclosures on performance measures. This diverse body of literature highlights that ESG factors can vary widely depending on the specific measures and thresholds considered.

2.2 Environmental practices and corporate performance

Environmental practices have been a focal point for researchers (Pinheiro *et al.*, 2022), with their financial results have yielded contrasting results, being described as positive, negative, insignificant, U-shaped, and even reverse U-shaped (Guenther and Hoppe, 2014; Lee *et al.*, 2016; Lewandowski, 2017; Wagner, 2005). Some studies reveal that environmentally friendly activities yield superior profits (Blanco *et al.*, 2009; Molina-Azorín *et al.*, 2009; Nakao *et al.*, 2007; Schultze and Trommer, 2012). As a pioneering study, Cohen *et al.* (1997) analyzed S&P 500 constituents between 1987 and 1989 and found that environmental compliance positively affected return on assets and equity. Subsequent studies have corroborated these findings with meta-analyses, confirming a linear and positive association between corporate environmental performance and financial profitability. This positive is observed in developed countries (Campbell, 2007; Fijałkowska *et al.*, 2018) and emerging economies (Manrique and Martí-Ballester, 2017; Voinea *et al.*, 2020). The natural resource-based viewpoint is that corporate environmental practices offer a competitive advantage, resulting in better financial outcomes (Hart, 1995; Hart and Dowell, 2011; Russo and Fouts, 1997). Another strand of literature reports a negative association (Boakye *et al.*, 2021; Hasan *et al.*, 2022; Preston and O'Bannon, 1997; Wong *et al.*, 2018). For instance, Xie *et al.* (2019) found a significant negative effect of environmental commitment on profitability across nearly 7,000 firms in 74 countries. According to neoclassical theories, the costs of implementing ecological activities can outweigh the benefits, thus reducing competitiveness (Ayayi and Wijesiri, 2022; García-Sánchez and Prado-Lorenzo, 2012; Jo *et al.*, 2015; Lioui and Sharma, 2012), since environmental initiatives require time and cost.

A third tranche of the literature suggests that the relationship between environmental practices and financial profitability is more complex than simply negative and positive; some research argues for a neutral impact (Ortiz-de-Mandojana and Aragon-Correa, 2015; Zou *et al.*, 2015) while others U-shaped and inverse U-shaped effect (Boakye *et al.*, 2020; Guo and Lu, 2021). For instance, Horváthová (2012) found that the relationship between environmental practices and profitability initially appears hostile but turns positive over time. The lack of consensus partly stems from the heterogeneity of measures to capture the environmental impact. Different methodologies and contexts make it challenging to generalize findings.

2.3 Social practices and corporate performance

The relationship between CSR and firm performance continues to be a topic of significant debate (Adamkaite *et al.*, 2023), with studies offering three main perspectives: positive, negative, and non-linear. The first tranche of literature supports stakeholder theory and supports a positive association between CSR and financial performance. According to Freeman (1984), firms exist within an ecosystem of stakeholders, including employees, customers, suppliers, communities, and shareholders. Engaged stakeholders contribute to

better financial outcomes, ensuring long-term success (Mhlanga and Moloi, 2020) and sustainability (Kitchen, 2020). Conversely, another tranche posits that CSR harms financial performance (Aupperle *et al*., 1985; Graves and Waddock, 2000). A recent study using data from 15 countries (2011–2020) reveals that CSR hinders firm performance (Hsiao *et al*., 2024). This view is supported by the trade-off theory, which argues that pro-social activities impose additional costs on firms, ultimately reducing profitability (Cai *et al*., 2016; Friedman, 1970).

Research by Nollet *et al*. (2016) and Mittal *et al*. (2008) concludes that firms committed to CSR generate significantly less economic value. The agency theory further supports this view, suggesting that managers pursue social contribution for personal reputation at the expense of shareholders, leading to deterioration in firm wealth (Barnea and Rubin, 2010). Consequently, social activities are often seen as problems with the managerial agency (Benbou and Tirole, 2010; Masulis and Reza, 2013). Other studies indicate no association (Garcia-Castro *et al*., 2010; McWilliams and Siegel, 2000) or non-linear relationships (Flammer, 2015). For instance, Wagner and Schaltegger (2004) show that CSR and firm outcomes have an inverted U-shaped curve relationship based on data from the EU manufacturing industry. Similarly, Kim and Statman (2012) report that social commitment increases financial profitability only up to a certain point, after which it has a negative performance effect, aligning with this agency's view (Flammer, 2015).

2.4 Governance practices and corporate performance

This relationship has been extensively reviewed, with studies presenting various perspectives (Martínez-Ferrero and Frías-Aceituno, 2015; Nguyen *et al*., 2021; Rodriguez-Fernandez, 2016). Corporate governance arises from the separation of business ownership (principals/shareholders) and the management (agents), encompassing the rules and practices that guide firm operations and align stakeholders' interests (Cuervo, 2002). Effective corporate governance promotes sustainable development and ethical business practices (Lambertini and Tampieri, 2015; Manasakis *et al*., 2014) while mitigating sustainability-related controversies. The board of directors is critical in ensuring the firm's financial viability (Agnese *et al*., 2023).

A significant body of literature suggests a positive association between corporate governance and financial profitability. Stewardship theory posits that employees are intrinsically motivated to perform the tasks, leading to improved business track records (Davis *et al*., 1997; Muth and Donaldson, 1998). Donaldson and Preston (1995) claims that top management does not need to oversee manager's work closely, as they aim to achieve superior financial results. Alternatively, resource dependency theory supports the positive relationship, emphasizing that governance must interact with external organizations to secure essential resources for firm survival (Pfeffer and Salancik, 1978). Implementing effective management frameworks and governance structures can improve financial results and competitive advantage (Hillman *et al*., 2000). Conversely, some studies identify a negative relationship (Agrawal and Knoeber, 1996; Saidat *et al*., 2019). Cherian *et al*. (2019) find that specific corporate board characteristics negatively impact ROE. Similarly, Paniagua *et al*. (2018) analyzed 1,207 companies from 59 countries across 19 sectors (2013–2015) and observed the adverse effects of corporate governance on performance. The agency cost theory explains these results, posing that those managers may not act in the owners' best interest due to a lack of incentives (Jensen and Meckling, 1976). Although appropriate rewards can mitigate the agency problem (Grossman and Hart, 1986), they might backfire, leading managers to pursue personal goals over the firm's best interest (Brown, 2006; Krüger, 2015).

2.5 ESG disclosure and controversial environmentally sensitive firms

The variability in social or environmental disclosure content across different industries is a focal point of current research (Branco and Rodrigues, 2008; Teng *et al*., 2021). Studies by

Han *et al*. (2021) and Huang *et al*. (2024) emphasize the challenges faced by inherently more polluting industries such as thermal power generation and steel production. These sectors, characterized by significant environmental footprints, are under intense stakeholder pressure and social scrutiny, which drive them towards extensive ESG disclosure. Kuo and Chang (2021) note that firms in SI face unique challenges in ESG reporting, using Sustainable Development Goals (SDG) reporting as a strategy to mitigate negative public image. Since 2015, research has focused on ESG within specific sectors, particularly the energy industry (Horobet *et al*., 2024; Makridou *et al*., 2024). For instance, Zhao *et al*. (2017) find that high ESG can enhance the financial performance of firms in Chinese energy power. In contrast, Adamkaite *et al*. (2023) conclude that sustainability efforts do not impact the economic performance of nine Lithuanian energy companies. Similarly, Constantinescu *et al*. (2021) discover a varied connection between ESG and market values across various sectors, such as oil and gas, renewable energy, and utilities, indicating both positive and a negative link. Another study in the Indian context finds that while ESG disclosure positively affects firm performance, the relationship is moderated by industry climate sensitivity, with the effect being more substantial for less climate-sensitive industries (Maji and Lohia, 2024).

Although the global interest in ESG (Arianpoor and Salehi, 2020; Chen and Zhang, 2024; Chen *et al*., 2023a, b; Nie *et al*., 2024; Salehi *et al*., 2018; Salehi and Arianpoor, 2021; Shafeeq Nimr Al-Maliki *et al*., 2023; Zhou *et al*., 2024; Zimon *et al*., 2022), few studies have differentiated the impact between environmentally sensitive and non-sensitive industries (Naeem *et al*., 2022). Research indicates that firms in SI exhibit higher environmental performance, where they define sensitive industries as those operating in sectors marked by societal taboos, moral conflicts, and political pressure, including weapons, nuclear, biotech, gambling, alcohol, and adult entertainment. Moreover, environmentally sensitive industries encompass energy, mining, metals, construction, chemicals, pulp, and paper, known for their significant environmental impact and regulatory scrutiny (Garcia *et al*., 2017). Companies in these industries invest heavily in environmental and social management to mitigate their operational impact and face stricter regulatory frameworks and social pressures due to their substantial environmental footprints (Miralles-Quirós *et al*., 2018). However, this study defines environmentally sensitive industries as those whose business activities significantly impact socio-environmental landscapes (Halkos and Skouloudis, 2016). In line with prior studies, Gracia and Siregar (2021), Miralles-Quirós *et al*. (2018), and Yoon *et al*. (2018), we classify industries within the energy, material, and utilities sectors as environmentally sensitive industries, following the GICS. All other industries are categorized as non-environmentally sensitive industries.

Previous studies point out different firm characteristics due to the firm's affiliation with SI (Iatridis, 2013; Miralles-Quirós *et al*., 2018; Wang *et al*., 2022; Yoon *et al*., 2018). ESG disclosure in these industries may not significantly contribute to corporate sustainable growth, potentially due to the stricter regulatory scrutiny and accusation of "greenwashing" (Ren *et al*., 2022b). This scrutiny leads firms to use impression management tactics in their sustainable disclosures (Appiagyei and Donkor, 2024; Hahn and Lülfs, 2013). Xue *et al*. (2023) highlight that regardless of the disclosure regime, an adverse relationship exists between disclosures and firm value, with the effect being more pronounced in SI companies.

Numerous academic studies indicate a positive association between a firm's market value and its ESG scores (Abdi *et al*., 2022; Bissoondoyal-Bheenick *et al*., 2023; Carnini Pulino *et al*., 2022; Chouaibi *et al*., 2021). However, these findings do not account for whether a firm belongs to an SI or NSI. The link between robust corporate ESG performance and increased corporate outcomes varies in both industries, highlighting the need for further research into the nuanced impacts of ESG on corporate performance. Given the significant influence on stakeholders and regulatory frameworks, particularly in SI companies (Wichianrak *et al*., 2023). Hence, it is imperative and novel to explore the differentiated impacts of ESG on SI and

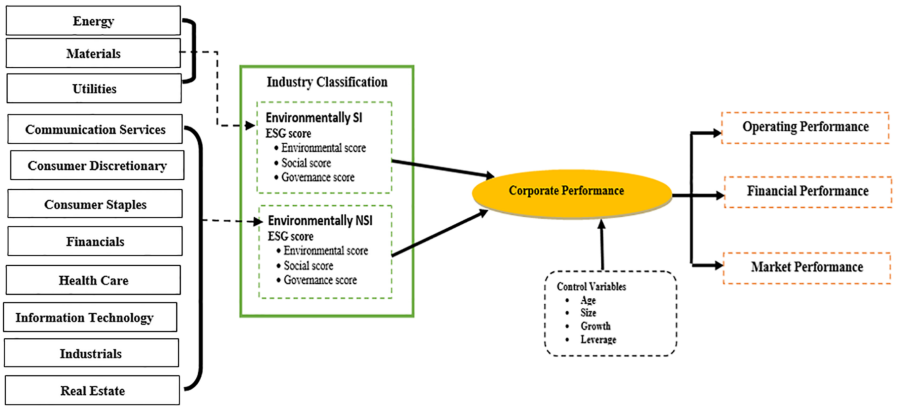
NSI corporate performance in India. To the best of our knowledge, no prior study has addressed this distinction analysis, as shown in the conceptual framework (Figure 1), which is crucial for developing targeted strategies to enhance corporate performance and sustainable growth.

3. Research design and methodology

3.1 Sample and data description

The study examines the differential impact of ESG and dimensions on corporate performance across SI and NSI industries among Indian companies listed on the NSE from 2015 to 2023. This investigation's initial period is prompted by the implementation of laws mandating the disclosure of information in 2014–15. A sample comprising 605 Indian-listed corporations, those who have headquarters in India. These corporations were selected based on the corporate metrics from Prowess databases and the availability of ESG disclosure data sourced from Refinitiv Eikon. Notably, the availability of ESG scores varies across corporations and years based on the availability of ESG data for selected companies. Consequently, sample size fluctuates annually based on the data availability of chosen companies.

To delve into the distinct industry contexts, the study classifies corporations based on the GICS classification provided by Morgan Stanley Capital International (MSCI) and Standard and Poor (S&P), which are utilized by the global community. Table 1 illustrates the sample's industry classification, ensuring a representative and unbiased portrayal of the broader economy. As shown in Panel A of Table 1, the companies represented in the sample are well-distributed across several sectors, with materials (19.67%), industrials (17.52%), and financials (15.7%) as the most prominent segments. At the same time, real estate, utilities, and information technology have the lowest concentration. Panel B classified industries into SI and NSI. The analysis explicitly defines the businesses operating in the energy, materials, and utility sectors as environmentally sensitive companies, following the framework employed by Yoon *et al*. (2018) and Gracia and Siregar (2021). The energy sectors encompass energy equipment and services and the oil, gas, and consumable fuel industries; the materials sector includes chemicals, construction materials, containers and packaging, metals and mining, and papers and forest products industries; the utility sectors include electric, gas, water, and power and renewable electric producers' industries. The categorization allows for a nuanced examination of how SI industries may interact with ESG factors in driving corporate performance.



Source(s): Authors' own work

Figure 1. Conceptual framework

Table 1. Distribution of the sample by the industry

	Panel A Sample distribution across industries				Panel B Sample classification between environmentally sensitive industries and non-sensitive industries	
	GICS sector name	Companies	Percent	Cum	Criteria	No. of companies
1	Communication Services	19	3.14	3.14	*Total listed companies in the sample*	605
2	Consumer Discretionary	94	15.54	18.68	Companies in sensitive industries include energy (4),	152
3	Consumer Staples	42	6.94	25.62	materials (9), and utilities (11)	
4	Energy	13	2.15	27.77		
5	Financials	95	15.7	43.47		
6	Health Care	50	8.26	51.74		
7	Industrials	106	17.52	69.26		
8	Information Technology	29	4.79	74.05	Companies in non-sensitive industries, all companies	453
9	Materials	119	19.67	93.72	excluding energy, materials, and	
10	Real Estate	18	2.98	96.69	utility sectors	
11	Utilities	20	3.31	100		
	Total	*605*	*100*			

Source(s): Authors' own work

3.2 Variable and measurement

The research has been facilitated by the utilization of pertinent empirical variables, which are elaborated upon in the subsequent description:

3.2.1 Dependent variables. Table 2 outlines the variables in this study, focusing on corporate performance, which is assessed using three key metrics: operating (indicated by ROA), financial (ROCE), and market (Tobin's Q) as defined by Narula *et al*. (2024), Veeravel *et al*. (2024a, b). Prior research has commonly utilized these metrics as outcome variables to examine the performance of corporations, employing appropriate panel data regression (Buallay, 2019; Alareeni and Hamdan, 2020; Pätäri *et al*., 2012; Choi and Wang, 2009).

We utilize ROA as a measure to evaluate operational performance; it assesses how effectively and efficiently a corporation utilizes its total assets to generate profit, thereby providing insights into its overall operational efficiency and performance (Velte, 2020; Wang and Sarkis, 2017). Additionally, ROCE is employed as an indicator of financial performance; it evaluates the firm's profitability and efficiency in utilizing its capital. It provided insights into how efficiently a firm generates returns from capital invested (Maqbool and Bakr, 2019; Veeravel *et al*., 2024a, b). Tobin's Q is employed as a measure of market performance; it is a market-based metric for assessing a company's performance and market valuation. In addition to market performance, it accounts for the long-term replacement cost of a country's assets, making it a significant indicator in sustainable investment strategies, as defined by Narula *et al*. (2024) and Veeravel *et al*. (2024a, b).

3.2.2 Independent variables. ESG scores from Refinitiv Eikon-DataStream reflect a composite measure derived from publicly available data, such as annual reports, stock market filings, corporate websites, newspapers, and CSR reports from reputable corporations across environmental, social, and corporate governance dimensions. This database objectively and transparently evaluates a company's ESG practices, accountability, and effectiveness. The scoring system evaluates companies using 186 key metrics and industry-specific criteria, generating a percentage score ranging from 0 to 100 (Bruna *et al*., 2022; Rajesh and Rajendran,

Table 2. Definitions of key variables

Symbol	Definition	Measurement	Source
Dependent Variables			
ROA	Return on Assets	Operating Performance: Earnings Before interest and tax divided by Total Assets	Prowess
ROCE	Return on Capital Employed	Financial Performance: Earnings Before interest and tax divided by Capital Employed	Prowess
TQ	Tobin's Q	(Market Capitalization + Debt)/ Total Assets	Prowess
Independent variables			
ESG	Environmental, Social, and Governance score	The aggregate score consists of parameters from Thomson Reuters	Refinitiv Eikon
ES	Environmental Score	The environmental score consists of parameters such as natural systems and ecosystems. Thomson Reuters has analyzed 57 indicators to measure the score	Refinitiv Eikon
SS	Social Score	The social score is the company's ability to increase loyalty and trust with its workforce, customers, and society. Thomson Reuters has utilized 60 indicators to measure the score	Refinitiv Eikon
GS	Governance Score	The governance score consists of parameters such as the systems and processes of a company, which are in the best interests of its shareholders. Thomson Reuters has assessed 48 indicators to measure the score	Refinitiv Eikon
Control variables			
Age	Firm Age	log of the fiscal year minus the incorporation year	Prowess
Size	Firm Size	Natural logarithm of total assets	Prowess
Growth	Firm Growth	determined by the growth rate of core business revenue	Prowess
DE	Firm Leverage	Total company debt divided by shareholder's equity	Prowess

Source(s): Authors' own work

2020; Sarkis and Cordeiro, 2001; Shakil, 2021). These scores reflect a firm's performance across ten critical E, S, and G themes. Where the E discloses the emissions, innovation, and resource usage. The S pillar includes human rights, workforce, community, and product responsibility. The G pillar covers shareholders, management, and CSR strategy. These data are calculated based on publicly available data utilizing Refinitiv's specific methodology [4].

3.2.3 Control variables. Moreover, this study incorporates control variables to enhance the study's internal validity, thus mitigating the confounding factors and bolstering the robustness of the findings, as shown in Table 2. We include the following control variables to enhance the efficiency of the measurement model. Firm age, measured as the natural logarithm of the fiscal year minus the incorporate year, reflects the influence of mature firms, which may possess more established operational structures and reputational capital, affecting their sustainability initiatives and financial performance (Doshi *et al*., 2024). The firm size, determined by the natural logarithm of total assets (Buallay, 2019; Hillman *et al*., 2000; Velte, 2017), is a key factor influencing corporate performance. Larger firms generally exhibit superior disclosure strategies compared to small companies (Veeravel *et al*., 2024a, b). Leverage, measured by debt-to-equity ratio, captures financial risk, with evidence suggesting that firms with lower debt levels tend to adopt stronger ESG practices than highly leveraged firms (Bhandari, 1988). High leverage may reduce operational efficiency by increasing interest expenses, which limits investment in growth and innovation. Moreover, it also impacts market performance, as investors may view highly leveraged firms as riskier. Another control variable is sales growth, which is calculated as annual changes in sales included to account for a firm's growth trajectory, which can impact the ability to allocate resources to sustainability initiatives (Bansal *et al*., 2021; Chen and Ibhagui, 2019).

3.3 Model specification

The interplay between ESG and corporate performance, along with relevant control variables, involves potential endogeneity issues, such as mutual influence, reverse causality, and simultaneity (Kahloul *et al*., 2022; Martínez-Ferrero and Frías-Aceituno, 2015). To navigate this, we employed the system-GMM, recommended by Arellano and Bond (1991) and refined by Arellano and Bover (1995) and Blundell and Bond (1998). This approach uses lagged variables as instruments, combining equations in first differences and levels, as shown in the given equations.

System GMM is widely used in ESG research (Agnese *et al*., 2023; Sinha Ray and Goel, 2023) due to its robust ability to mitigate econometric issues like weak instruments, size distortion, and omitted variable bias. By leveraging instruments from both level and first differenced equations, this technique effectively handles challenges in dynamic panel data, ensuring robust and precise estimates through asymptotic normality (Dwibedi *et al*., 2024).

$$ROA_{it}(SI) = \varnothing ROA_{it-1} + \beta_1 ESG_{it} + \beta_2 Age_{it} + \beta_3 Size_{it} + \beta_4 Growth_{it} + \beta_5 leverage_{it} + \varepsilon_{it}$$

$$ROCE_{it}(SI) = \varnothing ROCE_{it-1} + \beta_1 ESG_{it} + \beta_2 Age_{it} + \beta_3 Size_{it} + \beta_4 Growth_{it} + \beta_5 leverage_{it} + \varepsilon_{it}$$

$$Tobin'sQ_{it}(SI) = \varnothing TObin'sQ_{it-1} + \beta_1 ESG_{it} + \beta_2 Age_{it} + \beta_3 Size_{it} + \beta_4 Growth_{it} + \beta_5 leverage_{it} + \varepsilon_{it}$$

$$ROA_{it}(NSI) = \varnothing ROA_{it-1} + \beta_1 ESG_{it} + \beta_2 Age_{it} + \beta_3 Size_{it} + \beta_4 Growth_{it} + \beta_5 leverage_{it} + \varepsilon_{it}$$

$$ROCE_{it}(NSI) = \varnothing ROCE_{it-1} + \beta_1 ESG_{it} + \beta_2 Age_{it} + \beta_3 Size_{it} + \beta_4 Growth_{it} + \beta_5 leverage_{it} + \varepsilon_{it}$$

$$Tobin'sQ_{it}(NSI) = \varnothing Tobin'sQ_{it-1} + \beta_1 ESG_{it} + \beta_2 Age_{it} + \beta_3 Size_{it} + \beta_4 Growth_{it} + \beta_5 leverage_{it} + \varepsilon_{it}$$

ROA, ROCE, and Tobin's Q are operational, financial, and market performance indicators. Here, SI and NSI represent environmentally sensitive and non-sensitive industries. Concurrently, ESG (and E, S, and G pillars individually) scores are explanatory variables, with age, size, growth, and leverages as controls. Parameters $\varnothing and\ \beta$ represents coefficients, i denotes cross-sections, t represents time, and ε is the error term.

To ensure robustness, data were winsorized at the 1 and 99 percentiles. VIF (Variance Inflation Factor) and correlation tests have been employed to test multicollinearity, while Wooldridge and modified Wald tests addressed serial correlation and heteroscedasticity. Instrument validity was confirmed using Sargon and Hansen tests (Arellano and Bond, 1991), ensuring the reliability of the dynamic analyses.

4. Descriptive statistics and correlation matrix

Table 3 presents descriptive statistics of this study. The mean ROA is 0.111, indicating an average return on assets for the sample companies, ranging from – 16.1% to 43.3%, while ROCE averages 18.3%, surpassing the ROA return. Tobin's Q ratio averages 2.789, signifying that the market value of companies exceeds their replacement cost. The average ESG

Table 3. Descriptive statistics

Variable	Obs	Mean	SD	Min	Max
ROA	5,311	0.1111	0.0930	−0.1612	0.4328
ROCE	5,311	0.1830	0.1608	−0.2296	0.8423
Tobin's Q	3,430	2.7897	2.9620	0.1327	16.8364
ESG	2,100	0.4600	0.1605	0.1199	0.8224
Env	2,101	0.3913	0.2267	0.0000	0.9112
Soc	2,100	0.5001	0.2038	0.0894	0.9279
Gov	2,100	0.5018	0.2249	0.0692	0.9333
Age	5,409	1.4943	0.3299	0.0000	2.2041
Size	5,311	4.6612	0.7843	2.9308	6.8626
Growth	4,775	0.1615	0.4352	−0.6601	2.9679
DE	5,207	0.7141	1.3419	0.0000	8.6300

Source(s): Authors' own work

disclosure is 46 out of 100, reflecting a commendable level of performance among Indian corporations, with pillar-wise averages of 39.1 (environmental), 50 (social), and 50.2 (governance), indicating a balanced and upbeat performance. Control variables include age (mean 1.494), size (4.661), growth (0.162), and leverage (0.714), offering a comprehensive overview of the sample's financial and non-financial attributes.

Table 4 presents the Pearson correlation matrix, revealing varying associations between ESG and corporate performance across both industries. ESG and performance indicators exhibit negative correlations in SI industries, whereas NSI is generally favorable. However, the correlation matrix does not provide a definitive analysis. Therefore, we employed multivariate models to investigate these dynamics further. As a preliminary diagnostic tool, we employed the Pearson correlation matrix to identify potential multicollinearity among the variables within SI and NSI companies. Multicollinearity occurs when two or more predictors in multiple regression models are highly correlated (Kennedy, 2008; Wooldridge, 2015). While the correlation matrix offers initial insights, it must comprehensively depict the relationship. We calculated and displayed the mean VIF in their respective model to address this limitation, presenting a more nuanced understanding of multicollinearity.

5. Empirical results

We examine the relationship between ESG and corporate performance among SI and NSI companies using a two-step system-GMM—Tables 5-7 present operating, financial, and market performance results, respectively. Here, Model (1) evaluates the overall ESG score, while Models (2), (3), and (4) assess individual dimensions' impact, respectively. The estimated coefficients β and corresponding t-values (in parenthesis) are reported.

One concern is the potential endogeneity issue due to reverse causality between the variables. To address endogeneity, measurement errors, and unobserved heterogeneity, we employed system-GMM, controlling for time-fixed effect. The AR tests confirm no second-order correlation (with p-value>0.05), and the Sargon and Hansen tests ($p > 0.05$) validate the instrument sets. These results ensure the robustness and unbiasedness of the models.

5.1 Impact of ESG on the operating performance across SI and NSI corporations

Table 5 delineates ESG's impact on the operating performance of corporations for SI (as shown in Panel A) and NSI (in Panel B). For SI, the results indicate that ESG has an

Table 4. Pairwise correlation matrix

	1	2	3	4	5	6	7	8	9	10	11
Panel A: Sensitive Industries											
ROA	1.000										
ROCE	0.9237*	1.000									
TQ	0.3751*	0.3390*	1.000								
ESG	−0.1311*	−0.1090*	−0.029	1.000							
Env	−0.1544*	−0.1136*	−0.1085*	0.8269*	1.000						
Soc	−0.065	−0.026	−0.1246*	0.7902*	0.7242*	1.000					
Gov	−0.046	−0.080	0.1325*	0.3653*	0.052	−0.010	1.000				
Age	0.045	0.045	−0.0813*	0.2086*	0.1882*	0.2245*	0.0858*	1.000			
Size	−0.2250*	−0.2007*	−0.2588*	0.4227*	0.5394*	0.6120*	−0.1573*	0.2670*	1.000		
Growth	0.1015*	0.0927*	0.1414*	−0.0909*	−0.0979*	−0.1345*	0.037	−0.1439*	−0.047	1.000	
DE	−0.3162*	−0.3151*	−0.1928*	0.068	0.0878*	0.1369*	−0.1475*	−0.0832*	0.1897*	0.021	1.000
Panel B: Non-Sensitive Industries											
ROA	1.000										
ROCE	0.7076*	1.000									
TQ	0.4912*	0.4679*	1.000								
ESG	0.1555*	0.1072*	0.1376*	1.000							
Env	0.0653*	0.0896*	0.029	0.7105*	1.000						
Soc	0.1326*	0.1339*	0.0861*	0.8423*	0.6562*	1.000					
Gov	0.1202*	0.012	0.1189*	0.6780*	0.2846*	0.3626*	1.000				
Age	0.026	0.1589*	−0.0470*	0.0648*	0.1682*	0.1216*	−0.1034*	1.000			
Size	−0.2108*	0.0436*	−0.2433*	0.1683*	0.2531*	0.3063*	−0.035	0.2954*	1.000		
Growth	0.033	0.032	0.1226*	−0.045	−0.0791*	−0.0593*	−0.004	−0.1646*	−0.080	1.000	
DE	−0.2241*	−0.2296*	−0.1986*	−0.0858*	−0.1268*	−0.029	−0.0884*	−0.0908*	0.2855*	0.023	1.000

Note(s): Panels A and B show the correlation matrix of the studied variables in sensitive and non-sensitive industries, respectively, where * denotes *p*-value <0.05
Source(s): Authors' own work

Table 5. ESG and ROA: system GMM two-step model

	Panel A: Environmentally sensitive industries				Panel B: Environmentally non-sensitive industries			
	1	2	3	4	5	6	7	8
	ROA	ROA	ROA	ROA	ROA	ROA	ROA	ROA
L.ROA	0.570***	0.560***	0.545***	0.575***	0.514***	0.501***	0.503***	0.531***
	(4.01)	(3.95)	(3.66)	(4.23)	(5.27)	(4.86)	(5.02)	(5.4)
ESG	−0.0228				0.0515***			
	(−1.12)				(3.53)			
Env		−0.0141				0.0259**		
		(−0.83)				−2.55		
Soc			0.00178				0.0483***	
			(0.08)				(3.7)	
Gov				−0.0317**				0.0165*
				(−2.26)				(1.74)
Age	0.0105	0.00884	0.011	0.0131	0.0134	0.0134	0.0138	0.0170*
	(0.82)	(0.69)	(0.8)	(1.04)	(1.32)	(1.28)	(1.36)	(1.65)
Size	−0.000556	−0.00061	−0.00406	−0.0035	−0.0170***	−0.0171***	−0.0200***	−0.0126**
	(−0.08)	(−0.08)	(−0.45)	(−0.54)	(−2.84)	(−2.78)	(−3.35)	(−2.20)
Growth	0.0523***	0.0514***	0.0508***	0.0525***	0.0204***	0.0201***	0.0201***	0.0203***
	(3.57)	(3.53)	(3.5)	(3.51)	(3.00)	(3.03)	(2.97)	(3.06)
DE	−0.0190***	−0.0192***	−0.0188***	−0.0192***	−0.00504**	−0.00523**	−0.00500**	−0.00573**
	(−3.67)	(−3.65)	(−3.60)	(−3.99)	(−2.05)	(−2.11)	(−2.05)	(−2.39)
constant	0.0458	0.0458	0.0505	0.0563	0.0984***	0.114***	0.114***	0.0850***
	(0.85)	(0.81)	(0.85)	(1.07)	(3.03)	(3.32)	(3.40)	(2.65)
N	503	503	503	503	1,222	1,222	1,222	1,222
Arellano-Bond test for AR(1)	0.021	0.019	0.020	0.021	0.000	0.000	0.000	0.000
Arellano-Bond test for AR(2)	0.678	0.723	0.714	0.693	0.625	0.620	0.683	0.686
Sargon test	0.081	0.083	0.100	0.102	0.391	0.385	0.367	0.405
Hansen test	0.343	0.356	0.349	0.379	0.827	0.819	0.841	0.780
Max VIF	1.48	1.71	1.88	1.27	1.25	1.33	1.35	1.16
Mean VIF	1.23	1.31	1.38	1.14	1.12	1.18	1.16	1.07
Year effect	YES	YES	YES	YES	YES	YES	YES	YES

Note(s): This table uses the model: ROA ~ constant + ESG scores (E. S or G) + control variables + year fixed effect. Year-fixed effects are omitted for brevity. Panel A represents highly environmentally sensitive companies, whereas Panel B focuses on less sensitive ones. Robust t-statistics are in paratheses, and ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively. VIF represents the variance inflation factor

Source(s): Authors' own work

Table 6. ESG and ROCE: System GMM two-step model

	Panel A: Environmentally sensitive industries				Panel B: Environmentally non-sensitive industries			
	1	2	3	4	5	6	7	8
	ROCE	ROCE	ROCE	ROCE	ROCE	ROCE	ROCE	ROCE
L.ROCE	0.487***	0.483***	0.508***	0.515***	0.769***	0.770***	0.763***	0.785***
	(3.04)	(3.12)	(3.47)	(3.37)	(7.60)	(7.32)	(7.37)	(7.66)
ESG	−0.0377				0.0303*			
	(−1.22)				(1.65)			
Env		−0.0145				0.0181		
		(−0.61)				(1.43)		
Soc			0.0101				0.0329**	
			(0.33)				(2.00)	
Gov				−0.0603***				0.00351
				(−3.14)				(0.29)
Age	0.0304*	0.0275	0.0285	0.0326*	0.0200**	0.0194**	0.0203**	0.0215**
	(1.68)	(1.52)	(1.58)	(1.90)	(2.09)	(2.06)	(2.13)	(2.31)
Size	0.00127	−0.00017	−0.003	−0.00249	−0.00561	−0.00574	−0.00819	−0.00206
	(0.11)	(−0.01)	(−0.22)	(−0.24)	(−0.77)	(−0.75)	(−1.03)	(−0.31)
Growth	0.0728***	0.0731***	0.0747***	0.0743***	0.0383***	0.0388***	0.0381***	0.0392***
	(3.76)	(3.79)	(3.96)	(3.87)	(3.07)	(3.11)	(3.06)	(3.16)
DE	−0.0342***	−0.0344***	−0.0345***	−0.0348***	−0.00836**	−0.00831**	−0.00826**	−0.00897**
	(−5.42)	(−5.54)	(−6.51)	(−6.17)	(−2.15)	(−2.14)	(−2.14)	(−2.36)
_cons	0.052	0.0537	0.0379	0.0609	0.0277	0.0358	0.0388	0.0176
	(0.64)	(0.65)	(0.47)	(0.79)	(0.68)	(0.80)	(0.89)	(0.44)
N	503	503	503	503	1,222	1,222	1,222	1,222
Arellano-Bond test for AR(1)	0.03	0.028	0.027	0.027	0	0	0	0
Arellano-Bond test for AR(2)	0.692	0.765	0.791	0.715	0.62	0.943	0.914	0.937
Sargon test	0.093	0.089	0.099	0.114	0.385	0.068	0.068	0.074
Hansen test	0.298	0.326	0.383	0.356	0.819	0.406	0.392	0.393
Max VIF	1.48	1.71	1.88	1.27	1.25	1.33	1.35	1.16
Mean VIF	1.23	1.31	1.38	1.14	1.12	1.18	1.16	1.07
Year effect	YES	YES	YES	YES	YES	YES	YES	YES

Note(s): This table uses the model: ROCE ∼ constant + ESG scores (E. S or G) + control variables + year fixed effect. Year-fixed effects are omitted for brevity. Panel A represents highly environmentally sensitive companies, whereas Panel B focuses on less sensitive ones. Robust t-statistics are in paratheses, and ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively. VIF represents the variance inflation factor

Source(s): Authors' own work

Table 7. ESG and Tobin's Q: System GMM two-step model

	Panel A: Environmentally sensitive industries				Panel B: Environmentally non-sensitive industries			
	1	2	3	4	5	6	7	8
	TQ	TQ	TQ	TQ	TQ	TQ	TQ	TQ
L. TQ	0.797***	0.807***	0.782***	0.789***	0.924***	0.925***	0.937***	0.926***
	(4.41)	(4.34)	(4.51)	(4.50)	(13.80)	(13.56)	(13.57)	(16.09)
ESG	0.752				1.612*			
	(1.26)				(1.94)			
Env		0.647				2.066**		
		(1.06)				(2.56)		
Soc			0.87				1.512**	
			(1.40)				(2.09)	
Gov				0.259				0.536
				(0.76)				(1.62)
Age	−0.614*	−0.601*	−0.642**	−0.690**	0.416	0.288	0.467	0.451
	(−1.87)	(−1.88)	(−2.02)	(−2.11)	(1.29)	(0.89)	(1.44)	(1.39)
Size	−0.458	−0.492	−0.539*	−0.398*	−0.944**	−1.212***	−1.031**	−0.678**
	(−1.53)	(−1.52)	(−1.66)	(−1.69)	(−2.42)	(−2.65)	(−2.43)	(−2.17)
Growth	0.123	0.126	0.13	0.123	0.156	0.187	0.133	0.169
	(0.89)	(0.92)	(1.03)	(0.90)	(0.70)	(0.82)	(0.57)	(0.75)
DE	−0.112	−0.0935	−0.116	−0.124	−0.0184	−0.0202	−0.0181	−0.0312
	(−0.69)	(−0.50)	(−0.76)	(−0.73)	(−0.23)	(−0.24)	(−0.21)	(−0.43)
_cons	3.769**	3.954**	4.177**	3.837**	2.25	4.772**	2.494	1.485
	(2.09)	(2.00)	(2.24)	(2.25)	(1.43)	(2.45)	(1.48)	(1.09)
N	445	445	445	445	951	951	951	951
Arellano-Bond test for AR(1)	0.048	0.049	0.044	0.046	0.001	0.001	0.001	0.000
Arellano-Bond test for AR(2)	0.481	0.485	0.495	0.474	0.604	0.432	0.604	0.689
Sargon test	0.156	0.145	0.181	0.153	0.706	0.725	0.706	0.581
Hansen test	0.14	0.109	0.203	0.168	0.116	0.189	0.116	0.113
Max VIF	1.46	1.72	1.82	1.25	1.18	1.29	1.28	1.09
Mean VIF	1.22	1.32	1.37	1.15	1.10	1.15	1.14	1.05
Year effect	YES	YES	YES	YES	YES	YES	YES	YES

Note(s): This table uses the model Tobin's Q ~ constant + ESG scores (E. S or G) + control variables + year fixed effect. Year-fixed effects are omitted for brevity. Panel A represents highly environmentally sensitive companies, whereas Panel B focuses on less sensitive ones. Robust t-statistics are in paratheses, and ***, **, and * indicate statistical significance at 1%, 5%, and 10%, respectively. VIF represents the variance inflation factor

Source(s): Authors' own work

insignificant (−0.0228) effect (Chung *et al*., 2024). Supporting this, Yoon *et al*. (2018) suggested that firms in sensitive industries are held to higher standards to address the environmentally harmful activities associated with their operations. These standards may lead to costs for these corporations rather than benefits. For instance, mining, cement, and chemical industries with stricter sustainability practices may impose significant compliance costs. In contrast, ESG positively and significantly influences ROA ($\beta = 0.514$, p-value<0.01) across NSI firms, indicating that investing in companies with vital ESG initiatives can yield higher returns (Xu and Zhu, 2024; Veeravel *et al.*, 2024a, b). Moreover, robust ESG practices enhance corporate reputation, bolster stakeholder confidence, and improve the efficiency of operations. This observation is consistent with stakeholder theory, which posits that ethical and moral considerations are critical in organizational decisions (Freeman *et al*., 2020). Consequently, businesses that effectively manage ESG risks and capitalize on ESG opportunities exhibit superior performance (Dong *et al*., 2023; Rahman *et al*., 2023; Zhao *et al*., 2018). These findings align with numerous seminal studies documenting a favorable association between ESG and firm performance (Albitar *et al*., 2020; Rahman *et al*., 2023). This finding underscores the strategic importance of integrating ESG considerations into management to achieve sustainable financial success.

On a standalone analysis, the analysis reveals that environmental ($\beta = 0.0259$, p-value<0.05), social ($\beta = 0.0483$, p-value<0.05), and governance ($\beta = 0.0165$, p-value<0.10) exhibit significant positive impacts NSI companies. Social factors, in particular, are critical to improving outcomes, highlighting stakeholders positively evaluating social concerns (Fosu *et al*., 2024). These findings underscore the importance of societal contributions in fostering value creation and enhancing public image, thereby encouraging stakeholder investment. Conversely, in SI firms, we find that while aggregate ESG disclosure does not have a significant impact (Masongweni and Simo-Kengne, 2024), governance factors ($\beta = -0.0317$, p-value<0.05) show a negative and significant impact, contradicting stakeholder theory, aligning with the findings of Yoon *et al*. (2018). This relationship is likely due to the burden of regulatory burdens, suggesting that increased focus on sustainability can detract from corporate objectives. These industries often have stringent disclosure requirements due to their negative potential environmental impacts, which can hinder performance (Yoon *et al*., 2018). These findings highlight ESG's varied effects across different industries, shaped by regulatory demands and stakeholders' expectations.

Control variables show a negative and significant association between size and leverage with the operating performance, indicating that smaller firms are generally more profitable and higher leverage increases risk, reducing profitability, aligning with the pecking order theory, which posits that higher debt levels reduce profit margins (Gupta and Aggarwal, 2024; Malik and Kashiramka, 2024; Veeravel *et al*., 2024a, b). Conversely, growth consistently shows a positive impact across industries.

5.2 Impact of ESG on the financial performance of across SI and NSI corporations

In Table 6, Panel A represents the impact of ESG disclosure and its dimensions on the financial performance of SI companies, whereas Panel B focuses on NSI companies.

The analysis indicates that ESG positively and significantly impacts ROCE for NSI ($\beta = 0.0303$, p-value<0.10) but is insignificant for SI (−0.0228), highlighting that NSI firms benefit more from sustainability initiatives (Bocken and Geradts, 2020). In contrast, SI companies face challenges like high compliance costs, substantial investments, and short-term financial focus, diluting ESG's immediate initiatives. Like ROA, we decompose ESG performance into individual components and find that only governance disclosure significantly negatively impacts sensitive corporations ($\beta = -0.0603$, p-value<0.01). This negative valuation effect is inconsistent with the stakeholder theory. However, it aligns with the trade-off theory, which may emphasize the substantial costs and complexities of maintaining strong governance standards. In NSI, social performance ($\beta = 0.0329$, p-value<0.05) significantly impacts financial performance, enhancing productivity, brand reputation, loyalty, and ethical investors.

These factors bolster companies' financial stability and growth prospects. Whereas in the SI industries, the impact of governance disclosure is significant but negative; in other words, we can say that the firms with better governance practices would have lower market value (Yoon *et al*., 2018).

Among control variables, growth and age significantly positively affect financial performance. This indicates that firms experiencing growth and older trends perform better financially. Conversely, leverage shows a negative and significant association with ROCE, suggesting that higher debt levels reduce profitability.

5.3 Impact of ESG on the market performance across SI and NSI corporations

As shown in Table 7, results based on market factors also demonstrate a stark contrast, measured by Tobin's Q, between SI and NSI corporations. ESG factors and their dimensions do not significantly affect SI companies' market performance (Atan *et al*., 2018). The reason could be that investors and stakeholders lack confidence and trust in these corporations' ESG practices and operations in emerging markets (Naeem *et al*., 2022). Moreover, another explanation could be that investors in these industries may prioritize immediate financial returns over long-term ESG improvements, leading to an insignificant impact. For SI firms, ESG disclosure may function as a signaling and reputation mechanism, fostering sustainable growth. However, stricter environmental in recent years have complicated their effectiveness. Instead of attracting stakeholders, such disclosures risk being perceived as "greenwashing", potentially undermining their role in fostering sustainable corporate success (Ren *et al*., 2022a; Wang *et al*., 2022).

In contrast, analysis indicates a significant positive impact of ESG ($\beta = 1.612$, p-value<0.10), with environmental ($\beta = 2.066$, p-value<0.05) and social ($\beta = 1.512$, p-value<0.05) dimensions playing a pivotal role. This implies that strong ESG practices are more highly valued by investors, boosting market valuation, as reflected by Tobin's Q (Malik and Kashiramka, 2024). Improvements in environmental and social practices enhance brand reputation and customer loyalty, which are recognized and rewarded. Additionally, firm age significantly impacts sensitive industries, whereas size negatively and substantially impacts non-sensitive sectors.

5.3.1 Robustness test. To test the robustness of this research, we simultaneously regressed the individual E, S, and G scores on corporate performance metrics (ROA, ROCE, and Tobin's Q) in a unified model. This approach corroborated previous conclusions, demonstrating consistent outcomes across analytical approaches. This not only strengthens the reliability of the study but also underscores the significance of considering each ESG dimension independently when assessing its impact on outcomes in SI and NSI. Therefore, the previous findings are robust, as shown in Table 8.

6. Discussion and contribution of the study

Intriguingly, the summarized results (as shown in Table 9) suggest the differentiating impacts of ESG factors on corporate outcomes across industry types. In SI companies, ESG initiatives may insignificantly influence performance indicators in emerging economies, as firms often engage in ESG activities primarily to meet regulatory obligations and social contracts rather than to drive profitability (Naeem *et al*., 2022). Additionally, in line with recent research on sustainable practices, factors such as organizational productivity, macro policies, and technological advancements have been identified as crucial for enhancing the performance of sustainability-driven models (Bagherpasandi *et al*., 2024a, b). Consequently, companies in these industries may face a challenging balance between regulatory demands and profitability in emerging nations. Furthermore, among ESG dimensions, Governance operations, in particular, exhibit an inverse relationship, aligning with shareholder theory (Friedman, 1970) and incurring reputational expenses (Mendiratta *et al*., 2023; Nirino *et al*., 2021; Wu *et al*., 2023), high innovational costs (Liu, 2023), and regulatory burdens (Gracia and Siregar, 2021;

Table 8. Robustness test: Model restructuring

	Panel A: Environmentally sensitive industries			Panel A: Environmentally non-sensitive industries		
	ROA	ROCE	TQ	ROA	ROCE	TQ
L.ROA	0.550***			0.507***		
	(3.89)			(5.15)		
L.ROCE		0.525***			0.760***	
		(3.68)			(7.38)	
L.TQ			0.773***			0.927***
			(4.40)			(12.82)
Env	−0.012	−0.0146	0.363	0.00155	0.00374	1.749**
	(−0.65)	(−0.63)	(0.54)	−0.14	(0.26)	(2.27)
Soc	0.0186	0.0318	0.603	0.0452***	0.0349*	0.623
	(0.77)	(0.88)	(1.06)	(2.99)	(1.77)	(0.95)
Gov	−0.0350**	−0.0628***	0.194	0.00403	−0.00817	−0.0147
	(−2.48)	(−3.49)	(0.56)	(0.42)	(−0.61)	(−0.04)
Age	0.0126	0.0296*	−0.587*	0.0137	0.0198**	0.299
	(1.00)	(1.82)	(−1.70)	(1.35)	(2.07)	(0.92)
Size	−0.00599	−0.00545	−0.592*	−0.0198***	−0.00911	−1.269***
	(−0.70)	(−0.44)	(−1.67)	(−3.31)	(−1.09)	(−2.63)
Growth	0.0505***	0.0746***	0.143	0.0203***	0.0382***	0.172
	(3.45)	(4.02)	(1.04)	(3.00)	(3.08)	(0.75)
DE	−0.0200***	−0.0360***	−0.0822	−0.00492**	−0.00818**	−0.0122
	(−4.05)	(−7.20)	(−0.45)	(−1.98)	(−2.11)	(−0.14)
constant	0.0772	0.0879	4.238**	0.112***	0.0464	4.857**
	(1.28)	(1.07)	(2.16)	(3.35)	(1.00)	(2.41)
N	503	503	445	1,222	1,222	951
Arellano-Bond test for AR(1)	0.024	0.028	0.048	0.000	0	0.001
Arellano-Bond test for AR(2)	0.667	0.732	0.491	0.688	0.922	0.458
Sargon test	0.108	0.125	0.184	0.355	0.068	0.701
Hansen test	0.388	0.467	0.139	0.838	0.39	0.178
Max VIF	2.49	2.49	2.46	2.04	2.04	2.12
Mean VIF	1.60	1.6	1.60	1.42	1.42	1.42
Year effect	YES	YES	YES	YES	YES	YES

Source(s): Authors' own work

Jucá *et al*., 2024; Nilipour *et al*., 2020), which might be seen as a cost rather than a performance driver (Makridou *et al*., 2024). Moreover, companies with high ESG scores are not immune to reputational issues arising from scandals involving inadequate social or governance practices and environmental mishaps (Liu, 2024; Xue *et al*., 2023). However, some studies argue that environmental innovation leads to long-term cost savings (Fiore *et al*., 2017) through improved production efficiency (Hojnik and Ruzzier, 2016), productivity (Arvanitis *et al*., 2017), and lower input costs. Doran and Ryan (2012) suggest that firms are willing to incur initial expenses for green technologies, following the same Hellström (2007) asserts that short-term sacrifices can lead to long-term cost reductions. Theoretically, the negative relationship between corporate performance and aggregate ESG performance of environmentally sensitive industries supports the shareholder theory (Barnea and Rubin, 2010; Friedman, 1970) and opposes the stakeholder theory (Freeman, 1984; Wang, 2024). ESG disclosures in environmentally sensitive corporations are driven by regulatory obligations and the potential to enhance corporate reputation, with key motivations including regulatory compliance and stakeholder pressures (Khamisu *et al*., 2024).

Most importantly, research findings contrast (Miralles-Quirós *et al*., 2018), who show a more significant positive valuation effect of CSR in environmentally sensitive industries by

Table 9. Summarized results

	Hypotheses	Results (SI)	Results (NSI)
H1	A significant relationship association exists between operating performance and its overall ESG performance	Insignificantly Negative	significantly Positive
H1a	A significant relationship association exists between operating performance and environmental performance	Insignificantly Negative	significantly Positive
H1b	A significant relationship association exists between operating performance and social performance	Insignificantly Positive	significantly Positive
H1c	A significant relationship association exists between operating performance and governance performance	Significantly Negative	significantly Positive
H2	A significant relationship association exists between financial performance and overall ESG performance	Insignificantly Negative	significantly Positive
H2a	A significant relationship association exists between financial performance and environmental performance	Insignificantly Negative	Insignificantly Positive
H2b	A significant relationship association exists between financial performance and social performance	Insignificantly Positive	Significantly Positive
H2c	A significant relationship association exists between financial performance and governance performance	Significantly Negative	Insignificantly Positive
H3	A significant relationship association exists between market performance and its overall ESG performance	insignificantly Positive	Significantly Positive
H3a	A significant relationship association exists between market performance and environmental performance	insignificantly Positive	Significantly Positive
H3b	A significant relationship association exists between market performance and social performance	insignificantly Positive	Significantly Positive
H3c	A significant relationship association exists between market performance and governance performance	insignificantly Positive	Insignificantly Positive

Source(s): Authors' own work

analyzing the sample of Brazilian firms. Furthermore, the negative effect of corporate governance practice on firm valuation in environmentally sensitive industries is well aligned with the studies of Yoon *et al*. (2018) based on Korean markets. Moreover, environmental, social, and governance dimensions have a distinctive impact.

Conversely, in NSI industries, ESG practices significantly enhance performance, highlighting their importance for long-term value creation and investor recognition. This result supports stakeholder theory (Freeman, 1984), which posits that strong ESG propositions help businesses attract and retain top talent, enhance employee motivation by fostering a sense of purpose, and boost overall productivity. The study findings indicate that implementing social responsibility and environmental practices may increase the return on assets, return on capital employed, and Tobin's Q in non-sensitive industries (Rahman *et al*., 2023). The varying effects of environmental, social, and governance (ESG) factors underscore the need for companies to develop customized sustainable strategies to see both environmentally sensitive and non-sensitive sectors achieve economic success and sustainability (Jucá *et al*., 2024). Furthermore, three core pillars of ESG also show a positive impact, resulting in those businesses that adopt ESG standards gaining advantages, including better corporate reputation with stakeholders and employees, higher returns on investments, competitive advantage over industry rivals, appeal to investors and lenders, improved financial performance, and increased customer loyalty (Maji and Lohia, 2023).

6.1 Contributions to theory, practice and policy

This paper makes significant theoretical contributions. First, to the authors' best knowledge, it is the first research to classify sectors according to environmental sensitivity and investigate ESG disclosure's influence on company performance in the Indian market. We explore the

listed firms rather than taking any indices that aid in replicating the findings for all companies in the securities basket. Secondly, the study's findings support the initial hypothesis that the impact of ESG and its dimensions on corporate performance has varied based on environmental sensitivity classification.

From a policy perspective, government and regulatory bodies are crucial in promoting clean technologies through carbon taxes and environmental mandates (Costa and Ferrão, 2010). Stringent ecological measures can facilitate the implementation of advanced green technologies (Farinelli *et al*., 2005) and enhance eco-technologies development, fostering optimization and innovation within the enterprise ecosystem (Jang *et al*., 2016). As clean technologies become more cost-effective over time (Chang, 2011), they help firms comply with stringent environmental regulations, offsetting initial acquisition and installation costs (Arvanitis *et al*., 2017). In this context, Ghisetti and Quatraro (2013) posit that even firms not directly affected by environmental regulation may invest in clean technologies to signal their commitment to social and ecological responsibilities and position themselves for future market opportunities. Additionally, polluting firms that integrate the latest eco-innovations will likely lead to research and development (R&D) for clean technologies in the long run. Hence, regulators should persist in supporting and strengthening ESG initiatives by offering firms guidance, assistance, and financial grants. Ultimately, compliance and disclosure of ESG benefits society as firms contribute value to the environment and community.

The work's economic implications encourage the advancement of ESG programs; the research provides insightful recommendations for businesses in both sectors, enhancing the body of knowledge on corporate sustainability. For instance, regarding ESG strategy implementations, corporations in SI industries should move beyond mere regulatory compliance by adopting proactive sustainability initiatives, such as developing carbon management strategies, investing in renewable energy, and pursuing eco-innovation to reduce further costs and risks. Additionally, transparency and investor trust may be increased by strengthening stakeholder involvement through specialized sustainability committees and coordinating ESG reporting with global frameworks like the Global Reporting Initiatives (GRI). NSI sectors should prioritize community involvement, diversity efforts, and employee-centric projects to improve their social reputations. Industry-wide ESG disclosure reliability may be increased, and digital advancements like AI-driven sustainability reporting systems can streamline disclosure.

The study offers key managerial implications, such as prioritizing ESG performance and improving disclosure, which can enhance financial success and market valuation relative to industry peers. In SI, strategic resource allocation is pivotal to balance immediate financial returns and long-term sustainability goals. Investing in areas with strong regulatory compliance potential despite high costs ensures long-term value. Furthermore, findings suggest that allocating resources to sustainable activities deepens stakeholder engagement and ensures crisis resilience. This research also offers key investors insights. For instance, NSI companies with superior ESG performance tend to exhibit better operating efficiency, financial performance, and market valuation; incorporating ESG performance into investment decisions can yield improved returns. Lastly, the findings emphasize the significance of implementing a robust legal framework and vigilant oversight to ensure effective ESG implementation. Governance streamlining is vital in balancing oversight with corporate outcomes, while reputation management underscores the necessity of ESG investment in building stakeholder trust and market competitiveness.

7. Conclusions and future research directions

Sustainable development initiatives are gradually reshaping the corporate landscapes, mitigating the climate-damaging effects of business activities and benefiting stakeholders. Firms worldwide face unprecedented hurdles due to social disruptions and climate change, necessitating regulatory and policy measures to encourage industries to adopt better practices. Prior research has predominantly focused on developed economies (Bassen *et al*., 2006; Cheng

et al., 2014; Dhaliwal *et al*., 2014). ESG practices within emerging nations remain relatively unexplored (Baughn *et al*., 2007; Dobers and Halme, 2009), despite their significance in global activities. This motivated us to conduct this study by employing system-GMM, offering timely and significant contribution by investigating the relationship between ESG disclosure and corporate performance across four different disclosures: (1) aggregate ESG disclosure, (2) Environmental disclosure, (3) Social disclosure and (4) Corporate governance disclosure. Based on data from 605 Indian-listed companies from 2015 to 2023, categorizing companies based on environmental sensitivity. The findings indicate that ESG factors affect corporate performance differently based on environmental sensitivity. While ESG disclosure in NSI firms significantly enhances profitability, market valuation, and long-term growth. It may translate differently for SI firms as compliance-driven ESG activities might not enhance profitability, given high governance costs and regulatory burdens (Yoon *et al*., 2018). Thus, customizing ESG strategies to industry-specific contexts is necessary to balance sustainability with profitability. It further underlines the reframing of the regulatory guidelines as per the environmental sensitivity of the industries (Appiagyei and Donkor, 2024). Moreover, larger firms with more resources to meet ESG criteria can leverage ESG standards for strategic growth and reduce regulatory pressure, highlighting the need for more rigorous ESG reporting frameworks in emerging nations than in developed countries (Chen *et al*., 2023a, b).

Academically, this study contributes to the literature on the ESG-corporate performance nexus within emerging nations, i.e. India. Examining outside developed nations, this study provides critical insights into the unique regulatory, cultural, and economic characteristics, offering findings that may be relevant to other emerging nations with comparable contexts. It is worth mentioning that our study reflects the snapshot of ESG performance over time, not capturing the evolving impact of ESG disclosure. Future research directions could include (1) investigating ESG performance in other developing nations where institutional frameworks, regulatory environments, and corporate governance structure differ to gain a comprehensive understanding of global ESG dynamics, (2) extending the time frame by leveraging newly available data to explore the evolving impact of ESG disclosures, enhancing the temporal scope of the analysis.; (3) further by conducting more granular sector-specific analyses to standardize better and understand the association between ESG and corporate outcomes.

Notes

1. India Overtakes UK as the World's Fifth-Largest Economy | World Economic Forum.
2. https://www.un.org/development/desa/dpad/publication/un-desa-policy-brief-no-153-india-overtakes-china-as-the-worlds-most-populous-country/
3. https://pib.gov.in/PressNoteDetails.aspx?NoteId=153357&ModuleId=3&reg=3&lang=1
4. https://www.lseg.com/content/dam/data-analytics/en_us/documents/methodology/lseg-esg-scores-methodology.pdf

References

Abdi, Y., Li, X. and Càmara-Turull, X. (2022), "Exploring the impact of sustainability (ESG) disclosure on firm value and financial performance (FP) in the airline industry: the moderating role of size and age", *Environment, Development and Sustainability*, Vol. 24 No. 4, pp. 5052-5079, doi: 10.1007/s10668-021-01649-w.

Adamkaite, J., Streimikiene, D. and udzioniene, K. (2023), "The impact of social responsibility on corporate financial performance in the energy sector: evidence from Lithuania", *Corporate Social Responsibility and Environmental Management*, Vol. 30 No. 1, pp. 91-104, doi: 10.1002/csr.2340.

Ademi, B. and Klungseth, N.J. (2022), "Does it pay to deliver superior ESG performance? Evidence from US S&P 500 companies", *Journal of Global Responsibility*, Vol. 13 No. 4, pp. 421-449, doi: 10.1108/JGR-01-2022-0006.

Agnese, P., Battaglia, F., Busato, F. and Taddeo, S. (2023), "ESG controversies and governance: evidence from the banking industry", *Finance Research Letters*, Vol. 53, 103397, doi: 10.1016/j.frl.2022.103397.

Agrawal, A. and Knoeber, C.R. (1996), "Firm performance and mechanisms to control agency problems between managers and shareholders", *The Journal of Financial and Quantitative Analysis*, Vol. 31 No. 3, p. 377, doi: 10.2307/2331397.

Alareeni, B.A. and Hamdan, A. (2020), "ESG impact on the performance of US S&P 500-listed firms", *Corporate Governance: The International Journal of Business in Society*, Vol. 20 No. 7, pp. 1409-1428, doi: 10.1108/CG-06-2020-0258.

Albitar, K., Hussainey, K., Kolade, N. and Gerged, A.M. (2020), "ESG disclosure and firm performance before and after IR", *International Journal of Accounting and Information Management*, Vol. 28 No. 3, pp. 429-444, doi: 10.1108/IJAIM-09-2019-0108.

Almulhim, A.A. and Aljughaiman, A.A. (2023), "Corporate sustainability and financial performance: the moderating effect of CEO characteristics", *Sustainability*, Vol. 15 No. 16, 12664, doi: 10.3390/su151612664.

Anklesaria-Dalal, K. and Thaker, N. (2019), "ESG and corporate financial performance: a panel study of Indian companies", Vol. 18, pp. 44-59.

Appiagyei, K. and Donkor, A. (2024), "Integrated reporting quality and sustainability performance: does firms' environmental sensitivity matter?", *Journal of Accounting in Emerging Economies*, Vol. 14 No. 1, pp. 25-47, doi: 10.1108/JAEE-02-2022-0058.

Aras, G., Tezcan, N. and Kutlu Furtuna, O. (2018), "The value relevance of banking sector multidimensional corporate sustainability performance", *Corporate Social Responsibility and Environmental Management*, Vol. 25 No. 6, pp. 1062-1073, doi: 10.1002/csr.1520.

Arellano, M. and Bond, S. (1991), "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations", *The Review of Economic Studies*, Vol. 58 No. 2, p. 277, doi: 10.2307/2297968.

Arellano, M. and Bover, O. (1995), "Another look at the instrumental variable estimation of error-components models", *Journal of Econometrics*, Vol. 68 No. 1, pp. 29-51, doi: 10.1016/0304-4076(94)01642-D.

Arian, A., Sands, J. and Tooley, S. (2023), "Industry and stakeholder impacts on corporate social responsibility (CSR) and financial performance: consumer vs. Industrial sectors", *Sustainability*, Vol. 15 No. 16, 12254, doi: 10.3390/su151612254.

Arianpoor, A. and Salehi, M. (2020), "A framework for business sustainability performance using meta-synthesis", *Management of Environmental Quality: An International Journal*, Vol. 32 No. 2, pp. 175-192, doi: 10.1108/MEQ-03-2020-0040.

Arianpoor, A., Salehi, M. and Daroudi, F. (2023), "Nonfinancial sustainability reporting, management legitimate authority and enterprise value", *Social Responsibility Journal*, Vol. 19 No. 10, pp. 1900-1916, doi: 10.1108/SRJ-09-2022-0374.

Arvanitis, S., Peneder, M., Rammer, C., Stucki, T. and Woerter, M. (2017), "Development and utilization of energy-related technologies, economic performance, and the role of policy instruments", *Journal of Cleaner Production*, Vol. 159, pp. 47-61, doi: 10.1016/j.jclepro.2017.04.162.

Atan, R., Alam, Md. M., Said, J. and Zamri, M. (2018), "The impacts of environmental, social, and governance factors on firm performance", *Management of Environmental Quality: An International Journal*, Vol. 29 No. 2, pp. 182-194, doi: 10.1108/MEQ-03-2017-0033.

Aupperle, K.E., Carroll, A.B. and Hatfield, J.D. (1985), "An empirical examination of the relationship between corporate social responsibility and profitability", *Academy of Management Journal*, Vol. 28 No. 2, pp. 446-463, doi: 10.2307/256210.

Ayayi, A.G. and Wijesiri, M. (2022), "Is there a trade-off between environmental performance and financial sustainability in microfinance institutions? Evidence from South and Southeast Asia", *Business Strategy and the Environment*, Vol. 31 No. 4, pp. 1552-1565, doi: 10.1002/bse.2969.

Bagherpasandi, M., Salehi, M., Hajiha, Z. and Hejazi, R. (2024a), "A model of increasing performance in sustainable supply chain management", *The TQM Journal*, Vol. 37 No. 3, pp. 653-679, doi: 10.1108/TQM-05-2023-0128.

Bagherpasandi, M., Salehi, M., Hajiha, Z. and Hejazi, R. (2024b), "Presenting a model for enhancing the performance of sustainable supply chain management using a data-driven approach", *Benchmarking: An International Journal*, doi: 10.1108/BIJ-12-2023-0846.

Bansal, M., Samad, T.A. and Bashir, H.A. (2021), "The sustainability reporting-firm performance nexus: evidence from a threshold model", *Journal of Global Responsibility*, Vol. 12 No. 4, pp. 491-512, doi: 10.1108/JGR-05-2021-0049.

Bansal, R., Wu, D.(A.) and Yaron, A. (2022), "Socially responsible investing in good and bad times", *The Review of Financial Studies*, Vol. 35 No. 4, pp. 2067-2099, doi: 10.1093/rfs/hhab072.

Barnea, A. and Rubin, A. (2010), "Corporate social responsibility as a conflict between shareholders", *Journal of Business Ethics*, Vol. 97 No. 1, pp. 71-86, doi: 10.1007/s10551-010-0496-z.

Bassen, A., Meyer, K. and Schlange, J. (2006), "The influence of corporate responsibility on the cost of capital", *SSRN Electronic Journal*, doi: 10.2139/ssrn.984406.

Baughn, C.C., Dusty) Bodie, N.L. and McIntosh, J.C. (2007), "Corporate social and environmental responsibility in Asian countries and other geographical regions", *Corporate Social Responsibility and Environmental Management*, Vol. 14 No. 4, pp. 189-205, doi: 10.1002/csr.160.

Behl, A., Kumari, P.S.R., Makhija, H. and Sharma, D. (2022), "Exploring the relationship of ESG score and firm value using cross-lagged panel analyses: case of the Indian energy sector", *Annals of Operations Research*, Vol. 313 No. 1, pp. 231-256, doi: 10.1007/s10479-021-04189-8.

Benbou, R. and Tirole, J. (2010), "Individual and corporate social responsibility", *Economica*, Vol. 77 No. 305, pp. 1-19, doi: 10.1111/j.1468-0335.2009.00843.x.

Bhandari, L.C. (1988), "Debt/equity ratio and expected common stock returns: empirical evidence", *The Journal of Finance*, Vol. 43 No. 2, pp. 507-528, doi: 10.1111/j.1540-6261.1988.tb03952.x.

Bissoondoyal-Bheenick, E., Brooks, R. and Do, H.X. (2023), "ESG and firm performance: the role of size and media channels", *Economic Modelling*, Vol. 121, 106203, doi: 10.1016/j.econmod.2023.106203.

Blanco, E., Rey-Maquieira, J. and Lozano, J. (2009), "The economic impacts of voluntary environmental performance of firms: a critical review", *Journal of Economic Surveys*, Vol. 23 No. 3, pp. 462-502, doi: 10.1111/j.1467-6419.2008.00569.x.

Blundell, R. and Bond, S. (1998), "Initial conditions and moment restrictions in dynamic panel data models", *Journal of Econometrics*, Vol. 87 No. 1, pp. 115-143, doi: 10.1016/S0304-4076(98)00009-8.

Boakye, D.J., Tingbani, I., Ahinful, G., Damoah, I. and Tauringana, V. (2020), "Sustainable environmental practices and financial performance: evidence from listed small and medium-sized enterprises in the United Kingdom", *Business Strategy and the Environment*, Vol. 29 No. 6, pp. 2583-2602, doi: 10.1002/bse.2522.

Boakye, D.J., Tingbani, I., Ahinful, G.S. and Nsor-Ambala, R. (2021), "The relationship between environmental management performance and financial performance of firms listed in the Alternative Investment Market (AIM) in the UK", *Journal of Cleaner Production*, Vol. 278, 124034, doi: 10.1016/j.jclepro.2020.124034.

Bocken, N.M.P. and Geradts, T.H.J. (2020), "Barriers and drivers to sustainable business model innovation: organization design and dynamic capabilities", *Long Range Planning*, Vol. 53 No. 4, 101950, doi: 10.1016/j.lrp.2019.101950.

Braam, G. and Peeters, R. (2018), "Corporate sustainability performance and assurance on sustainability reports: diffusion of accounting practices in the realm of sustainable development", *Corporate Social Responsibility and Environmental Management*, Vol. 25 No. 2, pp. 164-181, doi: 10.1002/csr.1447.

Brammer, S.J., Brooks, C. and Pavelin, S. (2005), "Corporate social performance and stock returns: UK evidence from disaggregate measures", *SSRN Electronic Journal*, doi: 10.2139/ssrn.739587.

Branco, M.C. and Rodrigues, L.L. (2008), "Factors influencing social responsibility disclosure by Portuguese companies", *Journal of Business Ethics*, Vol. 83 No. 4, pp. 685-701, doi: 10.1007/s10551-007-9658-z.

Brooks, C. and Oikonomou, I. (2018), "The effects of environmental, social and governance disclosures and performance on firm value: a review of the literature in accounting and finance", *The British Accounting Review*, Vol. 50 No. 1, pp. 1-15, doi: 10.1016/j.bar.2017.11.005.

Brown, T.J. (2006), "Identity, intended image, construed image, and reputation: an interdisciplinary framework and suggested terminology", *Journal of the Academy of Marketing Science*, Vol. 34 No. 2, pp. 99-106, doi: 10.1177/0092070305284969.

Bruna, M.G., Loprevite, S., Raucci, D., Ricca, B. and Rupo, D. (2022), "Investigating the marginal impact of ESG results on corporate financial performance", *Finance Research Letters*, Vol. 47, 102828, doi: 10.1016/j.frl.2022.102828.

Buallay, A. (2019), "Is sustainability reporting (ESG) associated with performance? Evidence from the European banking sector", *Management of Environmental Quality: An International Journal*, Vol. 30 No. 1, pp. 98-115, doi: 10.1108/MEQ-12-2017-0149.

Busch, T. and Hoffmann, V.H. (2011), "How hot is your bottom line? Linking carbon and financial performance", *Business and Society*, Vol. 50 No. 2, pp. 233-265, doi: 10.1177/0007650311398780.

Busch, T., Bauer, R. and Orlitzky, M. (2016), "Sustainable development and financial markets", *Business and Society*, Vol. 55 No. 3, pp. 303-329, doi: 10.1177/0007650315570701.

Cai, Y., Pan, C.H. and Statman, M. (2016), "Why do countries matter so much in corporate social performance?", *Journal of Corporate Finance*, Vol. 41, pp. 591-609, doi: 10.1016/j.jcorpfin.2016.09.004.

Campbell, J.L. (2007), "Why would corporations behave in socially responsible ways? An institutional theory of corporate social responsibility", *Academy of Management Review*, Vol. 32 No. 3, pp. 946-967, doi: 10.5465/amr.2007.25275684.

Carnini Pulino, S., Ciaburri, M., Magnanelli, B.S. and Nasta, L. (2022), "Does ESG disclosure influence firm performance?", *Sustainability*, Vol. 14 No. 13, p. 7595, doi: 10.3390/su14137595.

Cerciello, M., Busato, F. and Taddeo, S. (2023), "The effect of sustainable business practices on profitability. Accounting for strategic disclosure", *Corporate Social Responsibility and Environmental Management*, Vol. 30 No. 2, pp. 802-819, doi: 10.1002/csr.2389.

Chang, C. (2011), "Feeling ambivalent about going green", *Journal of Advertising*, Vol. 40 No. 4, pp. 19-32, doi: 10.2753/JOA0091-3367400402.

Chatterji, A.K., Levine, D.I. and Toffel, M.W. (2009), "How well do social ratings actually measure corporate social responsibility?", *Journal of Economics and Management Strategy*, Vol. 18 No. 1, pp. 125-169, doi: 10.1111/j.1530-9134.2009.00210.x.

Chelawat, H. and Trivedi, I.V. (2016), "The business value of ESG performance: the Indian context", *Asian Journal of Business Ethics*, Vol. 5 Nos 1-2, pp. 195-210, doi: 10.1007/s13520-016-0064-4.

Chen, Y. and Ibhagui, O.W. (2019), "R&D-firm performance nexus: new evidence from NASDAQ listed firms", *The North American Journal of Economics and Finance*, Vol. 50, 101009, doi: 10.1016/j.najef.2019.101009.

Chen, R. and Zhang, T. (2024), "Artificial intelligence applications implication for ESG performance: can digital transformation of enterprises promote sustainable development?", *Chinese Management Studies*, doi: 10.1108/CMS-11-2023-0653.

Chen, Y., Li, T., Zeng, Q. and Zhu, B. (2023a), "Effect of ESG performance on the cost of equity capital: evidence from China", *International Review of Economics and Finance*, Vol. 83, pp. 348-364, doi: 10.1016/j.iref.2022.09.001.

Chen, S., Song, Y. and Gao, P. (2023b), "Environmental, social, and governance (ESG) performance and financial outcomes: analyzing the impact of ESG on financial performance", *Journal of Environmental Management*, Vol. 345, 118829, doi: 10.1016/j.jenvman.2023.118829.

Cheng, B., Ioannou, I. and Serafeim, G. (2014), "Corporate social responsibility and access to finance", *Strategic Management Journal*, Vol. 35 No. 1, pp. 1-23, doi: 10.1002/smj.2131.

Cherian, J., Umar, M., Thu, P., Nguyen-Trang, T., Sial, M. and Khuong, N. (2019), "Does corporate social responsibility affect the financial performance of the manufacturing sector? Evidence from an emerging economy", *Sustainability*, Vol. 11 No. 4, p. 1182, doi: 10.3390/su11041182.

Cho, C.H. and Patten, D.M. (2007), "The role of environmental disclosures as tools of legitimacy: a research note", *Accounting, Organizations and Society*, Vol. 32 Nos 7-8, pp. 639-647, doi: 10.1016/j.aos.2006.09.009.

Cho, C.H., Michelon, G., Patten, D.M. and Roberts, R.W. (2014), "CSR report assurance in the USA: an empirical investigation of determinants and effects", *Sustainability Accounting, Management and Policy Journal*, Vol. 5 No. 2, pp. 130-148, doi: 10.1108/SAMPJ-01-2014-0003.

Choi, J. and Wang, H. (2009), "Stakeholder relations and the persistence of corporate financial performance", *Strategic Management Journal*, Vol. 30 No. 8, pp. 895-907, doi: 10.1002/smj.759.

Chouaibi, S., Rossi, M., Siggia, D. and Chouaibi, J. (2021), "Exploring the moderating role of social and ethical practices in the relationship between environmental disclosure and financial performance: evidence from ESG companies", *Sustainability*, Vol. 14 No. 1, p. 209, doi: 10.3390/su14010209.

Chung, R., Bayne, L. and Birt, J. (2024), "The impact of environmental, social and governance (ESG) disclosure on firm financial performance: evidence from Hong Kong", *Asian Review of Accounting*, Vol. 32 No. 1, pp. 136-165, doi: 10.1108/ARA-07-2022-0165.

Cohen, M., Fenn, S. and Konar, S. (1997), "Environmental and financial performance: are they related?".

Constantinescu, D., Caraiani, C., Lungua, C.I. and Mititean, P. (2021), "Environmental, social and governance disclosure associated with the firm value. Evidence from energy industry", *Journal of Accounting and Management Information Systems*, pp. 1-2021, doi: 10.24818/jamis.2021.01003.

Costa, I. and Ferrão, P. (2010), "A case study of industrial symbiosis development using a middle-out approach", *Journal of Cleaner Production*, Vol. 18 Nos 10-11, pp. 984-992, doi: 10.1016/j.jclepro.2010.03.007.

Cuervo, A. (2002), "Corporate Governance Mechanisms: a plea for less code of good governance and more market control", *Corporate Governance: An International Review*, Vol. 10 No. 2, pp. 84-93, doi: 10.1111/1467-8683.00272.

Danso, A., Adomako, S., Amankwah-Amoah, J., Owusu-Agyei, S. and Konadu, R. (2019), "Environmental sustainability orientation, competitive strategy and financial performance", *Business Strategy and the Environment*, Vol. 28 No. 5, pp. 885-895, doi: 10.1002/bse.2291.

Davis, J.H., Schoorman, F.D. and Donaldson, L. (1997), "Toward a stewardship theory of management", *The Academy of Management Review*, Vol. 22 No. 1, p. 20, doi: 10.2307/259223.

Dhaliwal, D., Li, O.Z., Tsang, A. and Yang, Y.G. (2014), "Corporate social responsibility disclosure and the cost of equity capital: the roles of stakeholder orientation and financial transparency", *Journal of Accounting and Public Policy*, Vol. 33 No. 4, pp. 328-355, doi: 10.1016/j.jaccpubpol.2014.04.006.

Dobers, P. and Halme, M. (2009), "Corporate social responsibility and developing countries", *Corporate Social Responsibility and Environmental Management*, Vol. 16 No. 5, pp. 237-249, doi: 10.1002/csr.212.

Donaldson, T. and Preston, L.E. (1995), "The stakeholder theory of the corporation: concepts, evidence, and implications", *The Academy of Management Review*, Vol. 20 No. 1, p. 65, doi: 10.2307/258887.

Dong, S., Burritt, R. and Qian, W. (2014), "Salient stakeholders in corporate social responsibility reporting by Chinese mining and minerals companies", *Journal of Cleaner Production*, Vol. 84, pp. 59-69, doi: 10.1016/j.jclepro.2014.01.012.

Dong, Y., Liang, C. and Wanyin, Z. (2023), "Board diversity and firm performance: impact of ESG activities in China", *Economic Research-Ekonomska Istraživanja*, Vol. 36 No. 1, pp. 1592-1609, doi: 10.1080/1331677X.2022.2090406.

Doran, J. and Ryan, G. (2012), "Regulation and firm perception, eco-innovation and firm performance", *European Journal of Innovation Management*, Vol. 15 No. 4, pp. 421-441, doi: 10.1108/14601061211272367.

Doshi, M., Jain, R., Sharma, D., Mukherjee, D. and Kumar, S. (2024), "Does ownership influence ESG disclosure scores?", *Research in International Business and Finance*, Vol. 67, 102122, doi: 10.1016/j.ribaf.2023.102122.

Dwibedi, P., Pahi, D. and Sahu, A. (2024), "Unveiling the relationship between ESG scores and firm performance in India: a system GMM approach", *Australasian Business, Accounting and Finance Journal*, Vol. 18 No. 3, pp. 26-49, doi: 10.14453/aabfj.v18i3.03.

Eccles, R. and Serafeim, G. (2013), "The performance frontier: innovating for a sustainable strategy", *Harvard Business Review*, Vol. 91 Nos 50-56, pp. 50-150, 58, 60, 150.

Eccles, R.G., Ioannou, I. and Serafeim, G. (2014), "The impact of corporate sustainability on organizational processes and performance", *Management Science*, Vol. 60 No. 11, pp. 2835-2857, doi: 10.1287/mnsc.2014.1984.

Farinelli, U., Johansson, T.B., McCormick, K., Mundaca, L., Oikonomou, V., Örtenvik, M., Patel, M. and Santi, F. (2005), "'White and Green': comparison of market-based instruments to promote energy efficiency", *Journal of Cleaner Production*, Vol. 13 Nos 10-11, pp. 1015-1026, doi: 10.1016/j.jclepro.2004.12.013.

Fijałkowska, J., Zyznarska-Dworczak, B. and Garsztka, P. (2018), "Corporate social-environmental performance versus financial performance of banks in Central and Eastern European countries", *Sustainability*, Vol. 10 No. 3, p. 772, doi: 10.3390/su10030772.

Fiore, M., Silvestri, R., Contò, F. and Pellegrini, G. (2017), "Understanding the relationship between green approach and marketing innovations tools in the wine sector", *Journal of Cleaner Production*, Vol. 142, pp. 4085-4091, doi: 10.1016/j.jclepro.2016.10.026.

Flammer, C. (2015), "Does corporate social responsibility lead to superior financial performance? A regression discontinuity approach", *Management Science*, Vol. 61 No. 11, pp. 2549-2568, doi: 10.1287/mnsc.2014.2038.

Fosu, E., Yi, K. and Asiedu, D. (2024), "The effect of CSR on corporate social performance: mediating role of corporate image, green innovation and moderating role of corporate identity", *Corporate Social Responsibility and Environmental Management*, Vol. 31 No. 1, pp. 69-88, doi: 10.1002/csr.2553.

Franco, S., Caroli, M.G., Cappa, F. and Del Chiappa, G. (2020), "Are you good enough? CSR, quality management and corporate financial performance in the hospitality industry", *International Journal of Hospitality Management*, Vol. 88, 102395, doi: 10.1016/j.ijhm.2019.102395.

Freeman, R.E. (1984), *Strategic Management: A Stakeholder Approach*, Pitman.

Freeman, R.E. (2010), *Strategic Management*, Cambridge University Press, doi: 10.1017/CBO9781139192675.

Freeman, R.E., Phillips, R. and Sisodia, R. (2020), "Tensions in stakeholder theory", *Business and Society*, Vol. 59 No. 2, pp. 213-231, doi: 10.1177/0007650318773750.

Friede, G., Busch, T. and Bassen, A. (2015), "ESG and financial performance: aggregated evidence from more than 2000 empirical studies", *Journal of Sustainable Finance and Investment*, Vol. 5 No. 4, pp. 210-233, doi: 10.1080/20430795.2015.1118917.

Friedman, M. (1970), "The social responsibility of business is to increase its profits", *New York Times Magazine*, pp. 122-126.

Garcia, A.S., Mendes-Da-Silva, W. and Orsato, R.J. (2017), "Sensitive industries produce better ESG performance: evidence from emerging markets", *Journal of Cleaner Production*, Vol. 150, pp. 135-147, doi: 10.1016/j.jclepro.2017.02.180.

Garcia-Castro, R., Ariño, M.A. and Canela, M.A. (2010), "Does social performance really lead to financial performance? Accounting for endogeneity", *Journal of Business Ethics*, Vol. 92 No. 1, pp. 107-126, doi: 10.1007/s10551-009-0143-8.

García-Sánchez, I. and Prado-Lorenzo, J. (2012), "Greenhouse gas emission practices and financial performance", *International Journal of Climate Change Strategies and Management*, Vol. 4 No. 3, pp. 260-276, doi: 10.1108/17568691211248720.

Gartia, U. and Panda, A.K. (2024), "Do the firm characteristics moderate the nexus between the firm's sustainable practices and financial performance?", *Business Strategy and Development*, Vol. 7 No. 2, doi: 10.1002/bsd2.376.

Ge, Y., Knittel, C., MacKenzie, D. and Zoepf, S. (2016), "Racial and gender discrimination in transportation network companies", doi: 10.3386/w22776.

Ghisetti, C. and Quatraro, F. (2013), "Beyond inducement in climate change: does environmental performance spur environmental technologies? A regional analysis of cross-sectoral differences", *Ecological Economics*, Vol. 96, pp. 99-113, doi: 10.1016/j.ecolecon.2013.10.004.

Giannopoulos, G., Kihle Fagernes, R.V., Elmarzouky, M. and Afzal Hossain, K.A.B.M. (2022), "The ESG disclosure and the financial performance of Norwegian listed firms", *Journal of Risk and Financial Management*, Vol. 15 No. 6, p. 237, doi: 10.3390/jrfm15060237.

Gigante, G. and Manglaviti, D. (2022), "The ESG effect on the cost of debt financing: a sharp RD analysis", *International Review of Financial Analysis*, Vol. 84, 102382, doi: 10.1016/j.irfa.2022.102382.

Gillet-Monjarret, C. (2015), "Assurance of sustainability information: a study of media pressure", *Accounting in Europe*, Vol. 12 No. 1, pp. 87-105, doi: 10.1080/17449480.2015.1036894.

Gracia, O. and Siregar, S.V. (2021), "Sustainability practices and the cost of debt: evidence from ASEAN countries", *Journal of Cleaner Production*, Vol. 300, 126942, doi: 10.1016/j.jclepro.2021.126942.

Graves, S.B. and Waddock, S.A. (2000), "Beyond built to last . Stakeholder relations in 'built-to-last' companies", *Business and Society Review*, Vol. 105 No. 4, pp. 393-418, doi: 10.1111/0045-3609.00090.

Grossman, S.J. and Hart, O.D. (1986), "The costs and benefits of ownership: a theory of vertical and lateral integration", *Journal of Political Economy*, Vol. 94 No. 4, pp. 691-719, doi: 10.1086/261404.

Guenther, E.M. and Hoppe, H. (2014), "Merging limited perspectives", *Journal of Industrial Ecology*, Vol. 18 No. 5, pp. 689-707, doi: 10.1111/jiec.12125.

Guo, H. and Lu, W. (2021), "The inverse U-shaped relationship between corporate social responsibility and competitiveness: evidence from Chinese international construction companies", *Journal of Cleaner Production*, Vol. 295, 126374, doi: 10.1016/j.jclepro.2021.126374.

Gupta, S. and Aggarwal, D. (2024), "Shrinking the capital costs and beta risk impediments through ESG: study of an emerging market", *Asian Review of Accounting*, Vol. 32 No. 2, pp. 249-277, doi: 10.1108/ARA-05-2023-0130.

Gupta, A., Sharma, U. and Gupta, S.K. (2021), "The role of ESG in sustainable development: an analysis through the lens of machine learning", *2021 IEEE International Humanitarian Technology Conference (IHTC)*, pp. 1-5, doi: 10.1109/IHTC53077.2021.9698939.

Hahn, R. and Lülfs, R. (2013), "Legitimizing negative aspects in GRI-oriented sustainability reporting: a qualitative study", *Academy of Management Proceedings*, Vol. 2013 No. 1, 10658, doi: 10.5465/ambpp.2013.10658abstract.

Halkos, G. and Skouloudis, A. (2016), "Exploring the current status and key determinants of corporate disclosure on climate change: evidence from the Greek business sector", *Environmental Science and Policy*, Vol. 56, pp. 22-31, doi: 10.1016/j.envsci.2015.10.011.

Hamrouni, A., Boussaada, R. and Ben Farhat Toumi, N. (2019), "Corporate social responsibility disclosure and debt financing", *Journal of Applied Accounting Research*, Vol. 20 No. 4, pp. 394-415, doi: 10.1108/JAAR-01-2018-0020.

Han, S., Pan, Y., Mygrant, M. and Li, M. (2021), "Differentiated environmental regulations and corporate environmental responsibility: the moderating role of institutional environment", *Journal of Cleaner Production*, Vol. 313, 127870, doi: 10.1016/j.jclepro.2021.127870.

Hang, M., Geyer-Klingeberg, J. and Rathgeber, A.W. (2019), "It is merely a matter of time: a meta-analysis of the causality between environmental performance and financial performance", *Business Strategy and the Environment*, Vol. 28 No. 2, pp. 257-273, doi: 10.1002/bse.2215.

Hart, S.L. (1995), "A natural-resource-based view of the firm", *The Academy of Management Review*, Vol. 20 No. 4, p. 986, doi: 10.2307/258963.

Hart, S.L. and Dowell, G. (2011), "Invited editorial: a natural-resource-based view of the firm", *Journal of Management*, Vol. 37 No. 5, pp. 1464-1479, doi: 10.1177/0149206310390219.

Hasan, R., Miah, M.D. and Hassan, M.K. (2022), "The nexus between environmental and financial performance: evidence from gulf cooperative council banks", *Business Strategy and the Environment*, Vol. 31 No. 7, pp. 2882-2907, doi: 10.1002/bse.3053.

Hellström, T. (2007), "Dimensions of environmentally sustainable innovation: the structure of eco-innovation concepts", *Sustainable Development*, Vol. 15 No. 3, pp. 148-159, doi: 10.1002/sd.309.

Hillman, A.J., Cannella, A.A. and Paetzold, R.L. (2000), "The resource dependence role of corporate directors: strategic adaptation of board composition in response to environmental change", *Journal of Management Studies*, Vol. 37 No. 2, pp. 235-256, doi: 10.1111/1467-6486.00179.

Hojnik, J. and Ruzzier, M. (2016), "The driving forces of process eco-innovation and its impact on performance: insights from Slovenia", *Journal of Cleaner Production*, Vol. 133, pp. 812-825, doi: 10.1016/j.jclepro.2016.06.002.

Horobet, A., Bulai, V., Radulescu, M., Belascu, L. and Dumitrescu, D.G. (2024), "ESG actions, corporate discourse, and market assessment nexus: evidence from the oil and gas sector", *Journal of Business Economics and Management*, Vol. 25 No. 1, pp. 153-174, doi: 10.3846/jbem.2024.21070.

Horváthová, E. (2012), "The impact of environmental performance on firm performance: short-term costs and long-term benefits?", *Ecological Economics*, Vol. 84, pp. 91-97, doi: 10.1016/j.ecolecon.2012.10.001.

Hsiao, H.-F., Zhong, T. and Wang, J. (2024), "Does national culture influence corporate social responsibility on firm performance?", *Humanities and Social Sciences Communications*, Vol. 11 No. 1, p. 5, doi: 10.1057/s41599-023-02538-5.

Huang, J., Li, Y. and Han, F. (2024), "Walk well and talk well: impact of the consistency of ESG performance and disclosure on firms' stock price crash risk", *International Review of Economics and Finance*, Vol. 93, pp. 1154-1174, doi: 10.1016/j.iref.2024.03.067.

Hussaini, M., Hussain, N., Nguyen, D.K. and Rigoni, U. (2021), "Is corporate social responsibility an agency problem? An empirical note from takeovers", *Finance Research Letters*, Vol. 43, 102007, doi: 10.1016/j.frl.2021.102007.

Iatridis, G.E. (2013), "Environmental disclosure quality: evidence on environmental performance, corporate governance and value relevance", *Emerging Markets Review*, Vol. 14, pp. 55-75, doi: 10.1016/j.ememar.2012.11.003.

Jaeger, B. and Upadhyay, A. (2020), "Understanding barriers to circular economy: cases from the manufacturing industry", *Journal of Enterprise Information Management*, Vol. 33 No. 4, pp. 729-745, doi: 10.1108/JEIM-02-2019-0047.

Jain, M., Sharma, G.D. and Srivastava, M. (2019), "Can sustainable investment yield better financial returns: a comparative study of ESG indices and MSCI indices", *Risks*, Vol. 7 No. 1, p. 15, doi: 10.3390/risks7010015.

Jang, J., McSparren, J. and Rashchupkina, Y. (2016), "Global governance: present and future", *Palgrave Communications*, Vol. 2 No. 1, 15045, doi: 10.1057/palcomms.2015.45.

Jensen, M.C. and Meckling, W.H. (1976), "Theory of the firm: managerial behavior, agency costs and ownership structure", *Journal of Financial Economics*, Vol. 3 No. 4, pp. 305-360, doi: 10.1016/0304-405X(76)90026-X.

Jha, M.K. and Rangarajan, K. (2020), "Analysis of corporate sustainability performance and corporate financial performance causal linkage in the Indian context", *Asian Journal of Sustainability and Social Responsibility*, Vol. 5 No. 1, p. 10, doi: 10.1186/s41180-020-00038-z.

Jo, H., Kim, H. and Park, K. (2015), "Corporate environmental responsibility and firm performance in the financial services sector", *Journal of Business Ethics*, Vol. 131 No. 2, pp. 257-284, doi: 10.1007/s10551-014-2276-7.

Jucá, M.N., Muren, P.D., Valentinčič, A. and Ichev, R. (2024), "The impact of ESG controversies on the financial performance of firms: an analysis of industry and country clusters", *Borsa Istanbul Review*, Vol. 24 No. 6, pp. 1305-1315, doi: 10.1016/j.bir.2024.08.001.

Jyoti, G. and Khanna, A. (2021), "Does sustainability performance impact financial performance? Evidence from Indian service sector firms", *Sustainable Development*, Vol. 29 No. 6, pp. 1086-1095, doi: 10.1002/sd.2204.

Kahloul, I., Sbai, H. and Grira, J. (2022), "Does Corporate Social Responsibility reporting improve financial performance? The moderating role of board diversity and gender composition", *The Quarterly Review of Economics and Finance*, Vol. 84, pp. 305-314, doi: 10.1016/j.qref.2022.03.001.

Kennedy, P. (2008), *A Guide to Econometrics*, Wiley.

Khamisu, M.S., Paluri, R.A. and Sonwaney, V. (2024), "Environmental social and governance (ESG) disclosure motives for environmentally sensitive industry: an emerging economy perspective", *Cogent Business and Management*, Vol. 11 No. 1, doi: 10.1080/23311975.2024.2322027.

Khan, M.A. (2022), "ESG disclosure and Firm performance: a bibliometric and meta analysis", *Research in International Business and Finance*, Vol. 61, 101668, doi: 10.1016/j.ribaf.2022.101668.

Khandelwal, V., Sharma, P. and Chotia, V. (2023), "ESG disclosure and firm performance: an asset-pricing approach", *Risks*, Vol. 11 No. 6, p. 112, doi: 10.3390/risks11060112.

Kim, Y. and Statman, M. (2012), "Do corporations invest enough in environmental responsibility?", *Journal of Business Ethics*, Vol. 105 No. 1, pp. 115-129, doi: 10.1007/s10551-011-0954-2.

Kitchen, J. (2020), "A stakeholder theory perspective on business models: value creation for sustainability", *Social and Environmental Accountability Journal*, Vol. 40 No. 1, p. 98, doi: 10.1080/0969160X.2020.1711334.

Kolk, A. and Perego, P. (2010), "Determinants of the adoption of sustainability assurance statements: an international investigation", *Business Strategy and the Environment*, Vol. 19 No. 3, pp. 182-198, doi: 10.1002/bse.643.

Krüger, P. (2015), "Corporate goodness and shareholder wealth", *Journal of Financial Economics*, Vol. 115 No. 2, pp. 304-329, doi: 10.1016/j.jfineco.2014.09.008.

Kuo, L. and Chang, B.-G. (2021), "The affecting factors of circular economy information and its impact on corporate economic sustainability-Evidence from China", *Sustainable Production and Consumption*, Vol. 27, pp. 986-997, doi: 10.1016/j.spc.2021.02.014.

Lahouel, B.B., Zaied, Y.B., Song, Y. and Yang, G. (2021), "Corporate social performance and financial performance relationship: a data envelopment analysis approach without explicit input", *Finance Research Letters*, Vol. 39, 101656, doi: 10.1016/j.frl.2020.101656.

Lambertini, L. and Tampieri, A. (2015), "Incentives, performance and desirability of socially responsible firms in a Cournot oligopoly", *Economic Modelling*, Vol. 50, pp. 40-48, doi: 10.1016/j.econmod.2015.05.016.

Lee, M.T. and Raschke, R.L. (2023), "Stakeholder legitimacy in firm greening and financial performance: what about greenwashing temptations?☆", *Journal of Business Research*, Vol. 155, 113393, doi: 10.1016/j.jbusres.2022.113393.

Lee, K., Cin, B.C. and Lee, E.Y. (2016), "Environmental responsibility and firm performance: the application of an environmental, social and governance model", *Business Strategy and the Environment*, Vol. 25 No. 1, pp. 40-53, doi: 10.1002/bse.1855.

Lewandowski, S. (2017), "Corporate carbon and financial performance: the role of emission reductions", *Business Strategy and the Environment*, Vol. 26 No. 8, pp. 1196-1211, doi: 10.1002/bse.1978.

Lin, W.-L., Cheah, J.-H., Azali, M., Ho, J.A. and Yip, N. (2019), "Does firm size matter? Evidence on the impact of the green innovation strategy on corporate financial performance in the automotive sector", *Journal of Cleaner Production*, Vol. 229, pp. 974-988, doi: 10.1016/j.jclepro.2019.04.214.

Lioui, A. and Sharma, Z. (2012), "Environmental corporate social responsibility and financial performance: disentangling direct and indirect effects", *Ecological Economics*, Vol. 78, pp. 100-111, doi: 10.1016/j.ecolecon.2012.04.004.

Liu, L. (2023), "Green innovation, firm performance, and risk mitigation: evidence from the USA", *Environment, Development and Sustainability*, Vol. 26 No. 9, pp. 24009-24030, doi: 10.1007/s10668-023-03632-z.

Liu, L. (2024), "Environmental performance factors: insights from CSR-linked compensation, committees, disclosure, targets, and board composition", *Journal of Sustainable Finance and Investment*, pp. 1-36, doi: 10.1080/20430795.2024.2313497.

Lu, S. and Cheng, B. (2023), "Does environmental regulation affect firms' ESG performance? Evidence from China", *Managerial and Decision Economics*, Vol. 44 No. 4, pp. 2004-2009, doi: 10.1002/mde.3796.

Maji, S.G. and Lohia, P. (2023), "Environmental, social and governance (ESG) performance and firm performance in India", *Society and Business Review*, Vol. 18 No. 1, pp. 175-194, doi: 10.1108/SBR-06-2022-0162.

Maji, S.G. and Lohia, P. (2024), "Unveiling the financial effect of ESG disclosure on financial performance in India: climate-sensitive corporates' perspective", *International Journal of Ethics and Systems*, doi: 10.1108/IJOES-02-2024-0039.

Makridou, G., Doumpos, M. and Lemonakis, C. (2024), "Relationship between ESG and corporate financial performance in the energy sector: empirical evidence from European companies", *International Journal of Energy Sector Management*, Vol. 18 No. 4, pp. 873-895, doi: 10.1108/IJESM-01-2023-0012.

Malik, N. and Kashiramka, S. (2024), "Impact of ESG disclosure on firm performance and cost of debt: empirical evidence from India", *Journal of Cleaner Production*, Vol. 448, 141582, doi: 10.1016/j.jclepro.2024.141582.

Manasakis, C., Mitrokostas, E. and Petrakis, E. (2014), "Strategic corporate social responsibility activities and corporate governance in imperfectly competitive markets", *Managerial and Decision Economics*, Vol. 35 No. 7, pp. 460-473, doi: 10.1002/mde.2635.

Manrique, S. and Martí-Ballester, C.-P. (2017), "Analyzing the effect of corporate environmental performance on corporate financial performance in developed and developing countries", *Sustainability*, Vol. 9 No. 11, p. 1957, doi: 10.3390/su9111957.

Maqbool, S. and Bakr, A. (2019), "The curvilinear relationship between corporate social performance and financial performance", *Journal of Global Responsibility*, Vol. 10 No. 1, pp. 87-100, doi: 10.1108/JGR-11-2018-0060.

Martínez-Ferrero, J. and Frías-Aceituno, J.V. (2015), "Relationship between sustainable development and financial performance: international empirical research", *Business Strategy and the Environment*, Vol. 24 No. 1, pp. 20-39, doi: 10.1002/bse.1803.

Masongweni, V.V. and Simo-Kengne, B.D. (2024), "The impact of sustainable investment on firm performance in South Africa", *South African Journal of Accounting Research*, Vol. 38 No. 2, pp. 146-173, doi: 10.1080/10291954.2024.2306701.

Masulis, R.W. and Reza, S.W. (2013), "Agency problems of corporate philanthropy", *SSRN Electronic Journal*, doi: 10.2139/ssrn.2234221.

McWilliams, A. and Siegel, D. (2000), "Corporate social responsibility and financial performance: correlation or misspecification?", *Strategic Management Journal*, Vol. 21 No. 5, pp. 603-609, doi: 10.1002/(SICI)1097-0266(200005)21:5<603::AID-SMJ101>3.0.CO;2-3.

Mendiratta, A., Singh, S., Yadav, S.S. and Mahajan, A. (2023), "ESG controversies and firm performance in India: the moderating impact of government effectiveness", *Global Business Review*, doi: 10.1177/09721509231151490.

Mhlanga, D. and Moloi, T. (2020), "The stakeholder theory in the Fourth industrial revolution", *International Journal of Economics and Finance Studies*, Vol. 12, pp. 1309-8055, doi: 10.34109/ijefs.202012207.

Miralles-Quirós, M., Miralles-Quirós, J. and Valente Gonçalves, L. (2018), "The value relevance of environmental, social, and governance performance: the Brazilian case", *Sustainability*, Vol. 10 No. 3, p. 574, doi: 10.3390/su10030574.

Mishra, G., Patro, A. and Tiwari, A.K. (2024), "Does climate governance moderate the relationship between ESG reporting and firm value? Empirical evidence from India", *International Review of Economics and Finance*, Vol. 91, pp. 920-941, doi: 10.1016/j.iref.2024.01.059.

Mittal, R.K., Sinha, N. and Singh, A. (2008), "An analysis of linkage between economic value added and corporate social responsibility", *Management Decision*, Vol. 46 No. 9, pp. 1437-1443, doi: 10.1108/00251740810912037.

Mohammad, W.M.W. and Wasiuzzaman, S. (2021), "Environmental, Social and Governance (ESG) disclosure, competitive advantage and performance of firms in Malaysia", *Cleaner Environmental Systems*, Vol. 2, 100015, doi: 10.1016/j.cesys.2021.100015.

Molina-Azorín, J.F., Claver-Cortés, E., López-Gamero, M.D. and Tarí, J.J. (2009), "Green management and financial performance: a literature review", *Management Decision*, Vol. 47 No. 7, pp. 1080-1100, doi: 10.1108/00251740910978313.

Muth, M. and Donaldson, L. (1998), "Stewardship theory and board structure: a contingency approach", *Corporate Governance: An International Review*, Vol. 6 No. 1, pp. 5-28, doi: 10.1111/1467-8683.00076.

Naeem, N., Cankaya, S. and Bildik, R. (2022), "Does ESG performance affect the financial performance of environmentally sensitive industries? A comparison between emerging and developed markets", *Borsa Istanbul Review*, Vol. 22, pp. S128-S140, doi: 10.1016/j.bir.2022.11.014.

Nakao, Y., Amano, A., Matsumura, K., Genba, K. and Nakano, M. (2007), "Relationship between environmental performance and financial performance: an empirical analysis of Japanese corporations", *Business Strategy and the Environment*, Vol. 16 No. 2, pp. 106-118, doi: 10.1002/bse.476.

Narula, K. (2012), "'Sustainable investing' via the FDI route for sustainable development", *Procedia - Social and Behavioral Sciences*, Vol. 37, pp. 15-30, doi: 10.1016/j.sbspro.2012.03.271.

Narula, R., Rao, P., Kumar, S. and Matta, R. (2024), "ESG scores and firm performance- evidence from emerging market", *International Review of Economics and Finance*, Vol. 89, pp. 1170-1184, doi: 10.1016/j.iref.2023.08.024.

Nenavani, J., Prasuna, A., Siva Kumar, S.N.V. and Kasturi, A. (2024), "ESG measures and financial performance of logistics companies", *Letters in Spatial and Resource Sciences*, Vol. 17 No. 1, p. 5, doi: 10.1007/s12076-023-00358-4.

Nguyen, L.-T., Doan, A.N.P. and Frömmel, M. (2021), "Boards of directors and corporate sustainability performance: evidence from the emerging East Asian markets", *International Journal of Disclosure and Governance*, Vol. 18 No. 2, pp. 95-105, doi: 10.1057/s41310-020-00102-0.

Nie, C., Luo, W., Chen, Z. and Feng, Y. (2024), "Intellectual property protection and corporate ESG performance: evidence from a quasi-natural experiment in China", *Business Process Management Journal*, Vol. 31 No. 1, pp. 245-266, doi: 10.1108/BPMJ-01-2024-0041.

Nilipour, A., de Silva, T.-A. and Li, X. (2020), "The readability of sustainability reporting in New Zealand over time", *Australasian Accounting, Business and Finance Journal*, Vol. 14 No. 3, pp. 86-107, doi: 10.14453/aabfj.v14i3.7.

Nirino, N., Santoro, G., Miglietta, N. and Quaglia, R. (2021), "Corporate controversies and company's financial performance: exploring the moderating role of ESG practices", *Technological Forecasting and Social Change*, Vol. 162, 120341, doi: 10.1016/j.techfore.2020.120341.

Nollet, J., Filis, G. and Mitrokostas, E. (2016), "Corporate social responsibility and financial performance: a non-linear and disaggregated approach", *Economic Modelling*, Vol. 52, pp. 400-407, doi: 10.1016/j.econmod.2015.09.019.

Omar, B.F. and Zallom, N.O. (2016), "Corporate social responsibility and market value: evidence from Jordan", *Journal of Financial Reporting and Accounting*, Vol. 14 No. 1, pp. 2-29, doi: 10.1108/JFRA-11-2014-0084.

Orlitzky, M. and Shen, J. (2013), "Corporate social responsibility, industry, and strategy", *Industrial and Organizational Psychology*, Vol. 6 No. 4, pp. 346-350, doi: 10.1111/iops.12064.

Orlitzky, M., Schmidt, F.L. and Rynes, S.L. (2003), "Corporate social and financial performance: a meta-analysis", *Organization Studies*, Vol. 24 No. 3, pp. 403-441, doi: 10.1177/0170840603024003910.

Ortiz-de-Mandojana, N. and Aragon-Correa, J.A. (2015), "Boards and sustainability: the contingent influence of director interlocks on corporate environmental performance", *Business Strategy and the Environment*, Vol. 24 No. 6, pp. 499-517, doi: 10.1002/bse.1833.

Palmate, S.S., Pandey, A., Pandey, R.P. and Mishra, S.K. (2021), "Assessing the land degradation and greening response to changes in hydro-climatic variables using a conceptual framework: a case-study in central India", *Land Degradation and Development*, Vol. 32 No. 14, pp. 4132-4148, doi: 10.1002/ldr.4014.

Paniagua, J., Rivelles, R. and Sapena, J. (2018), "Corporate governance and financial performance: the role of ownership and board structure", *Journal of Business Research*, Vol. 89, pp. 229-234, doi: 10.1016/j.jbusres.2018.01.060.

Pätäri, S., Jantunen, A., Kyläheiko, K. and Sandström, J. (2012), "Does sustainable development foster value creation? Empirical evidence from the global energy industry", *Corporate Social Responsibility and Environmental Management*, Vol. 19 No. 6, pp. 317-326, doi: 10.1002/csr.280.

Peloza, J. (2009), "The challenge of measuring financial impacts from investments in corporate social performance", *Journal of Management*, Vol. 35 No. 6, pp. 1518-1541, doi: 10.1177/0149206309335188.

Pfeffer, J. and Salancik, G. (1978), *The External Control of Organizations: A Resource Dependence Perspective*, Harper & Row, New York.

Pham, H.S.T. and Tran, H.T. (2020), "CSR disclosure and firm performance: the mediating role of corporate reputation and moderating role of CEO integrity", *Journal of Business Research*, Vol. 120, pp. 127-136, doi: 10.1016/j.jbusres.2020.08.002.

Pinheiro, M.A.P., Jugend, D., Lopes de Sousa Jabbour, A.B., Chiappetta Jabbour, C.J. and Latan, H. (2022), "Circular economy-based new products and company performance: the role of stakeholders and Industry 4.0 technologies", *Business Strategy and the Environment*, Vol. 31 No. 1, pp. 483-499, doi: 10.1002/bse.2905.

Pisani, N., Kourula, A., Kolk, A. and Meijer, R. (2017), "How global is international CSR research? Insights and recommendations from a systematic review", *Journal of World Business*, Vol. 52 No. 5, pp. 591-614, doi: 10.1016/j.jwb.2017.05.003.

Preston, L.E. and O'Bannon, D.P. (1997), "The corporate social-financial performance relationship", *Business and Society*, Vol. 36 No. 4, pp. 419-429, doi: 10.1177/000765039703600406.

Quang, P.T., Rasoulinezhad, E., Linh, N.N. and Thao, D.P. (2022), "Investigating the determining factors of sustainable FDI in Vietnam", *China Finance Review International*, Vol. 12 No. 2, pp. 334-350, doi: 10.1108/CFRI-10-2021-0207.

Rahman, H.U., Zahid, M. and Al-Faryan, M.A.S. (2023), "ESG and firm performance: the rarely explored moderation of sustainability strategy and top management commitment", *Journal of Cleaner Production*, Vol. 404, 136859, doi: 10.1016/j.jclepro.2023.136859.

Rajesh, R. and Rajendran, C. (2020), "Relating Environmental, Social, and Governance scores and sustainability performances of firms: an empirical analysis", *Business Strategy and the Environment*, Vol. 29 No. 3, pp. 1247-1267, doi: 10.1002/bse.2429.

Ramirez, A.G., Monsalve, J., González-Ruiz, J.D., Almonacid, P. and Peña, A. (2022), "Relationship between the cost of capital and environmental, social, and governance scores: evidence from Latin America", *Sustainability*, Vol. 14 No. 9, p. 5012, doi: 10.3390/su14095012.

Rao, A., Dagar, V., Sohag, K., Dagher, L. and Tanin, T.I. (2023), "Good for the planet, good for the wallet: the ESG impact on financial performance in India", *Finance Research Letters*, Vol. 56, 104093, doi: 10.1016/j.frl.2023.104093.

Ren, X., Li, Y., Shahbaz, M., Dong, K. and Lu, Z. (2022a), "Climate risk and corporate environmental performance: empirical evidence from China", *Sustainable Production and Consumption*, Vol. 30, pp. 467-477, doi: 10.1016/j.spc.2021.12.023.

Ren, X., Zhang, X., Yan, C. and Gozgor, G. (2022b), "Climate policy uncertainty and firm-level total factor productivity: evidence from China", *Energy Economics*, Vol. 113, 106209, doi: 10.1016/j.eneco.2022.106209.

Rodriguez-Fernandez, M. (2016), "Social responsibility and financial performance: the role of good corporate governance", *BRQ Business Research Quarterly*, Vol. 19 No. 2, pp. 137-151, doi: 10.1016/j.brq.2015.08.001.

Russo, M.V. and Fouts, P.A. (1997), "A resource-based perspective on corporate environmental performance and profitability", *Academy of Management Journal*, Vol. 40 No. 3, pp. 534-559, doi: 10.2307/257052.

Saidat, Z., Silva, M. and Seaman, C. (2019), "The relationship between corporate governance and financial performance", *Journal of Family Business Management*, Vol. 9 No. 1, pp. 54-78, doi: 10.1108/JFBM-11-2017-0036.

Salehi, M. and Arianpoor, A. (2021), "The relationship among financial and non-financial aspects of business sustainability performance: evidence from Iranian panel data", *The TQM Journal*, Vol. 33 No. 6, pp. 1447-1468, doi: 10.1108/TQM-08-2020-0175.

Salehi, M., Lari DashtBayaz, M. and Khorashadizadeh, S. (2018), "Corporate social responsibility and future financial performance", *EuroMed Journal of Business*, Vol. 13 No. 3, pp. 351-371, doi: 10.1108/EMJB-11-2017-0044.

Sarkis, J. and Cordeiro, J.J. (2001), "An empirical evaluation of environmental efficiencies and firm performance: pollution prevention versus end-of-pipe practice", *European Journal of Operational Research*, Vol. 135 No. 1, pp. 102-113, doi: 10.1016/S0377-2217(00)00306-4.

Savio, R., D'Andrassi, E. and Ventimiglia, F. (2023), "A systematic literature review on ESG during the COVID-19 pandemic", *Sustainability*, Vol. 15 No. 3, p. 2020, doi: 10.3390/su15032020.

Schultze, W. and Trommer, R. (2012), "The concept of environmental performance and its measurement in empirical studies", *Journal of Management Control*, Vol. 22 No. 4, pp. 375-412, doi: 10.1007/s00187-011-0146-3.

Shafeeq Nimr Al-Maliki, H., Salehi, M. and Kardan, B. (2023), "The relationship between board characteristics and social responsibility with firm innovation", *European Journal of Management and Business Economics*, Vol. 32 No. 1, pp. 113-129, doi: 10.1108/EJMBE-04-2020-0094.

Shakil, M.H. (2021), "Environmental, social and governance performance and financial risk: moderating role of ESG controversies and board gender diversity", *Resources Policy*, Vol. 72, 102144, doi: 10.1016/j.resourpol.2021.102144.

Sharma, D., Bhattacharya, S. and Thukral, S. (2019), "Resource-based view on corporate sustainable financial reporting and firm performance: evidences from emerging Indian economy", *International Journal of Business Governance and Ethics*, Vol. 13 No. 4, p. 323, doi: 10.1504/IJBGE.2019.099565.

Sinha Ray, R. and Goel, S. (2023), "Impact of ESG score on financial performance of Indian firms: static and dynamic panel regression analyses", *Applied Economics*, Vol. 55 No. 15, pp. 1742-1755, doi: 10.1080/00036846.2022.2101611.

Stefan, A. and Paul, L. (2008), "Does it pay to Be green? A systematic overview", *Academy of Management Perspectives*, Vol. 22 No. 4, pp. 45-62, doi: 10.5465/amp.2008.35590353.

Suprabha, K.R., Sreepriya, J. and Prasad, K. (2024), "The impact of ESG disclosure on mitigating financial distress: exploring the moderating role of firm life cycle", *International Journal of Disclosure and Governance*, doi: 10.1057/s41310-024-00225-8.

Tadros, H. and Magnan, M. (2019), "How does environmental performance map into environmental disclosure?", *Sustainability Accounting, Management and Policy Journal*, Vol. 10 No. 1, pp. 62-96, doi: 10.1108/SAMPJ-05-2018-0125.

Teng, X., Chang, B.-G. and Wu, K.-S. (2021), "The role of financial flexibility on enterprise sustainable development during the COVID-19 crisis—a consideration of tangible assets", *Sustainability*, Vol. 13 No. 3, p. 1245, doi: 10.3390/su13031245.

United Nations (2020), "The sustainable development agenda – United Nations Sustainable Development", available at: https://www.un.org/sustainabledevelopment/development-agenda/

Veeravel, V., Murugesan, V.P. and Narayanamurthy, V. (2024a), "Does ESG disclosure really influence the firm performance? Evidence from India", *The Quarterly Review of Economics and Finance*, Vol. 95, pp. 193-202, doi: 10.1016/j.qref.2024.03.008.

Veeravel, V., Sadharma, E.K.S. and Kamaiah, B. (2024b), "Do ESG disclosures lead to superior firm performance? A method of moments panel quantile regression approach", *Corporate Social Responsibility and Environmental Management*, Vol. 31 No. 1, pp. 741-754, doi: 10.1002/csr.2598.

Velte, P. (2017), "Does ESG performance have an impact on financial performance? Evidence from Germany", *Journal of Global Responsibility*, Vol. 8 No. 2, pp. 169-178, doi: 10.1108/JGR-11-2016-0029.

Velte, P. (2020), "Corporate social responsibility and earnings management: a literature review", *Corporate Ownership and Control*, Vol. 17 No. 2, pp. 8-19, doi: 10.22495/cocv17i2art1.

Voinea, C.L., Hoogenberg, B.-J., Fratostiteanu, C. and Bin Azam Hashmi, H. (2020), "The relation between environmental management systems and environmental and financial performance in emerging economies", *Sustainability*, Vol. 12 No. 13, p. 5309, doi: 10.3390/su12135309.

Wagner, M. (2005), "How to reconcile environmental and economic performance to improve corporate sustainability: corporate environmental strategies in the European paper industry", *Journal of Environmental Management*, Vol. 76 No. 2, pp. 105-118, doi: 10.1016/j.jenvman.2004.11.021.

Wagner, M. and Schaltegger, S. (2004), "The effect of corporate environmental strategy choice and environmental performance on competitiveness and economic performance", *European Management Journal*, Vol. 22 No. 5, pp. 557-572, doi: 10.1016/j.emj.2004.09.013.

Wang, C. (2024), "The relationship between ESG performance and corporate performance - based on stakeholder theory", *SHS Web of Conferences*, Vol. 190, 03022, doi: 10.1051/shsconf/202419003022.

Wang, Z. and Sarkis, J. (2017), "Corporate social responsibility governance, outcomes, and financial performance", *Journal of Cleaner Production*, Vol. 162, pp. 1607-1616, doi: 10.1016/j.jclepro.2017.06.142.

Wang, N., Li, D., Cui, D. and Ma, X. (2022), "Environmental, social, governance disclosure and corporate sustainable growth: evidence from China", *Frontiers in Environmental Science*, Vol. 10, doi: 10.3389/fenvs.2022.1015764.

Who Cares Wins (n.d.),

Wichianrak, J., Khan, T., Teh, D. and Dellaportas, S. (2023), "Critical perspectives of NGOs on voluntary corporate environmental reporting: Thai public listed companies", *Sustainability*, Vol. 15 No. 7, p. 6195, doi: 10.3390/su15076195.

Wong, C.W.Y., Miao, X., Cui, S. and Tang, Y. (2018), "Impact of corporate environmental responsibility on operating income: moderating role of regional disparities in China", *Journal of Business Ethics*, Vol. 149 No. 2, pp. 363-382, doi: 10.1007/s10551-016-3092-z.

Wong, W.C., Batten, J.A., Ahmad, A.H., Mohamed-Arshad, S.B., Nordin, S. and Adzis, A.A. (2021), "Does ESG certification add firm value?", *Finance Research Letters*, Vol. 39, 101593, doi: 10.1016/j.frl.2020.101593.

Wooldridge, J.M. (2015), *Introductory Econometrics A Modern Approach*, Nelson Education, Toronto, Canada, References - Scientific Research Publishing, (n.d.), available at: https://www.scirp.org/reference/referencespapers?referenceid=2496102 (accessed 21 June 2024).

Wu, Z., Lin, S., Chen, T., Luo, C. and Xu, H. (2023), "Does effective corporate governance mitigate the negative effect of ESG controversies on firm value?", *Economic Analysis and Policy*, Vol. 80, pp. 1772-1793, doi: 10.1016/j.eap.2023.11.018.

Xie, J., Nozawa, W., Yagi, M., Fujii, H. and Managi, S. (2019), "Do environmental, social, and governance activities improve corporate financial performance?", *Business Strategy and the Environment*, Vol. 28 No. 2, pp. 286-300, doi: 10.1002/bse.2224.

Xu, Y. and Zhu, N. (2024), "The effect of environmental, social, and governance (ESG) performance on corporate financial performance in China: based on the perspective of innovation and financial constraints", *Sustainability*, Vol. 16 No. 8, p. 3329, doi: 10.3390/su16083329.

Xue, R., Wang, H., Yang, Y., Linnenluecke, M.K., Jin, K. and Cai, C.W. (2023), "The adverse impact of corporate ESG controversies on sustainable investment", *Journal of Cleaner Production*, Vol. 427, 139237, doi: 10.1016/j.jclepro.2023.139237.

Yadav, P. and Jain, A. (2023), "Sustainability disclosures and corporate boards: a stakeholder approach to decision-making", *Journal of Applied Accounting Research*, Vol. 24 No. 5, pp. 1027-1047, doi: 10.1108/JAAR-10-2022-0279.

Yoon, B., Lee, J. and Byun, R. (2018), "Does ESG performance enhance firm value? Evidence from Korea", *Sustainability*, Vol. 10 No. 10, p. 3635, doi: 10.3390/su10103635.

Zhao, L., Zha, Y., Wei, K. and Liang, L. (2017), "A target-based method for energy saving and carbon emissions reduction in China based on environmental data envelopment analysis", *Annals of Operations Research*, Vol. 255 Nos 1-2, pp. 277-300, doi: 10.1007/s10479-016-2163-y.

Zhao, C., Guo, Y., Yuan, J., Wu, M., Li, D., Zhou, Y. and Kang, J. (2018), "ESG and corporate financial performance: empirical evidence from China's listed power generation companies", *Sustainability*, Vol. 10 No. 8, p. 2607, doi: 10.3390/su10082607.

Zhou, S. (2022), "Reporting and assurance of climate-related and other sustainability information: a review of research and practice", *Australian Accounting Review*, Vol. 32 No. 3, pp. 315-333, doi: 10.1111/auar.12387.

Zhou, G., Liu, L. and Luo, S. (2022), "Sustainable development, ESG performance and company market value: mediating effect of financial performance", *Business Strategy and the Environment*, Vol. 31 No. 7, pp. 3371-3387, doi: 10.1002/bse.3089.

Zhou, C., Wang, K., Liu, R., Shu, A. and Wang, D. (2024), "Impact of ESG ratings on Chinese market performance during the COVID-19 crisis", *Journal of Accounting Literature*, doi: 10.1108/JAL-05-2024-0095.

Zimon, G., Arianpoor, A. and Salehi, M. (2022), "Sustainability reporting and corporate reputation: the moderating effect of CEO opportunistic behavior", *Sustainability*, Vol. 14 No. 3, p. 1257, doi: 10.3390/su14031257.

Zou, H.L., Zeng, S.X., Lin, H. and Xie, X.M. (2015), "Top executives' compensation, industrial competition, and corporate environmental performance", *Management Decision*, Vol. 53 No. 9, pp. 2036-2059, doi: 10.1108/MD-08-2014-0515.

Further reading

Pomarici, E. and Vecchio, R. (2014), "Millennial generation attitudes to sustainable wine: an exploratory study on Italian consumers", *Journal of Cleaner Production*, Vol. 66, pp. 537-545, doi: 10.1016/j.jclepro.2013.10.058.

Corresponding author

Prachi Jain can be contacted at: prachijainphd2k23@gmail.com

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: **permissions@emeraldinsight.com**

Check for updates

Journal of Neurochemistry

WILEY

Journal of Neurochemistry JNC The Official Journal of the International Society for Neurochemistry

ORIGINAL ARTICLE

Donepezil Improves PSD95 Expression, Mitigates Neuroinflammation via PI3K/Akt/NF-κB and Mitochondrial Dysfunction in a Rodent Model of Subarachnoid Haemorrhage

Ahmed Shaney Rehman[1] | Ammar Tasleem[1] | Nemat Ali[2] | Rehan Khan[3] | Mohd. Salman[4] | Pravir Kumar[5] | Suhel Parvez[1]

[1]Department of Medical Elementology & Toxicology, School of Chemical and Life Sciences, Jamia Hamdard, New Delhi, India | [2]Department of Pharmacology and Toxicology, College of Pharmacy, King Saud University, Riyadh, Saudi Arabia | [3]Chemical Biology Unit, Institute of Nano Science and Technology, Mohali, Punjab, India | [4]Department of Anatomy and Neurobiology, College of Medicine, The University of Tennessee Health Science Center, Memphis, Tennessee, USA | [5]Molecular Neuroscience and Functional Genomics Laboratory, Department of Biotechnology, Delhi Technological University (Formerly DCE), Delhi, India

Correspondence: Suhel Parvez (sparvez@jamiahamdard.ac.in)

Received: 11 January 2025 | **Revised:** 14 March 2025 | **Accepted:** 22 March 2025

Funding: This work was supported by the Researchers Supporting Project, King Saud University, Riyadh, Saudi Arabia (RSPD2025R940); the Department of Science and Technology, Ministry of Science and Technology, India [SR/PURSE Phase2/39 (C)], [SR/FST/LS-I/2017/05(C)]; and the Council of Scientific and Industrial Research, India [09/591/(0170)/2020-EMR-I].

Keywords: apoptosis | donepezil | mitochondrial biogenesis | neuroinflammation | PSD 95 | spatial working memory | subarachnoid haemorrhage

ABSTRACT

Mitochondrial dysfunction is a known contributor to subarachnoid haemorrhage (SAH) induced early brain damage (EBI), leading to poor neurological outcomes. An experimental SAH model was induced in adult male Wistar rats using endovascular perforation. Donepezil, an acetylcholinesterase (AChE) inhibitor (1 or 2 mg/kg body weight), was administered intraperitoneally 4 h after SAH. The severity of cerebral cortex injury was assessed using blood clot grading, behavioral tests and H and E staining. We carried out an assessment of neuroinflammatory markers using western blotting and immunofluorescence. Additionally, we examined neuronal architecture using H and E staining, measured mitochondrial redox imbalance or ROS and membrane potential (Δψm) and analyzed mitochondrial morphology using transmission electron microscopy (TEM). Apoptotic markers and mitochondrial respiratory complexes were assessed by western blotting. Our results indicated that donepezil treatment significantly upregulated PSD95, α7-AChR, CaMKII, BDNF, CREB, and PI3K expression in cerebral cortical neurons in response to SAH. This was accompanied by improved neurological function, reduced brain edema, decreased neuronal degeneration, and increased levels of OXPHOS and ATP. In the cerebral cortex, donepezil inhibited mitochondria-associated neuronal apoptosis after SAH as revealed by increased membrane potential integrity of mitochondria, reducing the ratio of Bax to Bcl-2 and inhibiting caspase-3 activity. Additionally, donepezil upregulated synaptic proteins (PSD95), strengthening synaptic connections and supporting spatial working memory circuits via the neurotrophic factor BDNF in post-SAH rats. Our research concludes that donepezil has neuroprotective benefits by inhibiting SAH-induced mitochondrial-mediated cell death through the regulation of Drp1-mediated mitochondrial morphology changes.

Abbreviations: α7-AChR, alpha 7 nicotinic acetylcholine receptors; BBB, blood–brain barrier; BDNF, brain-derived neurotrophic factor.; DCFDA, dichlorodihydrofluorescein diacetate; DPZ, donepezil; DRP 1, dynamin-related protein 1; EBI, early brain injury; GFAP, glial fibrillary acidic protein; MMP (Δψm), mitochondrial membrane potential; OXPHOS, oxidative phosphorylation; PSD 95, post synaptic density 95 protein; ROS, reactive oxygen species; SAH, Subarachnoid haemorrhage; TEM, Transmission electron microscopy; TMRE, tetramethyl rhodamine, ethyl ester.

© 2025 International Society for Neurochemistry.

1 | Introduction

Subarachnoid haemorrhage (SAH) occurs when blood accumulates in the subarachnoid space as a result of an aneurysm rupture (Suarez et al. 2006). Although relatively rare, comprising only 3%–4% of all strokes, it carries major implications. Approximately 40% of those diagnosed with SAH experience fatality, while 95% of those who survive face enduring challenges such as impaired mobility (Connolly et al. 2012), altered sensory perception (Jeon et al. 2009), speech difficulties (Hütter et al. 1994), cognitive impairments (Al-Khindi et al. 2010), fluctuations in mood, and chronic fatigue (Egeto et al. 2018). The sequence of pathophysiological processes that are altered following SAH, such as inflammation, oxidative stress, apoptosis, and abnormalities in cell electrophysiology, exerts an important impact on the progression of anxiety, cognitive impairment, and long-term memory issues (Rouanet and Silva 2019; Al-Khindi et al. 2010).

Brain injury after a SAH is diverse, occurring directly as early brain injury (EBI), and also developing secondarily as delayed brain damage (Schneider et al. 2018). EBI is influenced by several factors, such as cerebral ischemia, BBB compromise, brain swelling, cellular oxidative damage, inflammatory reactions, and apoptosis (Sehba et al. 2012). Therefore, it is crucial to develop neuroprotective techniques to combat EBI to improve clinical treatment outcomes (Zhang et al. 2017). Researchers recently employed a mouse model of SAH to develop a framework for understanding the short- and long-term consequences of the disease. They performed a range of neurobehavioral tests to evaluate the functional results in mice after the circle of Willis was perforated following SAH (Regnier-Golanov et al. 2022). Nevertheless, neurocognitive assessments alone may be insufficient in detecting impairments caused by SAH (Turan et al. 2020) (Pace et al. 2018); it is necessary to use particular tests that focus on learning and memory to evaluate the effectiveness of therapy (Fanizzi et al. 2017; Chung et al. 2021).

The role of PSD 95 in post-ICH synaptic impairment and cognitive dysfunction is being studied in recent studies (Wang et al. 2019). PSD 95 is a scaffolding protein involved in the development, remodeling, and maturation of neural synapses. It is essential for the assembly of excitatory synaptic signaling complexes. It takes part in NMDAR-mediated cerebral ischemia, in which an enhancement in PSD95-NMDAR interaction intensifies the entry of Ca^{2+} and triggers signaling cascades that lead to ischemia injury and neuronal death (Loomis et al. 2020; Cui et al. 2007). Moreover, the activation of BDNF-CREB proteins in the hippocampus following SAH enhances the strengthening of synaptic connections. BDNF and CREB are crucial in synaptic plasticity, a process that is thought to be responsible for memory formation (Yang et al. 2019). Post-SAH also has a considerable impact on the validity of important proteins, including CaMKII, a key serine/threonine kinase, essential for synaptic plasticity and memory formation, emphasizing the importance of these enzymes in maintaining and improving cognitive functioning (Gao et al. 2024).

Mitochondria are vital organelles within cells that play a crucial role in energy generation within the central nervous system (CNS) (Koch et al. 2017). They regulate apoptosis, cell differentiation, and the generation of ROS (Sharma et al. 2018). Removing excess ROS is necessary to preserve the normal state of cells. On the other hand, abnormalities in mitochondrial biogenesis may result in the buildup of ROS (Xian et al. 2019; Pei et al. 2016). ROS-induced cytochrome C release from damaged mitochondria causes activation of a caspase-dependent mechanism, which in turn initiates apoptosis (Rashmi et al. 2019). While excessive fission may lead to ROS generation (Qin et al. 2019), balanced fission and fusion are necessary for optimal mitochondrial stability. It is essential to understand these mechanisms to manage CNS disorders. Mitochondrial quality is regulated by certain activities, such as mitofusion (Mfn1 and Mfn2) and mitofission (Dynamin-related protein 1, DRP1) (Meyer et al. 2017). Therefore, reducing ROS generation and neuronal death based on mitochondrial malfunction requires limiting mitochondrial fission and fusion. ROS production is linked to neuroinflammation, which is a crucial component of early brain damage (EBI) following SAH. Hemoglobin, a blood component, binds to pathogen recognition receptors (PRRs) including the fourth Toll-like receptor (TLR4) on microglia following an injury (Hanafy 2013). This attachment activates inflammatory pathways such as NF-κB, TNF-alpha, and IL-1 beta (Lin et al. 2012).

Donepezil, a second-generation acetylcholinesterase inhibitor, has been approved for the treatment of Alzheimer's disease related dementia (Birks and Harvey 2018). Additionally, it demonstrates the potential to reduce the damage caused by ischemic stroke through the stimulation of neurogenesis, angiogenesis, inhibition of inflammation, and apoptosis (Madani Neishaboori et al. 2021). Nevertheless, donepezil remains an unexplored candidate for repurposing in SAH, making it a possible drug development target. This study aimed to examine the possibility of using donepezil for SAH by evaluating its effectiveness in reducing post-SAH changes through mitochondrial-mediated Drp1 fission with cognitive impairment and short-term memory problems. We assessed the neuroprotective properties of donepezil on neurological impairments, mitochondrial oxidative stress, and histological alterations using an endovascular perforation model of SAH. It was assumed that being an acetylcholinesterase inhibitor, donepezil could be a potential candidate, for enhancing acetylcholine levels, which interacts with nicotinic receptors, specifically α7 nAChRs (Shen et al. 2010; Ongnok et al. 2021). These receptors have the potential to improve cognitive function by modulating the PSD-95 protein and the BDNF-CREB-CamkII signaling pathway. This research is the first to examine donepezil's role in promoting mitochondrial-mediated neuroprotection in a rodent SAH model, suggesting its promise as a therapeutic target for drug development.

2 | Material and Methods

2.1 | Chemicals

Donepezil (C. no. 120011-70-3) was acquired from Sigma. We purchased tetramethyl rhodamine ethyl ester [TMRE] (C. no. 115532-52-0 λex 540 nm; λem 595 nm in DMSO), 2,7-dichlorodihydrofluorescein diacetate (DCDHF-DA) (C. no. 4091-99-0, Sigma), dihydroethidium (DHE) (C. no. 38483-26-0),

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and RIPA lysis buffer (MFCD02100484) from Sigma Aldrich and cell signaling technology. Apoptotic markers such as BAX (SC-7480 mouse monoclonal IgG2b κ Bax antibody), Bcl2 (SC-7380 mouse monoclonal IgG1 κ Bcl-2 antibody), and Caspase-3 (SC-65497 mouse monoclofnal IgG2a κ) were obtained from Santa Cruz Biotechnology. Anti-Iba1(Rabbit mAb #17198, CST) is a marker for glial cells, while anti-GFAP (Cat #14–9892-82 source Mouse monoclonal IgG1, Invitrogen) is a marker specifically for astrocytes. The proteins PSD 95 (RRID: AB_2292883), Creb (RRID: AB_331277) (CST), and Oxphos (ab110413 Rodent WB Antibody Cocktail) (Abcam) are being referred to. The chemicals ethanol (110-73-6), diethyl ether (60-29-7), EGTA (60-00-4), bovine serum albumin (BSA) (9048-46-8), and orthophosphoric acid (OPA) (7664-38-2) were purchased from Loba and SRL as standard chemicals.

2.2 | Animal and Experimental Group

Male Wistar rats weighing between 230 and 250 g and aged 8–10 weeks were utilized in this study under the guidelines set by the Institutional Animal Ethics Committee (IAEC), Jamia Hamdard, New Delhi, India (CPCSEA Registration: 173/GO/ReBi/2000/CPCSEA). Male Wistar rats (weighing 230–250 g) were purchased from in-house Centre Animal House Facility (CAHF) Jamia Hamdard University. The rats were kept in hygienic plastic cages with metal covers, with each cage housing 3–4 rats. The temperature was maintained at 25°C ± 1°C, with a 12-h cycle of light and darkness and a humidity level of 60%. The food pellets were provided ad libitum by Arshiwad Industries Pvt. Ltd., located in Punjab, India. The water used was filtered and obtained from RIOS (Merck Millipore, Billerica, MA). All experimental protocols were approved by the IAEC (Protocol no. 1742). Two hundred male rats were used in this work, and they were divided into four groups at random: (a) a sham-operated control group that underwent all procedures except for the occlusion of the right ICA; (b) a SAH surgery group; (c) a SAH + donepezil (1 mg/kg) treatment group; and (d) a SAH + donepezil (2 mg/kg) treatment group. First, we arranged the rats descendingly according to their weight. We then systematically divided them between different groups using Microsoft Excel (version 2019) to ensure a balanced distribution. All groups have a standard error mean of ±5. As a result of simple randomization, every rat has an equal chance of being assigned to any group, reducing selection bias and assuring unbiased group allocation (Kim and Shin 2014). Donepezil hydrochloride was administered intraperitoneally every 12 h, with daily doses of 1 and 2 mg/kg body weight split into two equal injections. The first dose was given on the day of surgery, followed by the second dose 12 h later on the same day. The treatment started 4 h after the surgery and lasted for 24 h (as shown in Figure 1). Out of the 148 rats with SAH examined, 46 (31.08%) did not survive (Table S1). The endovascular perforation SAH model shows high mortality due to abrupt haemorrhage causing vasospasm, reduced cerebral blood flow, impaired autoregulation, increased ICP and global ischemia, especially within 24 h (Fabiana et al. 2003). In this study, the mortality timeline showed the highest death rate within the first 24 h post-SAH, with 19 animals dying on day one, followed by 15 and 12 animals on subsequent days. This pattern suggests that the acute phase of SAH is the most critical, with mortality driven by the immediate consequences of haemorrhage and associated complications. Animals with severe deformities and epileptic seizures (approximately 16) were excluded (Figure S1). The remaining rats (36 operated as a control) were classified as having either undergone a simulated operation or experienced a reduced amount of bleeding, resulting in SAH grades below 13. The experimenter was kept blind to the allocation of treatments during the entire process.

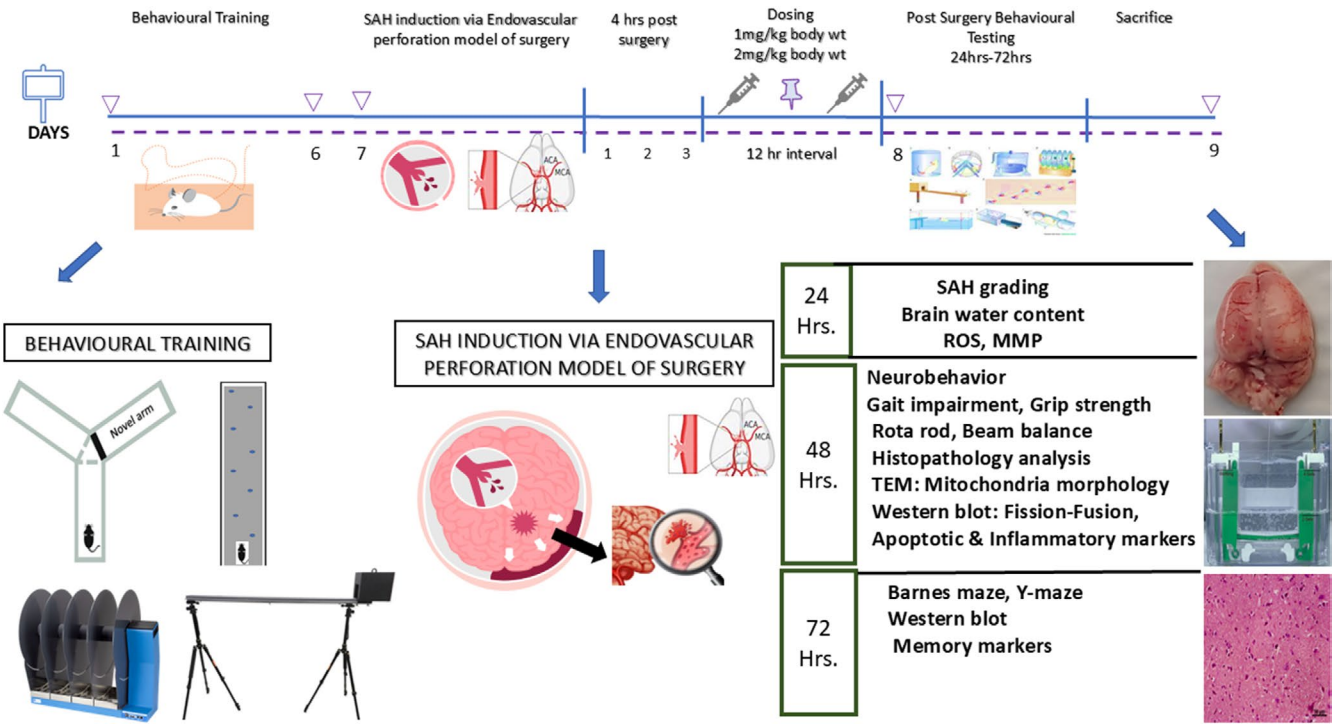


FIGURE 1 | Experimental design.

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.2.1 | SAH Model

Rats were used in the endovascular perforation model of SAH, with minor modifications made to the procedure previously described (Zhang et al. 2019). Anesthesia was administered by utilizing 4% isoflurane in a combination of 70% medical air and 30% oxygen. To sustain anesthesia during surgery, isoflurane was maintained at 1.5%. All animals were anesthetized with isoflurane (2%–3%) during surgery to guarantee total unconsciousness and painlessness. The rectal temperature was kept at 36.0°C ± 0.5°C using a warming pad. A 3- to 4-mm stump remained after the left external carotid artery was severed and ligated. After that, the external carotid artery stump was used to cannulate a sharpened 3–0 suture into the left internal carotid artery, puncturing the vessel wall where the anterior cerebral artery and middle cerebral artery split. Post-surgery, animals were monitored every 4 h for the first 24 h and twice daily thereafter for signs of pain or distress, including changes in posture, grooming, mobility, and body weight. The veterinarian quickly reassessed the animals if they showed any indications of distress. Environmental enrichment, including soft bedding and easy access to food and water, was provided to enhance recovery. Analgesics were not administered postoperatively to avoid confounding effects on neuroinflammation and neuronal recovery, which were key study outcomes. Donepezil, administered 4 h post-surgery, is known to reduce neuroinflammation and has mild analgesic properties. Nonpharmacological measures, including temperature-controlled environments, supported recovery, and human endpoints were predefined to minimize suffering (Kogure et al. 2017). Except for perforation, sham-operated rats received the same treatments. We maintained all the animals on ventilation until the anesthesia subsided.

2.2.2 | Drug Administration

Two distinct doses of Donepezil were given through intraperitoneal injection (ip), following a prior investigation (Madani Neishaboori et al. 2021). Donepezil, at a concentration of 3 mg/kg, was mixed in 1 mL of standard saline solution and delivered through an intraperitoneal injection at doses of 1 and 2 mg/kg, 4 h after SAH. Previous findings have reported on the pharmacokinetics of donepezil, where its concentration was measured in various organs, including the brain, following administration in Wistar rats (Krishna et al. 2018). Following that, donepezil was administered in doses of 1 mg/kg and 2 mg/kg at 12-h intervals to two different SAH-induced groups. The doses were not divided (as an instance, 1 mg/kg was not divided into two 0.5 mg/kg doses or 2 mg/kg into two 1 mg/kg doses). This approach aimed to assess short-term outcomes during a 24- to 72-h period.

2.2.3 | Neurological Score Evaluation

After 24 h following a subarachnoid haemorrhage (SAH), a neutral evaluator performed neurological assessments using the 6-point scale developed by Garcia et al. (Garcia et al. 1995). The evaluations included five distinct tests: the following assessments were conducted: (1) monitoring the natural activity after 5 min, (2) evaluating the voluntary movement of all limbs, (3) viewing the climbing behavior on the wire cage's wall, (4) measuring the extension of the fore limbs when the subject was held by the tail, and (5) evaluating the reflex response when both sides of the trunk were touched. The grading of each test was determined on a scale ranging from 0 to 3, taking into account the observable consequences that occurred after SAH. The scale spanned from negligible or minimum movement (0–1) to vigorous movement or performance equivalent to pre-SAH conditions (2–3), with uniform grading criteria utilized for all assessments.

2.2.4 | Neurobehavioral Assessment

To investigate whether donepezil improves the outcome of SAH, we assessed six neurobehavioral measures both before the surgery and at various time points between 24 and 72 h following the SAH procedure. After conducting a neurological assessment, we carried out grip strength, beam balance, and rotarod tests between 24 and 48 h post-SAH to evaluate motor coordination and balance. Additionally, we looked at gait impairment during this time to further assess any changes in motor function.

2.2.5 | Gait Analysis

Mendes et al. (2015) thoroughly explained the footprint method, which is also referred to as gait analysis. This method entailed submerging the fore paws and hind paws of rats in distinct nonharmful watercolor substances. Afterward, the rats traversed a distance of 1.2 m on a white paper surface, resulting in the imprinting of their paw prints. Stride length was determined by measuring the distance between successive imprints of the same paw, whereas stride width was evaluated by measuring the distance between the imprints of the left and right forelimbs or hind limbs. After a surgical procedure for SAH, gait analysis was performed at 48 h.

2.2.6 | Beam Balance

The assessment of vestibulomotor function after surgery (48 h) was conducted using a modified beam balance test, following the protocols described in the reference (Soltani et al. 2015). Animals received preoperative instruction. Each group performed five attempts on a beam measuring 160 cm in length and having a diameter of 2.5 cm. The average of three trials was calculated during the testing process. The experiment commenced with rats starting from one end of the beam and concluded upon their traversal of a distance of 1 m. The trials had a maximum duration of 60 s, during which the time spent maintaining balance was recorded. A score of 5 signifies the complete keeping of balance for 60 s, with each 12-s interval earning 1 point. Falls and traversal time were used as indicators of vestibulomotor function.

2.2.7 | Grip Strength

Evaluating grip strength is a widely used technique to determine the impact of medicines, toxins, or injuries on motor function.

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

To assess the strength of the forelimb and hind limb, grip tests were performed at 48 h after SAH. A one-meter string was securely stretched between two upright supports that were placed 50 cm above the base of the experiment. Rats were placed halfway down the string, and their performance was scored on a scale from 0 to 6. Here, 0 means one falls off within 30 s, 1 means one grips the string with one's forepaws for 30 s, and 2 is the same as 1 except one tries to climb the string. Three refers to the act of grasping the line using two front paws and either one or both hind paws; four indicates using all forepaws to grasp the string and encircling it with the tail; and five indicates a successful escape (Hall 1985).

2.2.8 | Rotarod

Before the surgical procedure, rats were subjected to a 3-day training program using a dynamic rotarod (Omni dynamic Rotor, Columbus, USA) set to maintain a consistent speed of 10 rpm (rpm). This regimen was designed to evaluate changes in motor coordination and balance following a subarachnoid haemorrhage (SAH) at 48 h. After a 1-min break between trials, a total of five trials were conducted, each lasting a maximum of 180 s. The latency to fall (measured in seconds) was recorded during each trial (Rosell et al. 2013).

2.2.9 | Barnes Maze

The Barnes maze used a 91.5-cm-wide black circle with 20 holes, each measuring 5 cm in diameter, spaced 3 cm from the circle's edge. An escape box (10.5 × 6 × 6 cm) could be placed beneath a designated hole. The rats were habituated to the starting location and escape box for 2 days. Rats were trained by placing them in the middle of the maze for 10 s, then exposing them to unpleasant stimuli to get them to enter the escape box. There were two trials and five daily sessions in each session. Rats were led to the escape box by the experimenter if needed, and trials had a maximum duration of 3 min. The final test evaluated the post-SAH rats (at 72 h) capacity to locate the escape box. Rats' travel times through any given maze were timed to determine the degree of anxiety and cognitive impairment. Odor cues were reduced by cleaning the maze with 70% ethanol in between each animal (Broadbelt et al. 2022).

2.2.10 | Y Maze

The Y-maze test quantifies the ability to remember spatial information in the short term (Sarter et al. 1988). Rats that were trained 3–6 days before, were given doses of donepezil and conducted Y-maze testing at 72 h after endovascular perforation for subarachnoid haemorrhage (SAH). The Y-maze comprises three arms that are equidistant from each other. Every rat was individually placed in one arm of the maze and given 5 min to freely explore. The researchers recorded both the spontaneous alternation and total entries. An actual alternation refers to a sequence in which consecutive entries appear in all three arms, such as ABC, CAB, and BAC. However, patterns like ABA are not included in this definition. The following formula was utilized to get the percentage of alternations: [(the number of alternations)/(the total number of arm entries-2)] × 100.

2.2.11 | SAH Grade

Rats were subjected to surgical induction of SAH and were euthanized 24 h after the surgery by cervical dislocation while under anesthesia induced with ketamine and xylazine (75 mg/kg and 10 mg/kg bt.wt.) respectively. The severity of SAH was blindly graded after euthanasia by an observer using a predetermined grading scale (24 h post-SAH). The basilar area around the circle of Willis was divided into six subsegments graded on a scale of 0–3 based on subarachnoid blood clot presence. The severity of SAH was assessed by calculating the overall score obtained from adding up the scores of each subsegment. SAH was classified as mild if the score ranged from 0 to 7, moderate if it fell between 8 and 12, and severe if it ranged from 13 to 18. Animals that had a cumulative score of eight or less were not included. Non-haemorrhaged or sham-operated animals received a score of zero. SAH grading was performed after sacrificing the animal after post-behavioral assessment.

2.2.12 | Brain Water Content

The rats were anesthetized with an intraperitoneal injection of a mixture of ketamine (75 mg/kg) and xylazine (10 mg/kg). Under anesthesia, the brains were instantly removed following decapitation. Subsequently, the brain's left and right hemispheres were separated (24 h post-SAH). The weight of the brain was promptly measured while it was still wet, and then the brain was dried at a temperature of 100°C for 48 h to ascertain its weight after it had dried out for 24 h following SAH. Brain water content % was calculated by [(wet weight − dry weight)/wet weight] × 100.

2.3 | Fluorescent Immunohistochemistry

Rats were anesthetized intraperitoneally with a mixture of ketamine (75 mg/kg bt.wt.) and xylazine (10 mg/kg bt.wt.) and subjected to transcardiac perfusion with 4% paraformaldehyde fixative for immunofluorescence staining. After the sacrifice, the brains were removed, preserved for 24 h in 4% paraformaldehyde, and then embedded in paraffin. Eight- to 10-micrometer sections were cut, xylene deparaffinized, and then dried. A solution of sodium citrate monohydrate with a concentration of 10 mM and a pH of 6.0 was utilized for antigen retrieval, and it was microwaved for 15 min. After that, 0.1% Triton X-100 was used to permeabilize the sample. After that, sections were treated for 60 min at 37°C with 5% normal donkey serum. Next, they were exposed to the primary antibody, GFAP (1:50 dilution; Invitrogen, UK), and anti-Iba1 (Cell for an overnight period at 4°C). The sections were subjected to treatment with a goat anti-rabbit Alexa Fluor 488 secondary antibody (diluted 1:500; Invitrogen, Scotland, UK) after being washed with PBS. Ultimately, antifade DAPI was applied to the slices, and a fluorescence microscope (Zeiss, Oberkochen, Germany) was used to take pictures of the ipsilateral cerebral cortical area.

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.4 | TEM

Samples of rats brain tissue, thin slices, or tiny cubes with a thickness of no more than 1 mm, were gathered, with an emphasis on the cerebral cortex area in order to evaluate SAH injuries (48 h post SAH). The samples were initially fixed by immersing them whole night at 4°C–5°C in a solution containing 2% glutaraldehyde in cacodylate buffer (consisting of 0.1 M sodium cacodylate and 2 mM MgCl2). Following three washes at 4°C using cacodylate buffer, the samples were then fixed for 1 h at 4°C in a 2% osmium tetroxide solution. Following fixation, the tissues were dried resin blocks were made, and ultramicrotomy was used to cut extremely thin 60 nm slices. Following their collection on metal mesh grids, these sections were dyed with uranyl lead staining. After staining, the sections were finally inspected under a TEM microscope (Talos).

2.5 | Mitochondrial and Cytosolic Fractionation

Differential centrifugation was used to separate the cytosolic and mitochondrial components (24 h post-SAH). With the aid of a glass dounce homogenizer, brain tissue was homogenized in a mitochondrial isolation solution that contained 0.07 M sucrose, 0.22 M mannitol, 10 miliMolar HEPES-KOH (pH 7.2), 0.1% fatty acid-free BSA, and 1 miliMolar EDTA. To separate intact cells and nuclei, the homogenate was centrifuged at 800 × g for 3 min at 4°C. After centrifuging the resultant supernatant for 10 min at 10 000 rcf, a loosely packed mitochondrial pellet was obtained; the cytosolic fraction was discarded. Using a discontinuous Ficoll gradient (3% layered over 6%) supplemented with 0.12 M mannitol, 0.03 M sucrose, and 25 mM EDTA, the mitochondrial fraction was further purified. After 30 min at 4°C and 10 400 × *g* of centrifugation, the mitochondrial fractions were successfully separated (Brown et al. 2004).

2.5.1 | Mitochondrial ROS and Mitochondrial MMP

Mitochondrial redox imbalance or ROS and membrane potential (MMP) were measured using BD Lyric flow cytometry (Becton Dickinson, San Jose, CA, USA). A 400 μL analytical buffer containing 250 miliMolar sucrose, 10 mM Tris base, 20 miliMolar 3-[N-morpholino] butane sulfonic acid, 100 M Pi[K], and 0.5 miliMolar Mg^{2+} was used to evaluate a 25 g sample of mitochondria. H2DCFDA (10 M) and TMRE (100 nM) were used to measure mitochondrial redox imbalance or ROS and MMP (Δψm), respectively. To measure ROS, excitation wavelengths of 495 and 540 nm were employed, while emission wavelengths of 525 and 595 nm were utilized for MMP. Fluorescence intensity was measured after mitochondrial labeling in six different sample sets, and 10 000 events were recorded in the flow cell. A logarithmic scale was used to create the scatter plot, which was then examined using FACS DIVA software.

2.5.2 | Cellular ATP Content Assay

For the assessment of cellular ATP concentration in brain tissue, a widely used ATP colorimetric assay kit from Elabscience in the USA was utilized. The brain sample was broken down into small fragments and then placed in hot distilled water at a w/v ratio of 1:9 for 10 min. Afterward, the mixture was vigorously mixed for 1 min, and the resulting liquid was used to quantify the cellular adenosine triphosphate (ATP) levels. The supernatant was acquired through centrifugation of the mixture at a speed of 10 000 rpm, following the methodology outlined in the kit's instructions.

2.6 | Western Blotting

Rats were sacrificed by cervical dislocation under anesthesia induced by intraperitoneal administration of ketamine (75 mg/kg bt.wt.) and xylazine (10 mg/kg bt.wt.). Samples of brain tissue were taken from several anesthetized groups of rats at 48–72 h after SAH surgery. The ipsilateral (left) cerebral hemisphere was suspended, homogenized, and sonicated in ice-cold RIPA (CST, USA) lysis buffer containing a protease and phosphatase inhibitor cocktail (20 mM Tris–HCl [pH 7.5], 150 mM NaCl, 1 mM Na_2EDTA, 1 mM EGTA, 1% NP-40, 1% sodium deoxycholate, 2.5 mM sodium pyrophosphate, 1 mM β-glycerophosphate, 1 mM Na_3VO_4, and 1 μg/mL leupeptin). All of the samples were then centrifuged for 30 min at 21 000*g*. The Bradford test was used to quantify the protein; 20–30 μg of protein were loaded onto 8%–15% SDS-PAGE gels, and then the gels were transferred onto 0.45 μm PVDF membranes. Additionally, after transferring the proteins onto the membrane, the membranes were blocked with 5% non-fat dry milk for 60 min at room temperature. Then, using distinct primary antibodies (Bax, Bcl2, Caspase 3, OXPHOS, TLR4, NfkB, Cytochrome C, AKT) that were acquired from multiple sources and targeted at different proteins, these membranes were incubated for an entire night. Secondary antibodies conjugated to horseradish peroxidase were added at a dilution ratio of 1:5000 for 90 min at room temperature following three rounds of PBST washing. The density of the bands was assessed using ImageJ software for densitometric measurements (version 1.50, NIH, USA) following observation using an enhanced chemiluminescence reaction.

2.7 | Histopathological Examination

H and E staining was carried out to assess neuronal degeneration using a previously established methodology. Histopathological examination started at 48 h. Rats were fixated with 4% paraformaldehyde after being anesthetized and having a transcardial perfusion of 0.9% saline. H and E was used to stain 5-μm thick coronal frozen slices for 10 min, or until the required staining intensity was reached. The stained specimens were viewed and photographed using an inverted optical microscope, and analysis was conducted using Image J software.

2.8 | Statistical Analysis

The statistical analysis was conducted using GraphPad Prism 6.0 software, developed in the United States. The data is displayed as the Mean ± SD. The statistical differences between groups were determined using ANOVA, followed by Tukey's test. A $p < 0.05$ value was deemed to be statistically significant. Based on the previously published reports and pilot study data (Uzair

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

et al. 2024; Weissgerber et al. 2015), effect sizes were calculated using G*Power analysis. An effect size of 0.299 was obtained for ANOVA, with an α error of 0.05 and a power of 95%. A sample size of $n = 3$ was used to assess the statistical significance of the data (Table S2). The data exhibited parametric characteristics, as confirmed by the non-significance of the Shapiro–Wilk test for normality (Table S3). No test for outliers was conducted and no data points were excluded from the analysis.

3 | Result

3.1 | Donepezil Administration Attenuates SAH-Induced Grading Score, Neurological Deficits, and Brain Water Content

Blood clots were primarily seen near the ventral brain stem and circle of Willis. Their severity was measured using a grading scale ranging from 1 to 18, following the onset of SAH. In this work, our attention was directed towards the specific area of the brain, that is, the right cortical region (Figure 2A,B). A statistically significant difference was seen between the SAH grades in rats and the sham-operated groups ($p < 0.0001$, F (3, 20) = 154.7; Figure 2A,B). Administration of donepezil resulted in a substantial decrease in clot formation in rats that underwent SAH (#$p < 0.05$ or ###$p < 0.001$). In terms of neurological outcomes, the induction of SAH resulted in a substantial rise in the neurological impairment score as compared to the sham-operated group ($p = 0.0676$, F (3, 32) = 37.87; Figure 2C). The administration of donepezil at the highest concentration resulted in a substantial improvement in the neurological impairment score (###$p < 0.001$). After inducing SAH, there was a significant increase in brain water content ($p = 0.0008$, F (3, 20) = 8.386; Figure 2D) and higher levels of AQP4 protein after 24 h, compared to the sham-operated group [$p = 0.0008$, F (3, 8) = 17.17; Figure 2E]. Administration of donepezil resulted in a notable decrease in brain water content (##$p < 0.01$) and AQP4 levels in the first 24 h after SAH (#$p < 0.05$ or ##$p < 0.01$).

3.2 | Donepezil Rescued Neurobehavioral Impairment Caused by SAH

To examine whether donepezil benefits SAH outcome, we analyzed six neurobehavioral measures both prior to the surgical procedure and at intervals between 24 and 72 h after SAH induction. After the neurological scoring assessment, grip strength, beam balance, and rotarod tests were performed between 24 and 48 h post-SAH surgery to evaluate motor coordination and balance. Gait impairment was also assessed during this period. Combined anxiety and memory deficits were investigated between 48 and 72 h post surgery using Barnes maze and Y-Maze tests, providing insight into cognitive recovery and spatial memory function. This sequential testing ensured a comprehensive evaluation of motor and cognitive outcomes while minimizing stress on the animals. To perform gait analysis, we recorded the footprints of the rats to quantify the stride length, which is the distance between two consecutive imprints formed by the same limb is referred to as step length, while the stride width is defined as the distance between the footprints of the left and right paws. The rats subjected to SAH showed a substantial decrease in both stride length ($p < 0.0001$, F (3, 32) = 9.858; F (3, 32) = 32.68 Figure 3A–C) and width ($p = 0.0024$, F (3, 32) = 5.935; $p = 0.0001$; F (3, 32) = 9.599; Figure 3A,D–E) when compared to the sham group. The administration of donepezil resulted in a significant improvement in these specific measures of walking (##$p < 0.01$), stride length forelimbs and hind limbs; (##$p < 0.01$ or ###$p < 0.001$, stride width forepaws and hind paws), leading to an enhanced walking pattern in rats that had experienced SAH. The results of the Rota rod test demonstrated that motor ability was considerably decreased following SAH surgery, as seen by an increased fall latency compared to rats who underwent a sham operation (F (3, 32) = 25.39, $p < 0.0001$ Figure 3F). Treatment with donepezil led to a significant improvement in motor function, resulting in a reduction in fall latency (##$p < 0.01$ or #$p < 0.05$). The beam balancing test was employed to evaluate motor coordination ($p < 0.0001$, F (3, 32) = 309.9; Figure 3G). After SAH, therapy with donepezil significantly enhanced motor coordination in a dose-responsive manner (##$p < 0.01$ or ###$p < 0.001$). The grip strength test was employed to evaluate muscular strength. The study found that SAH caused a substantial decrease in neuromuscular function, leading to a decrease in muscular strength compared to the sham-operated group ($p < 0.0001$, F (3, 32) = 17.34; Figure 3H). On the other hand, the administration of donepezil greatly improved grip strength in rats compared to the rats in the SAH group (##$p < 0.01$ or ##$p < 0.01$).

3.3 | Donepezil Upregulates Post-Synaptic Density Protein, Improving SAH-Induced Memory Deficits

In order to detect the expression of α7 nAChR, BDNF-CREB, and CamkII, a protein quantification or western blot analysis was conducted to elicit the effect of donepezil on the α7 nAChR receptor. In contrast to the Sham-operated group, the SAH rodents exhibited a substantial decrease in α7 nAChR ($p = 0.0001$, F (3, 8) = 29.77; Figure 4C,D), BDNF-CREB($p = 0.0004$, F (3, 8) = 20.83; $p = 0.0008$, F (3, 8) = 16.69 Figure 4C,G–H), and CamkII ($p = 0.0002$ F (3, 8) = 24.48; Figure 4C,F). Donepezil administration significantly enhanced the expression of these proteins (#$p < 0.05$, ##$p < 0.01$ or ###$p < 0.001$). Additionally, PSD95 expression, which declined in SAH rats compared to the Sham group ($p < 0.0001$, F (3, 12) = 159.4; $p = 0.0001$, F (3, 8) = 27.51; Figure 4–B,E) was significantly increased (###$p < 0.001$) with donepezil treatment, suggesting maintenance of post-synaptic circuits. As an additional method to confirm that donepezil improves memory deficits in rats, we used Barnes and Y maze tests to measure anxiety and working memory reductions at 72 h after SAH onset (Figure 4A–H). During the Barnes maze test, SAH rats exhibited significantly increased escape latency and entries in the target zone compared to the Sham group ($p = 0.0001$, F (3, 32) = 83.35; $p < 0.0001$, F (3, 32) = 88.29; Figure 5A–C). Treatment with donepezil dramatically decreased entry in the target zone and escape latency (##$p < 0.01$ or ###$p < 0.001$). Unlike SAH rats, Sham rats demonstrated considerable alternation and no arm preference in the Y-maze spontaneous alternation task ($p < 0.0001$, F (3, 32) = 20.14; Figure 5D,E). In SAH rats, donepezil therapy dramatically decreased arm preference and increased working memory (#$p < 0.05$ or ###$p < 0.001$). Donepezil-treated SAH rats showed improved working memory and reduced arm preference, confirming the cognitive benefits of donepezil in this model.

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

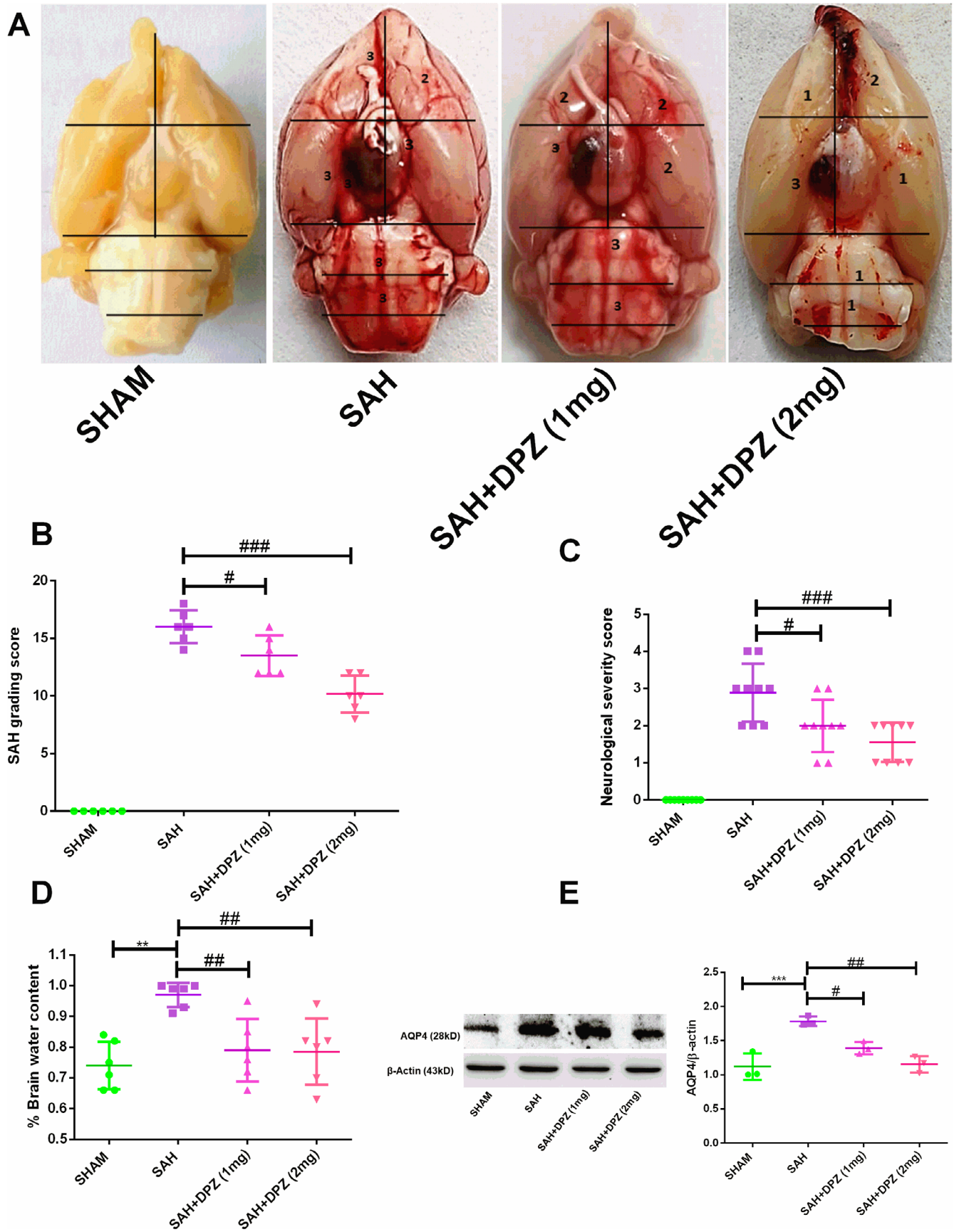


FIGURE 2 | Legend on next page.

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

FIGURE 2 | Effects of donepezil on SAH grading score, neurological severity, aquaporin expression, and brain edema. (A, B) Representative images of the brain across different experimental groups. Validation of the SAH model was conducted by quantifying the blood clot in cerebral cortex near the Circle of Willis with a grading score (1–18), $n = 6–9$ rats per group. (C) Neurological severity score. (D, E) Immunoblot showing that donepezil suppressed AQP4, reduced brain water content, and improved blood–brain barrier function 24 h after SAH ($n = 3$ rats per group). Experimental data are presented as means ± SD and were analyzed using a one-way ANOVA followed by the Tukey post hoc test. $^{**}p < 0.01$, $^{***}p < 0.001$ versus sham, $^{\#}p < 0.05$, $^{\#\#}p < 0.01$, $^{\#\#\#}p < 0.001$ versus SAH.

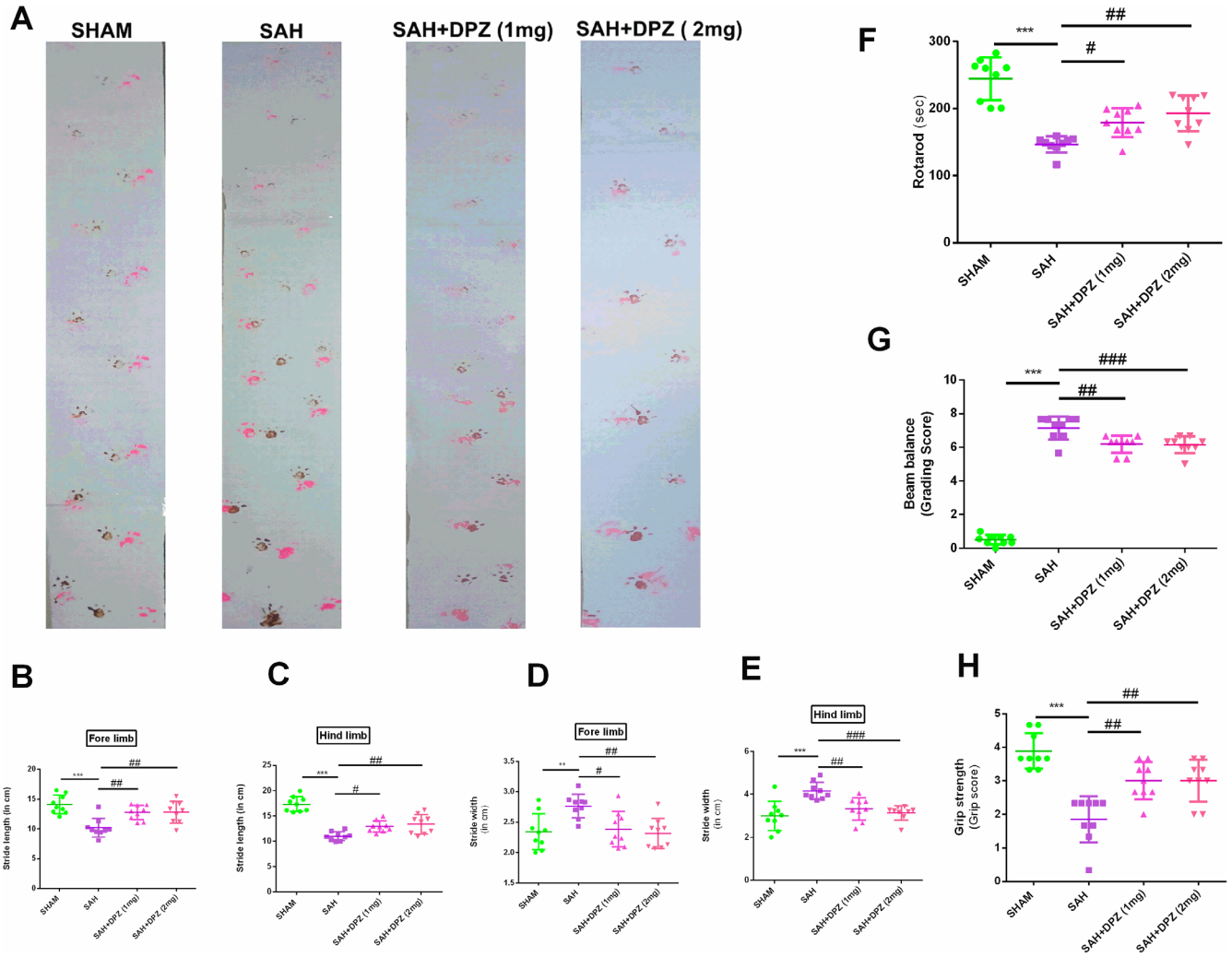


FIGURE 3 | Effect of donepezil on rats with SAH-induced short-term neurobehavioral dysfunction. (A) A diagrammatic representation of the rat's gait analysis, with the forepaws colored brown and the hind paws colored pink, illustrating the animal's foot printing pattern. Measurements include (B, C) stride length, indicating the distance between successive steps of the same paw; (D, E) stride width, the lateral distance between the left and right paws during movement; (F) rotarod performance, testing motor coordination and balance by measuring the duration the rat can stay on a rotating rod; (G) beam balance, evaluating the ability of the rat to maintain balance while walking on a narrow beam; and (H) grip strength, assessing the muscle strength of the forelimbs and hind limbs. Experimental data are presented as means ± SD ($n = 9$ rats per group) and were analyzed using a one-way ANOVA followed by the Tukey post hoc test. $^{**}p < 0.01$, $^{***}p < 0.001$ versus sham, $^{\#}p < 0.05$, $^{\#\#}p < 0.01$, $^{\#\#\#}p < 0.001$ versus SAH.

3.4 | Donepezil Regulates the TLR4/NF-κB/PI3K/AKT Pathway in Post-SAH Surgery

Inducing SAH-activated TLR4/NF-κB pathways, as indicated by increased expression, along with possible involvement of PI3K/AKT pathways in inflammation. The activation of TLR4 receptors ($p = 0.0005$, F (3, 8) = 19.44; Figure 6A,B) led to a high expression following NF-κB activation ($p = 0.0026$, F (3, 8) = 11.80; Figure 6A,C). Donepezil was assessed for its anti-inflammatory properties by examining intracellular TNF-alpha levels ($p = 0.0016$, F (3, 8) = 13.82; Figure 6A,D), as well as its antiapoptotic properties by measuring PI3K ($p = 0.0001$, F (3, 8) = 28.36; Figure 6A,E) and pAKT levels ($p < 0.0001$, F (3, 8) = 45.32; Figure 6A,F). Compared to the control group, these proteins were undetectable. Upon SAH induction in rats, the protein expression of TLR4, NF-κB, and TNF-α peaked,

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

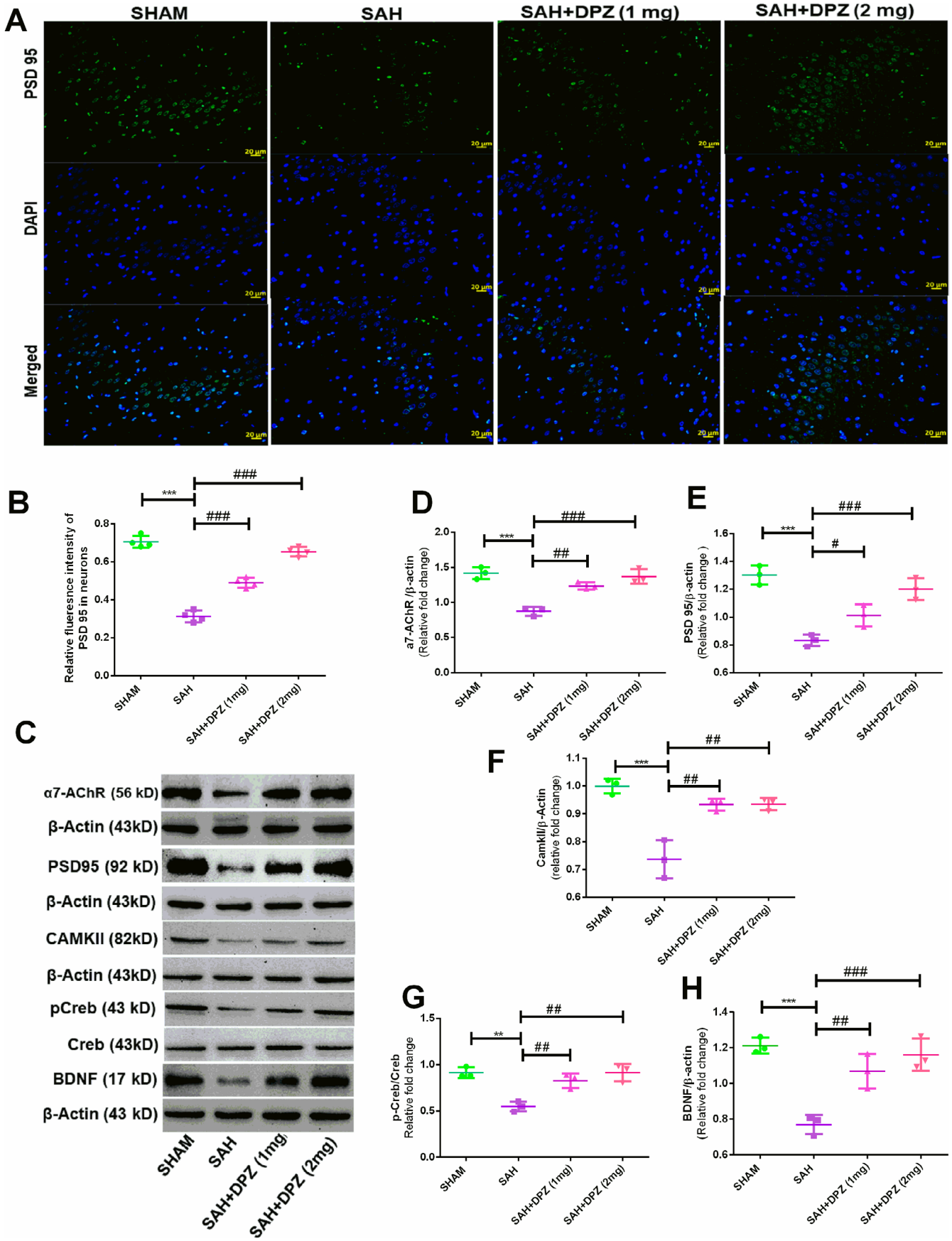


FIGURE 4 | Legend on next page.

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

FIGURE 4 | Donepezil enhances the expression of the α7 receptor along with improvements in memory markers PSD95, BDNF, CamkII, and Creb. (A–F) Representative western blot images and quantitative analyses of α7 nAChR, PSD95, BDNF, CamkII, and Creb from a time course study using samples obtained from the right hemisphere after SAH ($n = 3$ rats per group). (G, H) Representative microphotographs of immunofluorescence staining and quantification, showing the expression of PSD95 (green) in the damage neuron with DAPI (blue) between 48 and 72 h after SAH in different experimental groups ($n = 4$ rats per group). Expression was viewed at 20× magnification using a fluorescent microscope. Scale bar = 20 μm. Data are presented as means ± SD and were analyzed using a one-way ANOVA followed by the Tukey post hoc test. $^{**}p < 0.01$, $^{***}p < 0.001$ versus sham, $^{\#}p < 0.05$, $^{\#\#}p < 0.01$, $^{\#\#\#}p < 0.001$ versus SAH.

accompanied by a downregulation of PI3K/AKT protein levels. SAH rats were significantly more likely to express inflammatory proteins (TLR4, NF-B, and TNF-alpha) compared to control rats, along with decreased PI3K and pAKT levels, as shown in Figure 6A–F. Further, we observed that donepezil treatments significantly reduced levels of TLR4/NF-κB/TNF-α ($\#p < 0.05$, $\#\#p < 0.01$ or $\#\#\#p < 0.001$) and increased PI3K/pAKT levels ($\#p < 0.05$, $\#\#p < 0.01$ or $\#\#\#p < 0.001$) compared to SAH rats.

3.5 | Donepezil Ameliorates Increased GFAP and Iba-1 Expression in Post-SAH Rats

After observing signs of donepezil's ability to protect against inflammation, we confirmed this result by assessing the number of activated astrocytes and microglia in the vicinity of the clot site. At 48 h following SAH, there was a notable rise in the number of activated microglia, as identified by the iba1 antibody ($p < 0.0001$, $F(3, 12) = 132.2$; Figure 7A,B), and astrocytes, as detected by the GFAP antibody ($p < 0.0001$, $F(3, 12) = 242.3$; Figure 7C,D), in contrast to the control group. The administration of donepezil significantly decreased the expression of iba-1 and GFAP ($\#\#\#p < 0.001$) in rats having SAH.

3.6 | Effect of Post-Treatment of Donepezil on Mitochondrial ROS and MMP ($\Delta\Psi_m$) in SAH Rats

The endovascular perforation model of SAH showed a blood clot near the circle of Willis, caused by blood reperfusion injury, which increased mitochondrial ROS generation, leading to oxidative stress. To confirm ROS generation, DCFDH2 dye expression in mitochondria was measured using BDlyric. As compared to SHAM-operated animals, the SAH group had significantly higher ROS levels ($p < 0.0001$, $F(3, 20) = 72.84$; Figure 8A,B). The post-SAH rats group was treated with donepezil, and we found a significant reduction in ROS production. ROS generation was found to be the main cause of cellular membrane damage. It has been discovered that the generation of ROS decreased the Δψm and ATP production in post-SAH rats ($p = 0.0002$, $F(3, 8) = 25.41$; Figure 8C). The addition of TMRE, a fluorescent membrane potential indicator, to isolated mitochondria served as validation for the effect of ROS on mitochondrial membrane potential (Δψm). The mitochondrial membrane potential decreased with SAH injury, according to our results ($p < 0.0001$, $F(3, 20) = 73.70$; Figure 8D,E), but MMP was significantly recovered by donepezil treatment. These findings strengthen our hypothesis that donepezil works as a scavenging compound against ROS generation.

3.7 | Effect of Donepezil on Morphology and Quality Control of Mitochondria

Mitochondrial quality control is primarily regulated by the involvement of two proteins: dynamin-related GTPase (DRP1) and mitofusin1/2 (MFN1/2). To examine the impact of SAH induction on mitochondrial dynamics in rats, a western blotting analysis was conducted (Figure 9H–K). The results showed a significant increase in the translocation of Drp1 from the cytosol ($p = 0.0002$, $F(3, 8) = 24.02$; Figure 9H,I) to the mitochondria ($p = 0.0002$, F (3, 8) = 27.08 Figure 9H,J) in SAH-induced rats compared to rats who underwent a sham operation. The level of Mitofusin 2 (mfn2) was considerably reduced in the mitochondria of rats with SAH compared to the control group ($p = 0.0037$, $F(3, 8) = 10.61$; Figure 9H,K). It was shown that rats treated with donepezil inhibited the recruitment of drp1 into mitochondria and increased the level of mfn2 in mitochondria. To substantiate the hypothesis of mitochondrial dysfunction, the western blot analysis demonstrated that SAH directly impacts the stability of the electron transport chain (ETC) and mitochondrial respiratory complexes. Specifically, there was a more noticeable effect on the subunits CI ($p = 0.0051$, $F(3, 8) = 9.516$; Figure 9B,C), CII ($p = 0.0013$, $F(3, 8) = 14.78$; Figure 9B,D), and CIV ($p = 0.0025$, $F(3, 8) = 11.92$; Figure 9B,F). However, treatment with donepezil significantly mitigated the imbalance of mitochondrial respiratory complexes ($^{*}p < 0.05$, $^{**}p < 0.01$ or $\#p < 0.05$ or $\#\#p < 0.01$). To investigate the effectiveness of donepezil in protecting the brain's neurons and promoting the creation of new mitochondria, we utilized Electron microscopy to analyze any alterations in the structure of mitochondria, specifically looking for signs of fragmentation and fusion. TEM imaging in the SAH group demonstrated a notable increase in mitochondrial fragmentation and edema as compared to the sham group. Nevertheless, administration of donepezil led to a reduction in fragmentation and a boost in the framework of mitochondria in comparison to the SAH group (Figure 9A).

3.8 | Donepezil Reduces Mitochondrial-Mediated Cell Death in Post-SAH Rats

The proportion of pro- and anti-apoptotic Bcl-2 proteins is modified in response to stressful situations. Chronic stress exacerbates cellular apoptosis pathways, increasing the severity and decreasing the viability of organisms. The formation of ROS in cells can disturb the balance of ROS in mitochondria, which in turn triggers the apoptotic pathway. Following SAH, several apoptotic proteins including BCL-2, Cytochrome C, BAX, and cleaved caspase are triggered. Within 48 h following SAH, the caspase-dependent apoptotic cascade is triggered. This results

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

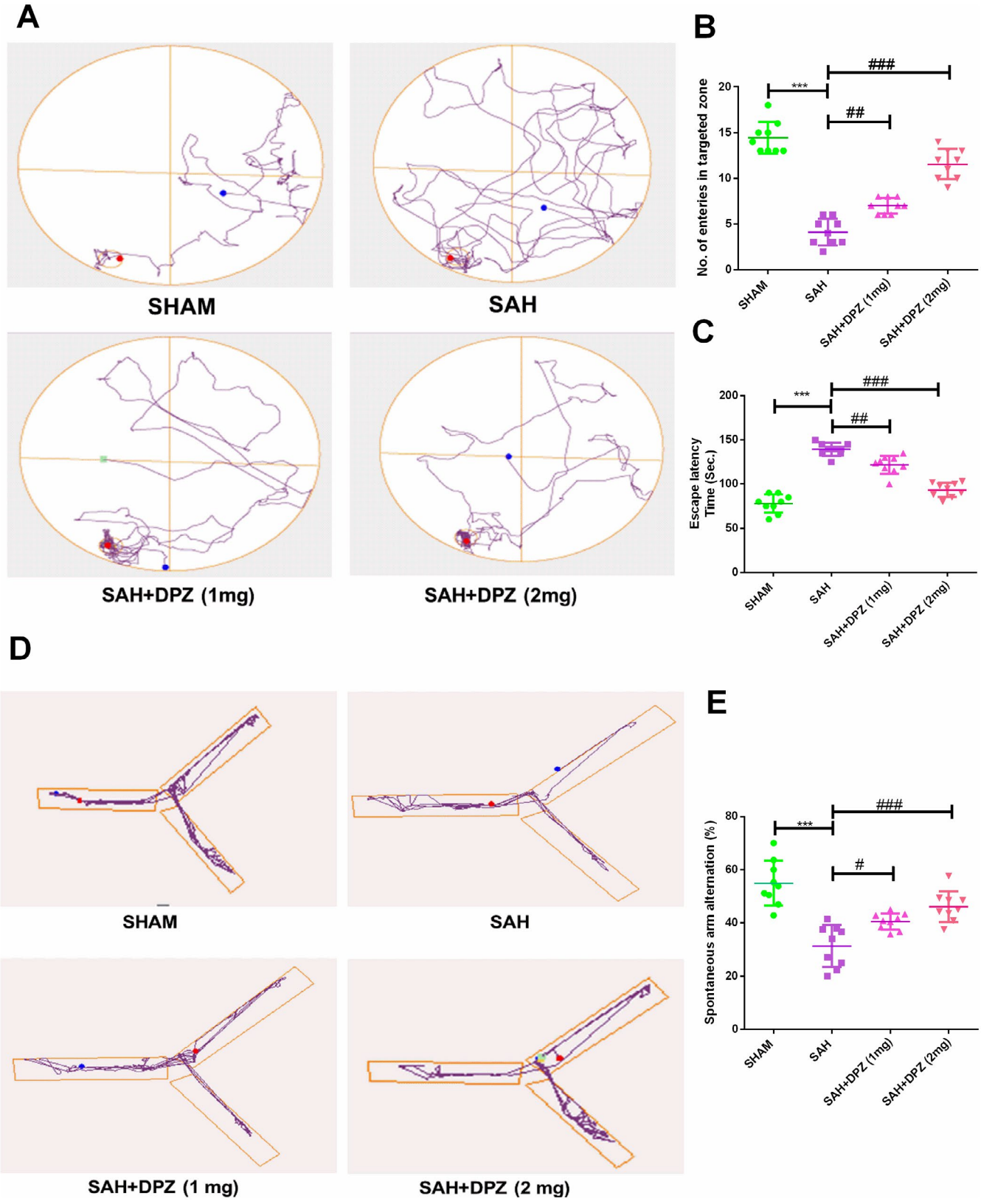


FIGURE 5 | Donepezil reduces anxiety and improves working memory in post-SAH rats. (A) Barnes maze schematic illustration. Rats in the SAH group had a significantly longer time to reach the escape box than rats in the sham group. However, sham rats reached the targeted hole faster than SAH rats. (B) Total number of hole entries. (C) Escape latency. (D, E) Schematic representation of the Y maze. The percent alternation was significantly higher for sham rats compared to SAH rats at 72 h post-surgery, and SAH rats alternated significantly less than sham rats. Experimental data are presented as means ± SD ($n = 9$ rats per group), and were analyzed using a one-way ANOVA followed by the Tukey post hoc test. $^{***}p < 0.001$ versus sham, $^{\#}p < 0.05$, $^{\#\#}p < 0.01$, $^{\#\#\#}p < 0.001$ versus SAH.

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

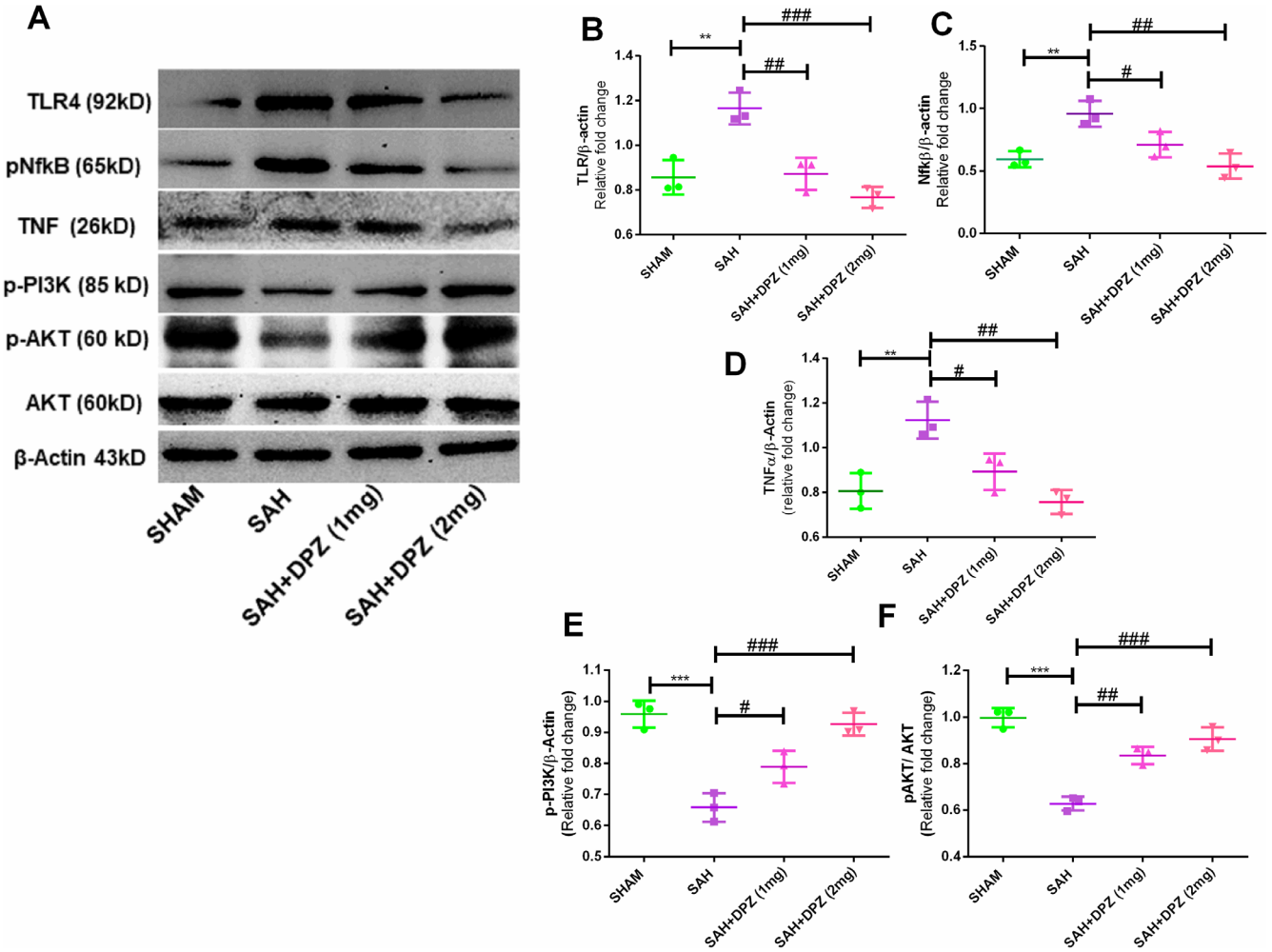


FIGURE 6 | Donepezil reduced the expression of neuroinflammatory markers and enhanced the expression of anti-apoptotic markers following SAH injury. (A) Representative western blot images. (B–F) Quantitative analyses of TLR4, NF-κβ, TNF-α, p-PI3K, pAKT ($n = 3$ rats per group). Data are presented as means ± SD and were analyzed using a one-way ANOVA followed by the Tukey post hoc test. $^{**}p < 0.01$, $^{***}p < 0.001$ versus sham, $^{\#}p < 0.05$, $^{\#\#}p < 0.01$, $^{\#\#\#}p < 0.001$ versus SAH.

in the translocation of Bax protein to the mitochondria through the mitochondrial membrane pore, causing damage to the mitochondria, and the subsequent release of Cytochrome C into the cytoplasm was verified using western blot analysis. In Figure 10A–F, the expression of BAX dramatically increases ($p = 0.0008$, F (3, 8) = 17.07 Figure 10A,C), whereas the levels of Cytochrome C significantly decrease in rats with SAH compared to animals who underwent a sham operation ($p = 0.0005$, F (3, 8) = 19.27; Figure 10D,F). The Western blot analysis revealed that donepezil caused a significant decrease in the movement of Bax protein ($p = 0.0006$, F (3, 8) = 18.18; Figure 10A,B) from cytosol to mitochondria, and also restored the expression levels of Cytochrome C in the mitochondria ($p < 0.0001$, F (3, 8) = 30.91; Figure 10D,E). To provide additional evidence for the anti-apoptotic function of donepezil, we noticed a notably increased Bax/Bcl-2 ratio ($p < 0.000$, F (3, 8) = 40.21; Figure 10G,H) and the production of activated cleaved caspase-3 ($p = 0.0003$, F (3, 8) = 22.36; Figure 10G,I) in whole cell lysates following SAH. Donepezil significantly reduced the proportion of Bax to Bcl-2 and the expression of cleaved caspase-3 in the basilar cortex following SAH.

3.9 | Donepezil Effects on Alterations in Neuron Morphology

Using hematoxylin and eosin staining, changes in the neuronal architecture of the cortex were investigated. In sham-operated mice, representative H and E pictures (Figure 11A–E) demonstrated normal neuronal morphology. SAH rats, on the other hand, showed increased vacuolated gaps, pyknotic neurons with reduced cytoplasm, and significant neuronal death. When SAH rats were compared to sham rats, the number of non-viable neurons rose considerably ($p < 0.0001$, F (3, 20) = 123.3 Figure 11A–E). When compared to SAH animals that were not treated, donepezil post-treatment significantly increased neuronal survivability and preserved the number of intact neurons ($^{\#\#\#}p < 0.001$).

4 | Discussion

Donepezil, an acetylcholinesterase inhibitor, has been demonstrated to reduce neuroinflammation and ROS generation after

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

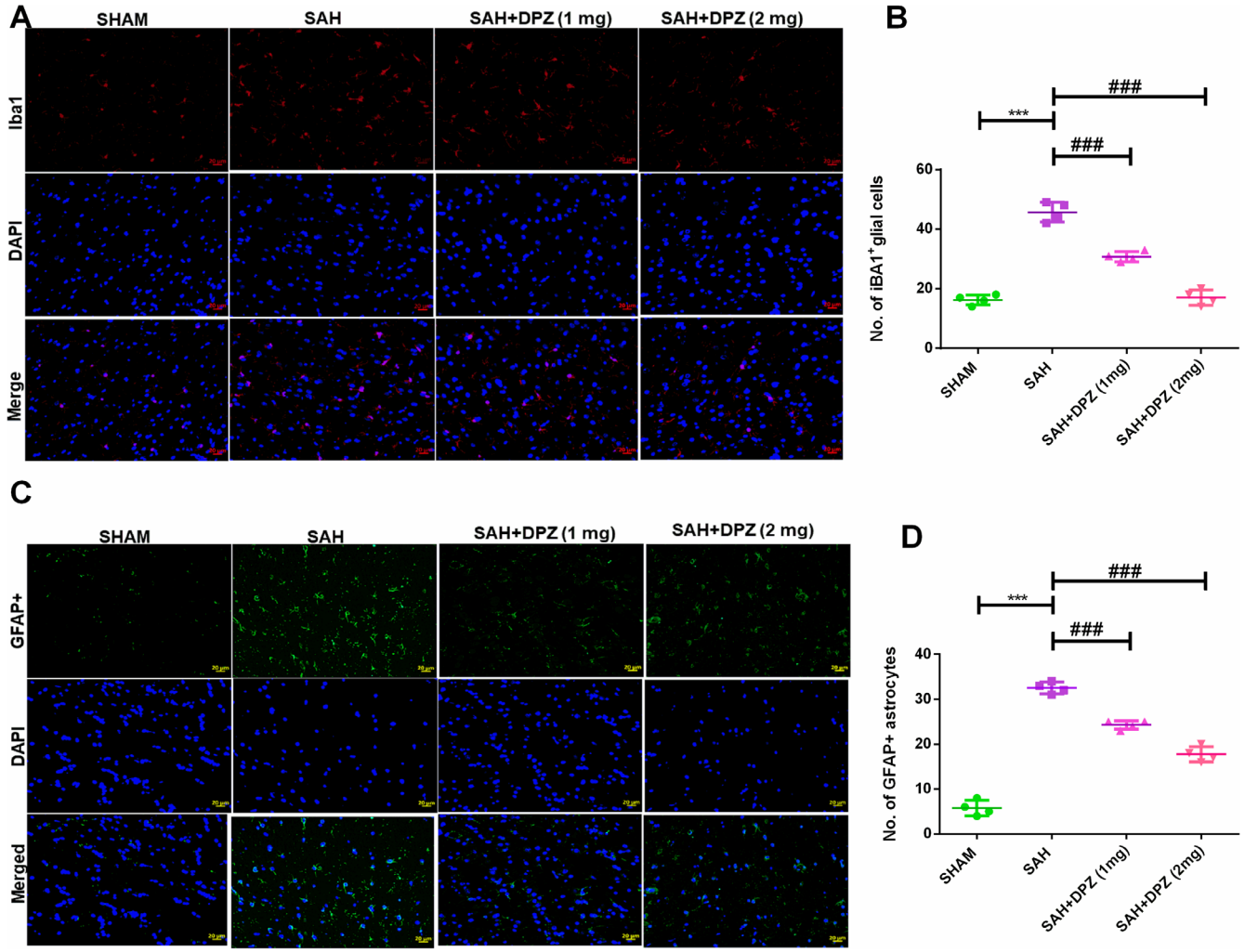


FIGURE 7 | Donepezil alleviates inflammation derived from glial cells and astrocytes induced by SAH. (A, B) Representative microphotographs and quantification of GFAP-positive astrocytes in the ipsilateral cerebral cortex of rat brain tissue, with GFAP (green) and DAPI (blue) immunofluorescence staining. (C, D) Representative microphotographs of immunofluorescence staining, showing Iba1-positive glial cells (red) with DAPI (blue), and their quantification at 48 h in different experimental groups ($n = 3–4$ rats per group). Data are presented as means ± SD and were analyzed using a one-way ANOVA followed by the Tukey post hoc test. $^{***}p < 0.001$ versus sham, $^{\#\#\#}p < 0.001$ versus SAH. Expression viewed at a 20× magnification using a fluorescent microscope. Scale bar = 20 μm.

brain haemorrhages. An FDA-approved cholinergic drug, donepezil was assessed for its therapeutic potential in subarachnoid hemorrhage (SAH). The decision to administer donepezil 4 h after SAH was based on multiple considerations. While neuroprotective drugs are often administered immediately following brain injury, we chose a 4-h window to better reflect clinical relevance, since it approximated the therapeutic timeframe available to patients following the onset of subarachnoid hemorrhage. Furthermore, donepezil's pharmacokinetics, including its ability to cross the blood–brain barrier and exert effects on acetylcholine levels and support its efficacy within this timeframe (Madani Neishaboori et al. 2021). The results of our study indicate that (a) SAH induction causes neurobehavioral deficits, particularly in the area of working memory and sensorimotor behavior. SAH rats treated with donepezil demonstrated improved functional outcomes, including working memory, postsynaptic density as well as prevention of neuroinflammation. (b) Donepezil regulates activated glial cells and astrocytes, while downregulating ROS that are responsible for improving outcomes after SAH. (c) Furthermore, donepezil enhanced the restoration of mitochondrial morphology and ultrastructure, regulated mitochondrial-mediated apoptosis and oxidative stress, preserved mitochondrial integrity, and promoted mitochondrial quality control in the brain after SAH.

Subarachnoid hemorrhage is the event where blood spreads into the subarachnoid space. This occurrence is associated with significant brain swelling and severe neurological deficits (Chen et al. 2016). The coagulation of blood cells results in the buildup of harmful compounds, leading to increased oxidative stress and subsequent neuronal apoptosis, ultimately resulting in persistent neurological impairments (Hamming et al. 2017). The formation of blood clots, which is measured using the SAH grading system, contributes to the severity of SAH in rats (Sugawara et al. 2008). Before euthanasia at either 24 or 72 h after surgery, the severity of brain edema, as evaluated by the amount of water in the brain, was found to be connected with various neurofunctional measures. These assessments included a compromise

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

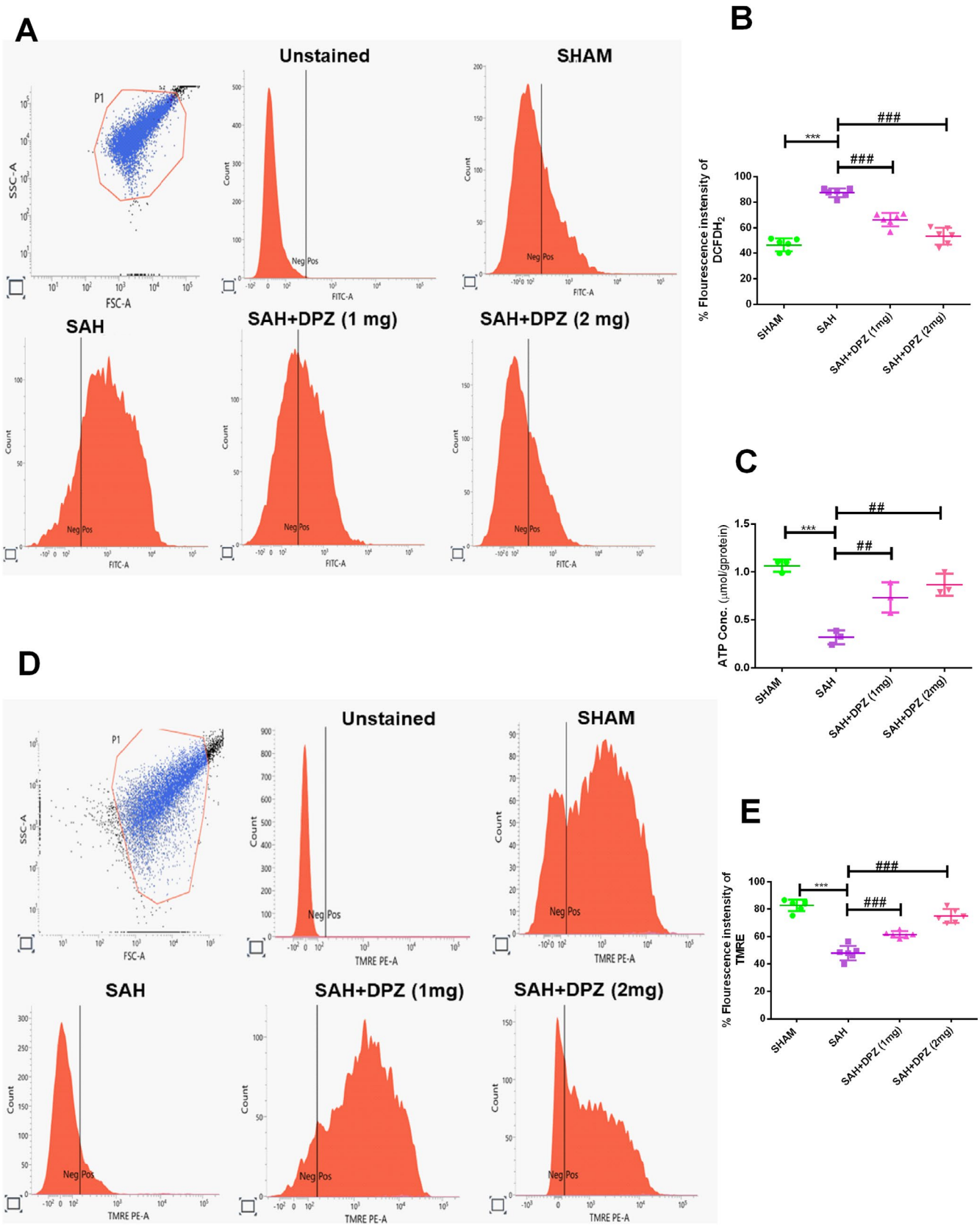


FIGURE 8 | Effect of donepezil on mitochondrial reactive oxygen species (ROS) and mitochondrial membrane potential (Δψm) in the cerebral cortex region following SAH. (A, B) Mitochondrial levels of ROS were evaluated by quantifying the percent variations in DCF fluorescence intensity, based on data obtained from 10000 events. (C) The level of ATP. (D-E) The estimation of Δψm was conducted by monitoring the percent changes in TMRE fluorescence intensity, while collecting data from 10000 events. Data are presented as means ± SD and were analyzed using a one-way ANOVA followed by the Tukey post hoc test ($n = 5–7$ rats per group) $^{***}p < 0.001$ versus sham, $^{\#\#}p < 0.01$, $^{\#\#\#}p < 0.001$ versus SAH.

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

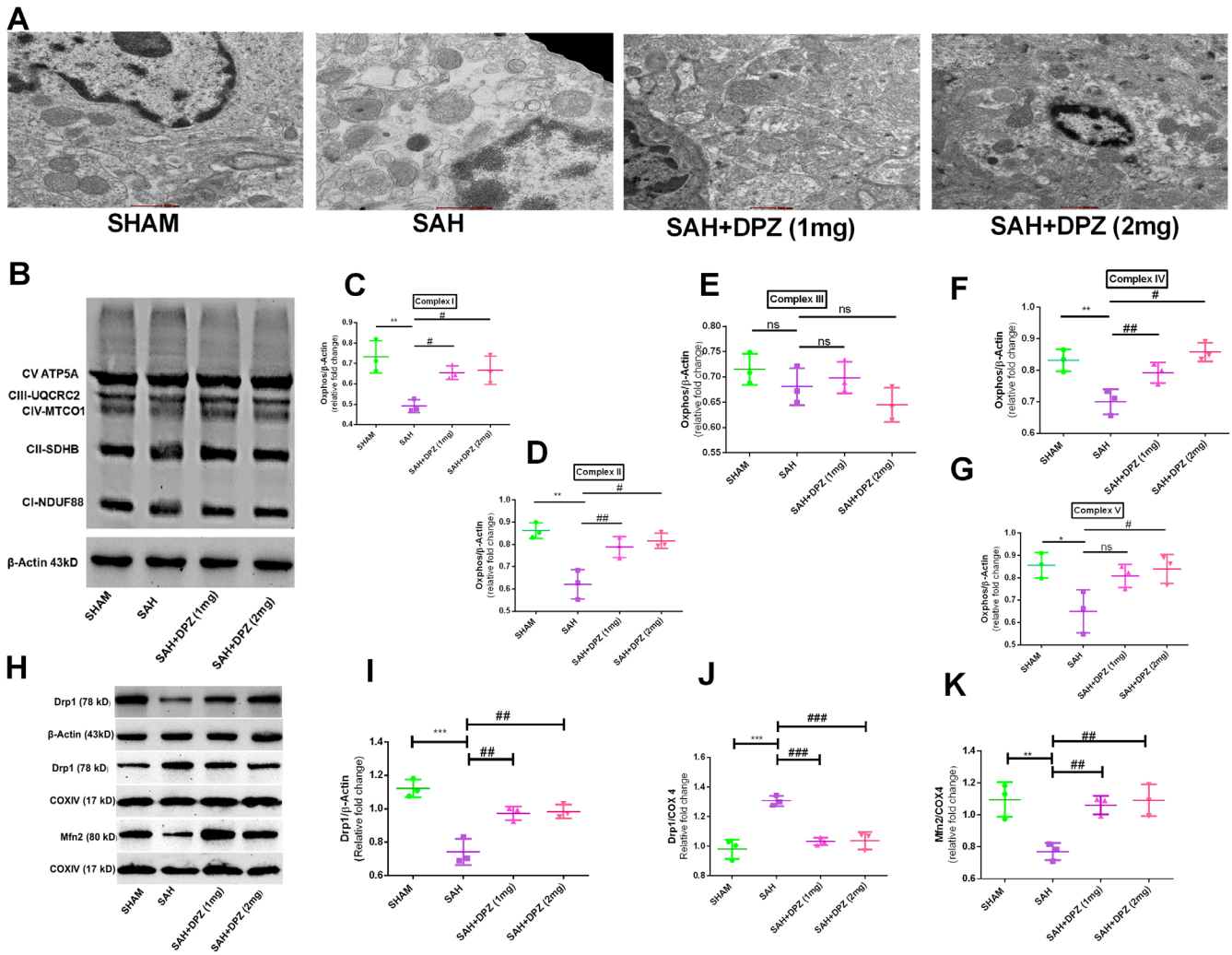


FIGURE 9 | Donepezil mitigated mitochondrial morphology alterations, respiratory complex impairment, and mitochondrial biogenesis disruption following SAH induction. (A) Transmission electron microscopy images display the morphological changes in mitochondria within brain tissues across different experimental groups (scale bar = 500 nm). (B) Immunoblot analysis of oxidative phosphorylation (OXPHOS) complexes. (C and G) Densitometric analysis of OXPHOS, normalized to β-actin. (H) Immunoblot analysis showing Drp1 translocation and mfn2 protein downregulation. (I–K) Densitometric analysis of Drp1 and mfn2, with relative optical densities of protein bands normalized to β-actin for the cytosol and COX-IV for the mitochondria ($n = 3$ rats per group). Experimental data are presented as means ± SD and were analyzed using a one-way ANOVA followed by the Tukey post hoc test. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ versus sham, $^{\#}p < 0.05$, $^{\#\#}p < 0.01$, $^{\#\#\#}p < 0.001$ versus SAH; ns(non-significant).

neuroscore, test for beam balance, gait impairment, rota rod performance, and grip strength (McBride et al. 2015). A higher degree of SAH (increased brain water content) led to decreased performance in these tests. Emerging data suggests that aquaporin 4 (AQP4), a protein responsible for transporting water in the brain, plays a vital role in regulating brain edema in different brain disorders. This underscores the significance of AQP4 in the progression and treatment of brain edema after SAH (Aoki et al. 2003). Our findings demonstrated that donepezil had a positive impact on the treatment outcomes in SAH-induced rats, leading to improvements in neurobehavioral deficits and motor coordination that worsened after surgery. In addition, donepezil effectively reduces the accumulation of blood clots and the increase in AQP4 caused by SAH. Previous research investigating the effects of Donepezil on ischemic stroke has shown that donepezil improved the loss of hippocampal neurons and cognitive impairments in Wistar rats after global cerebral ischemia, which is consistent with our findings (Madani Neishaboori et al. 2021b). In addition, galantamine, another agonist of the α7-AChR receptor, demonstrated comparable enhancements in functional outcomes, as indicated by a decrease in brain water content (Dienel et al. 2021).

The endovascular perforation model of SAH developed anxiety-like behavior that was detected at 72 h. The Barnes maze and Y maze were employed to quickly evaluate anxiety levels and working memory. Healthy rats generally exhibit elevated levels of arm alternation and reduced escape latency, which is indicative of strong working memory, in contrast to SAH animals. In the SAH perforation model, mice exhibited a slight elevation in alternation compared to the Sham group, as hippocampectomized animals developed a preference for one side and performed more effectively (Deacon and Rawlins 2006). The brain undergoes essential structural and functional rearrangement after intracerebral hemorrhage (ICH) to preserve brain plasticity (Tamakoshi et al. 2016). Enhanced synaptic function,

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

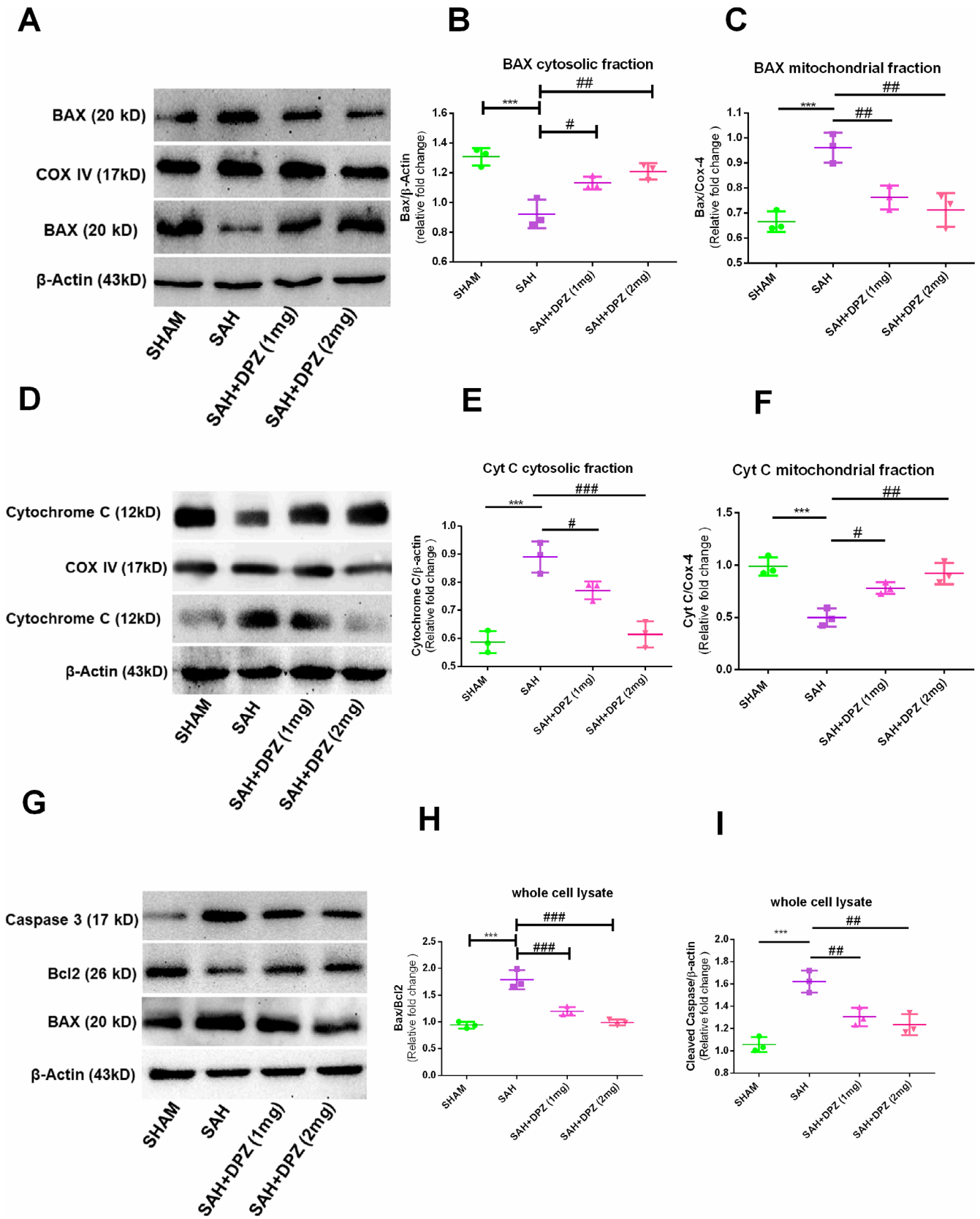


FIGURE 10 | Legend on next page.

which is linked to a higher quantity of synapses, facilitates neuron connectivity and inhibits post-ICH cognitive decline. The PSD-95 protein plays a crucial role in regulating synaptic maintenance following an ICH. Recent research suggests that PSD-95 experiences disruption following SAH surgery in rats (Shen et al. 2015). Similarly, BDNF protein together with PSD-95

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

FIGURE 10 | Donepezil treatment modulated the mitochondrial-dependent apoptotic pathway in a SAH model. (A) Immunoblots showing Bax translocation in cytosolic and mitochondrial fractions. (B, C) Scatter plot depicting the relative levels of Bax in cytosolic and mitochondrial fractions. (D) Immunoblots showing Cytochrome C translocation in mitochondrial and cytosolic fractions. (E, F) Scatter plot depicting the relative levels of Cytochrome C in cytosolic and mitochondrial fractions. (G) Relative levels of the Bax/Bcl-2 ratio and cleaved Caspase-3 in whole cell lysate. (H–I) Scatter plot depicting the relative levels of the Bax/Bcl-2 ratio and Caspase-3 in whole cell lysate ($n = 3$ rats per group). Data are presented as means ± SD and were analyzed using a one-way ANOVA followed by the Tukey post hoc test ***$p < 0.001$ versus sham, $^{\#}p < 0.05$, $^{\#\#}p < 0.01$, $^{\#\#\#}p < 0.001$ versus SAH.

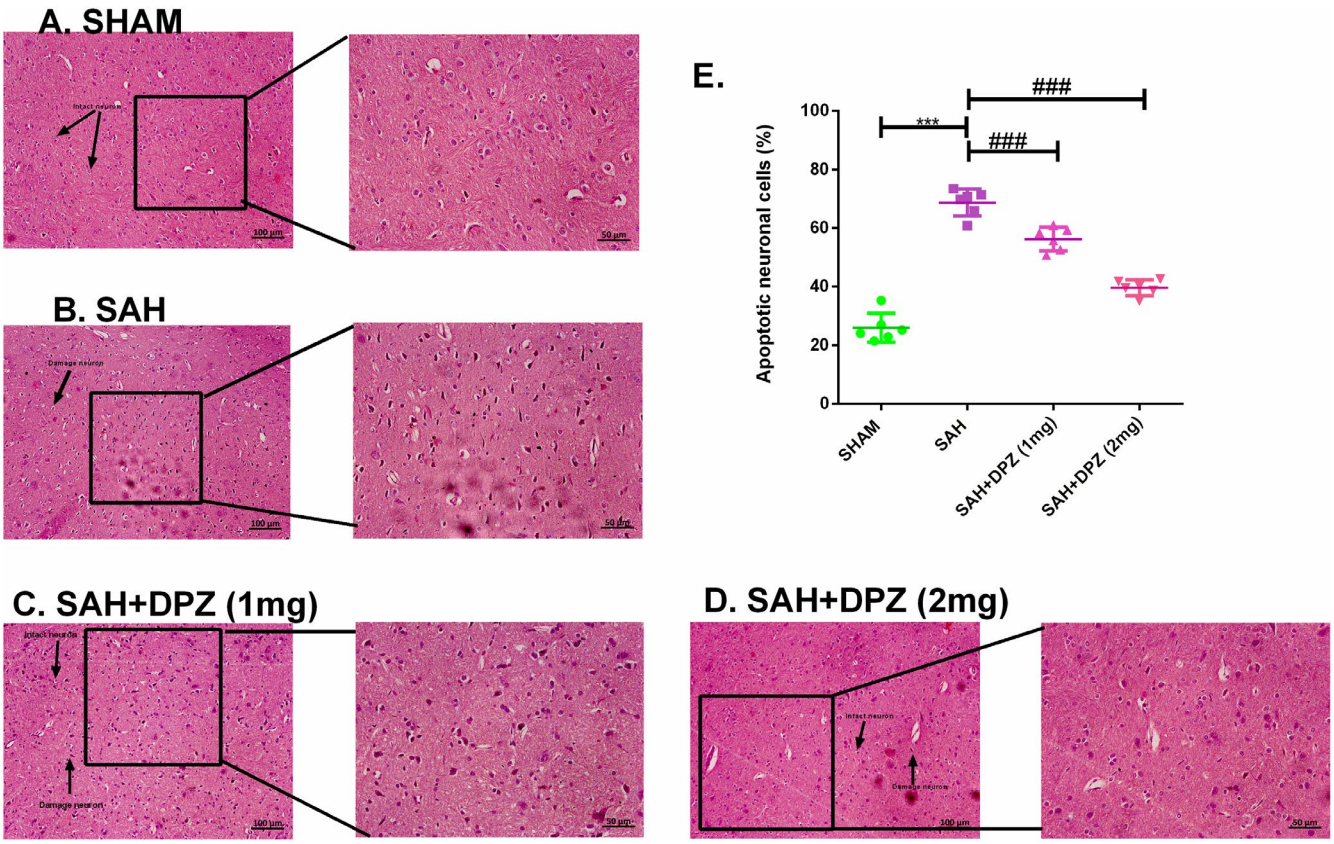


FIGURE 11 | Donepezil protects against histopathological abnormalities following SAH injury. (A–D) Coronal brain sections, stained with H and E, showing changes in the ipsilateral cerebral cortex region. (E) Scatter plot of H and E staining results across different experimental groups ($n = 6$ rats per group). Data are presented as means ± SD and were analyzed using a one-way ANOVA followed by the Tukey post hoc test ***$p < 0.001$ versus sham, $^{\#\#\#}p < 0.001$ versus SAH.

enhances synaptic plasticity and promotes the activation of the CaMKII-CREB pathway in mouse neurons (Hu et al. 2011).

Our research has uncovered that donepezil can increase the expression of PSD 95 protein and enhance the function of α7 nAChR receptors. Additionally, it promotes synaptic connection by activating memory markers such as CREB/CaMKII and BDNF protein in rats with SAH. In addition, donepezil demonstrated dose-dependent enhancements in escape latency and arm alternation in post-SAH rats. Prior research has shown that α7 nAChR overexpression can enhance the expression of PSD-95 in SH-SY5Y neuronal cells (Ren et al. 2020). A previous study reported that α7-nicotinic receptor activation enhances the growth of primary cultured rat cortical neurons (Takada-Takatori et al. 2008) and improves learning and memory through an adjunctive treatment strategy combining donepezil with PNU-120596 (Callahan et al. 2013). Furthermore, it has been shown to improve spatial memory deficits caused by temporary global cerebral ischemia in gerbils by elevating the expression of CaMKII-CREB (Min et al. 2012). Stimulation of presynaptic α7 nAChRs has been demonstrated to boost CaMKII activity, resulting in heightened Ca^{2+} entry and CaM production. This, in turn, may activate the CaMKII-CaM-CREB signaling pathway, leading to neuroprotective benefits and improved learning and memory abilities in mouse models (Wang et al. 2020).

SAH induction causes oxidative stress, which is then made worse by reperfusion injury and the overproduction of ROS within the cell. Overexposure to ROS causes the release of pro-inflammatory proteins, which in turn causes neuroinflammation. After SAH surgery, this disruption triggers astrocyte and glial cell activation, a significant early brain injury (EBI) event (Lucke-Wold et al. 2016). Through the activation of TLR4 receptors, SAH triggers the release of inflammatory proteins, including TNF-α and NF-κB, as well as activating anti-apoptotic pathways including PI3K and Akt. Reperfusion, on the other

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

hand, decreases these anti-apoptotic proteins action (Iadecola and Anrather 2011). Our research shows that donepezil therapy reduces the expression of Iba1 and GFAP in SAH damage. In SAH rats, the production of inflammatory mediators like TLR4, phosphorylated NF-κB, and TNF-α is downregulated by this therapy.

The prior report indicated that donepezil inhibits the nuclear translocation of NF-κB in LPS-stimulated macrophage cells, which lowers the production of proinflammatory mediators (Arikawa et al. 2016). Moreover, donepezil increases the expression of PI3K/Akt, which in turn triggers the activity of antioxidant enzymes, lowering oxidative stress and reducing the damage that ROS brings to neurons (Haraguchi et al. 2017). According to our findings, donepezil can significantly increase PI3K/Akt expression in post-SAH rats, suggesting that it may be used as a therapeutic agent for SAH-induced injury.

The maturation and development of neurons depend on mitochondrial dynamics. Oxidative damage during SAH causes mitochondrial permeability transition pores (PTPs) to open, which increases the release of ROS into the cytoplasm and eventually causes mitochondrial malfunction. An excess of ROS is produced as a result of this malfunction, upsetting the equilibrium between endogenous oxidants and antioxidants as well as the Drp1-mfn2 mediated fission and fusion process (Smith et al. 2013), (Xie et al. 2018). Consequently, oxidative alteration of proteins, lipids, and DNA occurs, further increasing ROS emission. The brain produces ATP by a process known as oxidative phosphorylation, or OXPHOS. OXPHOS downregulation is frequent in neurodegenerative disorders and impairs brain function (Hroudová et al. 2014). Better ATP availability is ensured by increased oxygen consumption, and this is important for preserving essential cellular functions and averting necrosis and apoptosis (Ali et al. 2022), (Kaushik et al. 2020). Our findings show that donepezil treatment successfully restores mitochondrial morphology, upregulates OXPHOS (C-I to C-V) protein expression, including ATP synthase activity, and reduces SAH-induced damage to the ETC. Donepezil has been shown to prevent oxidative stress and deficiencies in mitochondrial bioenergetics, which aligns with our findings, protecting mitochondrial integrity and biogenesis in the post-SAH brain (Ye et al. 2015).

The normal processes of mitochondrial fission and fusion are responsible for maintaining a functional and healthy mitochondrial network. Nevertheless, abnormal mitochondrial division facilitated by Drp1 results in the excessive formation of ROS and triggers the caspase-3-dependent process of cell death. This abnormal division also enables the movement of the proapoptotic Bcl-2 family member Bax from the cytosol inside the cell to the outer membrane of the mitochondria, increasing the release of cyt c into the cytosol inside the cell and encouraging cell death in different disease types (Clerc et al. 2014; Du et al. 2019).

Donepezil has demonstrated its efficacy as an anti-apoptotic drug in specific circumstances within the human brain (Sharifipour et al. 2014). Our hypothesis strongly indicates that donepezil has an immediate protective effect on mitochondrial function by preventing the movement of Bax. This is accomplished by enhancing the integrity of the mitochondrial membrane, thus restricting the release of cyt c from the mitochondria. In addition, donepezil decreases the Bax/Bcl-2 ratio and lowers the levels of cleaved caspase-3 in the brain tissue of rats after SAH. Similarly, the therapeutic approach of donepezil seems to mitigate the significant damage to the neuronal architecture that is commonly associated with SAH. In keeping with our findings, Madani Neishaboori et al. (2021) documented that donepezil can ameliorate the damage caused by ischemic stroke through the reduction of inflammation and enhancement of mitochondrial function, neurogenesis, and angiogenesis, highlighting its potential as a therapeutic agent.

There are different limitations in our investigation. Initially, the study concentrated on the immediate effects (between 24 and 72 h) of donepezil after SAH, highlighting the need for more research on long-term survival and functional recovery. Furthermore, although the involvement of PSD-95 in synaptic density and spatial memory has been recognized, a more in-depth investigation is needed to understand its exact molecular mechanisms. Furthermore, while we investigated the neuroprotective properties of donepezil concerning mitochondrial dysfunction, neuroinflammation, and neuronal death, a more comprehensive understanding of these molecular interconnections is necessary for therapeutic applications.

5 | Conclusion

In general, our idea suggests that donepezil helps to sustain the expression of PSD95 and promotes the levels of α7-AChR and BDNF-CaMKII. As a result, this restores neurobehavioral impairments and partially preserves anxiety and spatial memory. Our work presents novel data that show how donepezil can reinstate the dynamics of mitochondria, which helps protect neurons from damage caused by SAH at intervals of 24 and 72 h. This is achieved by minimizing a process called Drp1-mediated neuronal apoptosis, decreasing neuroinflammation, and maintaining the proper functioning of mitochondria. Thus, donepezil may be a promising therapeutic approach for averting and treating SAH.

Author Contributions

Ahmed Shaney Rehman: conceptualization investigation, methodology, software, data curation, formal analysis, validation, writing – original draft, writing – review and editing. **Ammar Tasleem:** methodology, investigation, software, formal analysis. **Nemat Ali:** methodology, software, formal analysis, funding acquisition, writing – review and editing. **Rehan Khan:** conceptualization, methodology, writing – review and editing. **Mohd. Salman:** resources, data curation. **Pravir Kumar:** supervision, writing – review and editing. **Suhel Parvez:** conceptualization, investigation, funding acquisition, writing – original draft, methodology, validation, visualization, writing – review and editing, software, formal analysis, project administration, data curation, supervision, resources.

Acknowledgements

This work was funded by the Indian Researchers Supporting Project, King Saud University, Riyadh, Saudi Arabia: Number (RSPD2025R940); the Department of Science and Technology, Ministry of Science and Technology, India, (Grant/Award Number: '[SR/PURSE Phase2/39 (C)]', '[SR/FST/LS-I/2017/05(C)]'); the Council of Scientific and Industrial Research, India (Grant/Award Number: '09/591/(0170)/2020-EMR-I').

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The authors are thankful to the Researchers Supporting Project Number (RSPD2025R940) King Saud University, Riyadh, Saudi Arabia for supporting this study.

Ethics Statement

The animal experiments were conducted in compliance with the established protocols of the Institutional Animal Ethics Committee (IAEC), which is officially recognized by the Committee for Control and Supervision of Experiments on Animals (CPCSEA) under registration no. 173/GO/Re/2000/CPCSEA.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The datasets supporting the findings of this study are included in the article.

Peer Review

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/jnc.70063.

References

Ali, M., H. Tabassum, M. M. Alam, and S. Parvez. 2022. "N-Acetyl-L-Cysteine Ameliorates Mitochondrial Dysfunction in Ischemia/Reperfusion Injury via Attenuating Drp-1 Mediated Mitochondrial Autophagy." *Life Sciences* 293: 120338.

Al-Khindi, T., R. L. Macdonald, and T. A. Schweizer. 2010. "Cognitive and Functional Outcome After Aneurysmal Subarachnoid Haemorrhage." *Stroke* 41, no. 8: 519–536.

Aoki, K., T. Uchihara, K. Tsuchiya, A. Nakamura, K. Ikeda, and Y. Wakayama. 2003. "Enhanced Expression of Aquaporin 4 in Human Brain With Infarction." *Acta Neuropathologica* 106: 121–124.

Arikawa, M., Y. Kakinuma, T. Noguchi, H. Todaka, and T. Sato. 2016. "Donepezil, an Acetylcholinesterase Inhibitor, Attenuates LPS-Induced Inflammatory Response in Murine Macrophage Cell Line RAW 264.7 Through Inhibition of Nuclear Factor Kappa B Translocation." *European Journal of Pharmacology* 789: 17–26.

Birks, J. S., and R. J. Harvey. 2018. "(2018) Donepezil for Dementia due to Alzheimer's Disease." *Cochrane Database of Systematic Reviews* 2018: CD001190.

Broadbelt, T., M. Mutlu-Smith, D. Carnicero-Senabre, T. C. Saido, T. Saito, and S.-H. Wang. 2022. "Impairment in Novelty-Promoted Memory via Behavioral Tagging and Capture Before Apparent Memory Loss in a Knock-In Model of Alzheimer's Disease." *Scientific Reports* 12: 22298.

Brown, M. R., P. G. Sullivan, K. A. Dorenbos, E. A. Modafferi, J. W. Geddes, and O. Steward. 2004. "Nitrogen Disruption of Synaptoneurosomes: An Alternative Method to Isolate Brain Mitochondria." *Journal of Neuroscience Methods* 137: 299–303.

Callahan, P. M., E. J. Hutchings, N. J. Kille, J. M. Chapman, and A. V. Terry. 2013. "Positive Allosteric Modulator of Alpha 7 Nicotinic-Acetylcholine Receptors, PNU-120596 Augments the Effects of Donepezil on Learning and Memory in Aged Rodents and Non-Human Primates." *Neuropharmacology* 67: 201–212. https://doi.org/10.1016/j.neuropharm.2012.10.019.

Chen, T., W. Wang, J.-R. Li, et al. 2016. "PARP Inhibition Attenuates Early Brain Injury Through NF-κB/MMP-9 Pathway in a Rat Model of Subarachnoid Haemorrhage." *Brain Research* 1644: 32–38. https://doi.org/10.1016/j.brainres.2016.05.005.

Chung, D. Y., F. Oka, G. Jin, et al. 2021. "Subarachnoid Haemorrhage Leads to Early and Persistent Functional Connectivity and Behavioral Changes in Mice." *Journal of Cerebral Blood Flow & Metabolism* 41, no. 5: 975–985. https://doi.org/10.1177/0271678X20940152.

Clerc, P., S. X. Ge, H. Hwang, et al. 2014. "Drp 1 Is Dispensable for Apoptotic Cytochrome *c* Release in Primed MCF 10 A and Fibroblast Cells but Affects Bcl–2 Antagonist-Induced Respiratory Changes." *British Journal of Pharmacology* 171: 1988–1999.

Connolly, E. S., A. A. Rabinstein, J. R. Carhuapoma, et al. 2012. "Guidelines for the Management of Aneurysmal Subarachnoid Haemorrhage." *Stroke* 43: 1711–1737.

Cui, H., A. Hayashi, H.-S. Sun, et al. 2007. "PDZ Protein Interactions Underlying NMDA Receptor-Mediated Excitotoxicity and Neuroprotection by PSD-95 Inhibitors." *Journal of Neuroscience* 27: 9901–9915.

Deacon, R. M. J., and J. N. P. Rawlins. 2006. "T-Maze Alternation in the Rodent." *Nature Protocols* 1, no. 1: 7–12. https://doi.org/10.1038/nprot.2006.2.

Dienel, A., R. A. Veettil, K. Matsumura, et al. 2021. "α7-Acetylcholine Receptor Signaling Reduces Neuroinflammation After Subarachnoid Haemorrhage in Mice." *Neurotherapeutics* 18: 1891–1904.

Du, J., P. Hang, Y. Pan, et al. 2019. "Inhibition of miR-23a Attenuates Doxorubicin-Induced Mitochondria-Dependent Cardiomyocyte Apoptosis by Targeting the PGC-1α/Drp1 Pathway." *Toxicology and Applied Pharmacology* 369: 73–81.

Egeto, P., R. Loch Macdonald, T. J. Ornstein, and T. A. Schweizer. 2018. "Neuropsychological Function After Endovascular and Neurosurgical Treatment of Subarachnoid Haemorrhage: A Systematic Review and Meta-Analysis." *Journal of Neurosurgery* 128, no. 3: 768–776. https://doi.org/10.3171/2016.11.jns162055.

Fabiana, G., T. Mathiesen, N. H. Diemer, and N.-A. Svendgaard. 2003. "Experimental Subarachnoid Haemorrhage: Subarachnoid Blood Volume, Mortality Rate, Neuronal Death, Cerebral Blood Flow, and Perfusion Pressure in Three Different Rat Models." *Neurosurgery* 52: 165–176.

Fanizzi, C., A. D. Sauerbeck, M. Gangolli, G. J. Zipfel, D. L. Brody, and T. T. Kummer. 2017. "Minimal Long-Term Neurobehavioral Impairments After Endovascular Perforation Subarachnoid Haemorrhage in Mice." *Scientific Reports* 7, no. 1: 7569. https://doi.org/10.1038/s41598-017-07701-y.

Gao, X., R. Li, L. Luo, C. Liao, H. Yang, and S. Mao. 2024. "Alpha-Asarone Ameliorates Neurological Dysfunction of Subarachnoid Hemorrhagic Rats in Both Acute and Recovery Phases via Regulating the CaMKII-Dependent Pathways." *Translational Stroke Research* 15: 476–494.

Garcia, J. H., S. Wagner, K.-F. Liu, and X. Hu. 1995. "Neurological Deficit and Extent of Neuronal Necrosis Attributable to Middle Cerebral Artery Occlusion in Rats." *Stroke* 26, no. 4: 627–635. https://doi.org/10.1161/01.str.26.4.627.

Hall, E. D. 1985. "High-Dose Glucocorticoid Treatment Improves Neurological Recovery in Head-Injured Mice." *Journal of Neurosurgery* 62: 882–887.

Hamming, A. M., A. Toorn, R. U. S. Van Der, et al. 2017. "Valproate Reduces Delayed Brain Injury in a Rat Model of Subarachnoid Haemorrhage." *Stroke* 48: 452–458.

Hanafy, K. A. 2013. "The Role of Microglia and the TLR4 Pathway in Neuronal Apoptosis and Vasospasm After Subarachnoid Haemorrhage." *Journal of Neuroinflammation* 10, no. 1: 868. https://doi.org/10.1186/1742-2094-10-83.

Haraguchi, Y., Y. Mizoguchi, M. Ohgidani, et al. 2017. "Donepezil Suppresses Intracellular Ca2+ Mobilization Through the PI3K Pathway

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

in Rodent Microglia." *Journal of Neuroinflammation* 14, no. 1: 258. https://doi.org/10.1186/s12974-017-1033-0.

Hroudová, J., N. Singh, and Z. Fišar. 2014. "Mitochondrial Dysfunctions in Neurodegenerative Diseases: Relevance to Alzheimer's Disease." *BioMed Research International* 2014: 1–9.

Hu, X., L. Ballo, L. Pietila, et al. 2011. "BDNF-Induced Increase of PSD-95 in Dendritic Spines Requires Dynamic Microtubule Invasions." *Journal of Neuroscience* 31: 15597–15603.

Hütter, B. O., J. M. Gilsbach, and I. Kreitschmann. 1994. "Is There a Difference in Cognitive Deficits After Aneurysmal Subarachnoid Haemorrhage and Subarachnoid Haemorrhage of Unknown Origin?" *Acta Neurochirurgica* 127: 129–135.

Iadecola, C., and J. Anrather. 2011. "The Immunology of Stroke: From Mechanisms to Translation." *Nature Medicine* 17: 796–808.

Jeon, H., J. Ai, M. Sabri, et al. 2009. "Neurological and Neurobehavioral Assessment of Experimental Subarachnoid Haemorrhage." *BMC Neuroscience* 10: 103.

Kaushik, P., M. Ali, H. Tabassum, and S. Parvez. 2020. "Post-Ischemic Administration of Dopamine D2 Receptor Agonist Reduces Cell Death by Activating Mitochondrial Pathway Following Ischemic Stroke." *Life Sciences* 261: 118349.

Kim, J., and W. Shin. 2014. "How to Do Random Allocation (Randomization)." *Clinics in Orthopedic Surgery* 6, no. 1: 103–109. https://doi.org/10.4055/cios.2014.6.1.103.

Koch, R. E., C. C. Josefson, and G. E. Hill. 2017. "Mitochondrial Function, Ornamentation, and Immunocompetence." *Biological Reviews* 92: 1459–1474.

Kogure, T., M. Sumitani, K. Ikegami, et al. 2017. "Donepezil, an Acetylcholinesterase Inhibitor, Can Attenuate Gabapentinoid-Induced Somnolence in Patients With Neuropathic Pain: A Retrospective Chart Review." *Journal of Pain & Palliative Care Pharmacotherapy* 31: 4–9.

Krishna, K. V., R. N. Saha, G. Singhvi, and S. K. Dubey. 2018. "Pre-Clinical Pharmacokinetic-Pharmacodynamic Modelling and Biodistribution Studies of Donepezil Hydrochloride by a Validated HPLC Method." *RSC Advances* 8: 24740–24749.

Lin, S., Q. Yin, Q. Zhong, et al. 2012. "Heme Activates TLR4-Mediated Inflammatory Injury via MyD88/TRIF Signaling Pathway in Intracerebral Haemorrhage." *Journal of Neuroinflammation* 9, no. 1: 548. https://doi.org/10.1186/1742-2094-9-46.

Loomis, C., A. Stephens, R. Janicot, U. Baqai, L. Drebushenko, and J. Round. 2020. "Identification of MAGUK Scaffold Proteins as Intracellular Binding Partners of Synaptic Adhesion Protein Slitrk2." *Molecular and Cellular Neuroscience* 103: 103465. https://doi.org/10.1016/j.mcn.2019.103465.

Lucke-Wold, B., A. Logsdon, B. Manoranjan, et al. 2016. "Aneurysmal Subarachnoid Haemorrhage and Neuroinflammation: A Comprehensive Review." *International Journal of Molecular Sciences* 17: 497.

Madani Neishaboori, A., S. Nasseri Maleki, M. Saberi Pirouz, S. Golmohammadi, D. Nazarinia, and N. Aboutaleb. 2021. "Donepezil Attenuates Injury Following Ischaemic Stroke by Stimulation of Neurogenesis, Angiogenesis, and Inhibition of Inflammation and Apoptosis." *Inflammopharmacology* 29, no. 1: 153–166. https://doi.org/10.1007/s10787-020-00769-5.

McBride, D. W., Y. Wang, P. Sherchan, J. Tang, and J. H. Zhang. 2015. "Correlation Between Subacute Sensorimotor Deficits and Brain Water Content After Surgical Brain Injury in Rats." *Behavioural Brain Research* 290: 161–171.

Mendes, C. S., I. Bartos, Z. Márka, T. Akay, S. Márka, and R. S. Mann. 2015. "Quantification of Gait Parameters in Freely Walking Rodents." *BMC Biology* 13: 50.

Meyer, J. N., T. C. Leuthner, and A. L. Luz. 2017. "Mitochondrial Fusion, Fission, and Mitochondrial Toxicity." *Toxicology* 391: 42–53.

Min, D., X. Mao, K. Wu, et al. 2012. "Donepezil Attenuates Hippocampal Neuronal Damage and Cognitive Deficits After Global Cerebral Ischemia in Gerbils." *Neuroscience Letters* 510: 29–33.

Ongnok, B., T. Khuanjing, T. Chunchai, et al. 2021. "Donepezil Provides Neuroprotective Effects Against Brain Injury and Alzheimer's Pathology Under Conditions of Cardiac Ischemia/Reperfusion Injury." *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* 1867, no. 1: 165975. https://doi.org/10.1016/j.bbadis.2020.165975.

Pace, A., S. Mitchell, E. Casselden, et al. 2018. "A Subarachnoid Haemorrhage-Specific Outcome Tool." *Brain* 141: 1111–1121.

Pei, H., Y. Yang, H. Zhao, et al. 2016. "The Role of Mitochondrial Functional Proteins in ROS Production in Ischemic Heart Diseases." *Oxidative Medicine and Cellular Longevity* 2016: 1–8.

Qin, X., Y. Zhao, J. Gong, et al. 2019. "Berberine Protects Glomerular Podocytes via Inhibiting Drp1-Mediated Mitochondrial Fission and Dysfunction." *Theranostics* 9: 1698–1713.

Rashmi, K. C., M. Harsha Raj, M. Paul, K. S. Girish, B. P. Salimath, and H. S. Aparna. 2019. "A New Pyrrole Based Small Molecule From Tinospora Cordifolia Induces Apoptosis in MDA-MB-231 Breast Cancer Cells via ROS Mediated Mitochondrial Damage and Restoration of p53 Activity." *Chemico-Biological Interactions* 299: 120–130. https://doi.org/10.1016/j.cbi.2018.12.005.

Regnier-Golanov, A. S., M. Gulinello, M. S. Hernandez, E. V. Golanov, and G. W. Britz. 2022. "Subarachnoid Haemorrhage Induces Sub-Acute and Early Chronic Impairment in Learning and Memory in Mice." *Translational Stroke Research* 13, no. 4: 625–640. https://doi.org/10.1007/s12975-022-00987-9.

Ren, J.-M., S.-L. Zhang, X.-L. Wang, Z.-Z. Guan, and X.-L. Qi. 2020. "Expression Levels of the α7 Nicotinic Acetylcholine Receptor in the Brains of Patients With Alzheimer's Disease and Their Effect on Synaptic Proteins in SH-SY5Y Cells." *Molecular Medicine Reports* 22: 2063–2075.

Rosell, A., V. Agin, M. Rahman, et al. 2013. "Distal Occlusion of the Middle Cerebral Artery in Mice: Are we Ready to Assess Long-Term Functional Outcome?" *Translational Stroke Research* 4: 297–307.

Rouanet, C., and G. S. Silva. 2019. "Aneurysmal Subarachnoid Haemorrhage: Current Concepts and Updates." *Arquivos de Neuro-Psiquiatria* 77: 806–814.

Sarter, M., G. Bodewitz, and D. N. Stephens. 1988. "Attenuation of Scopolamine-Induced Impairment of Spontaneous Alternation Behaviour by Antagonist but Not Inverse Agonist and Agonist?-Carbolines." *Psychopharmacology* 94: 491–495.

Schneider, U. C., R. Xu, and P. Vajkoczy. 2018. "Inflammatory Events Following Subarachnoid Haemorrhage (SAH)." *Current Neuropharmacology* 16: 1385–1395.

Sehba, F. A., J. Hou, R. M. Pluta, and J. H. Zhang. 2012. "The Importance of Early Brain Injury After Subarachnoid Haemorrhage." *Progress in Neurobiology* 97: 14–37.

Sharifipour, M., E. Izadpanah, B. Nikkhoo, et al. 2014. "A New Pharmacological Role for Donepezil: Attenuation of Morphine-Induced Tolerance and Apoptosis in Rat Central Nervous System." *Journal of Biomedical Science* 21: 1–9.

Sharma, A., K. Liaw, R. Sharma, Z. Zhang, S. Kannan, and R. M. Kannan. 2018. "Targeting Mitochondrial Dysfunction and Oxidative Stress in Activated Microglia Using Dendrimer-Based Therapeutics." *Theranostics* 8: 5529–5547.

Shen, H., Z. Chen, Y. Wang, et al. 2015. "Role of Neurexin-1β and Neuroligin-1 in Cognitive Dysfunction After Subarachnoid Haemorrhage in Rats." *Stroke* 46: 2607–2615.

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Shen, H., T. Kihara, H. Hongo, et al. 2010. "Neuroprotection by Donepezil Against Glutamate Excitotoxicity Involves Stimulation of α7 Nicotinic Receptors and Internalization of NMDA Receptors." *British Journal of Pharmacology* 161, no. 1: 127–139. https://doi.org/10.1111/j.1476-5381.2010.00894.x.

Smith, J. A., S. Park, J. S. Krause, and N. L. Banik. 2013. "Oxidative Stress, DNA Damage, and the Telomeric Complex as Therapeutic Targets in Acute Neurodegeneration." *Neurochemistry International* 62: 764–775.

Soltani, Z., M. Khaksari, E. Jafari, M. Iranpour, and N. Shahrokhi. 2015. "Is Genistein Neuroprotective in Traumatic Brain Injury?" *Physiology & Behavior* 152: 26–31.

Suarez, J. I., R. W. Tarr, and W. R. Selman. 2006. "Aneurysmal Subarachnoid Haemorrhage." *New England Journal of Medicine* 354: 387–396.

Sugawara, T., R. Ayer, V. Jadhav, and J. H. Zhang. 2008. "A New Grading System Evaluating Bleeding Scale in Filament Perforation Subarachnoid Haemorrhage Rat Model." *Journal of Neuroscience Methods* 167, no. 2: 327–334. https://doi.org/10.1016/j.jneumeth.2007.08.004.

Takada-Takatori, Y., T. Kume, Y. Ohgi, et al. 2008. "Mechanisms of α7-Nicotinic Receptor Up-Regulation and Sensitization to Donepezil Induced by Chronic Donepezil Treatment." *European Journal of Pharmacology* 590, no. 1-3: 150–156. https://doi.org/10.1016/j.ejphar.2008.06.027.

Tamakoshi, K., K. Kawanaka, H. Onishi, Y. Takamatsu, and K. Ishida. 2016. "Motor Skills Training Improves Sensorimotor Dysfunction and Increases Microtubule-Associated Protein 2 mRNA Expression in Rats With Intracerebral Haemorrhage." *Journal of Stroke and Cerebrovascular Diseases* 25, no. 8: 2071–2077. https://doi.org/10.1016/j.jstrokecerebrovasdis.2016.05.007.

Turan, N., R. A. Heider, M. Nadeem, et al. 2020. "Neurocognitive Outcomes in a Cisternal Blood Injection Murine Model of Subarachnoid Haemorrhage." *Journal of Stroke and Cerebrovascular Diseases* 29: 105249.

Uzair, M., C. Singhal, A. Ali, et al. 2024. "Myocardial Ischemia-Reperfusion Injury Released Cellular Fibronectin Containing Domain A (CFN-EDA): A Destructive Positive Loop Amplifying Arterial Thrombosis Formation and Exacerbating Myocardial Reperfusion Injury." *Thrombosis Research* 238: 117–128.

Wang, X.-L., Y.-X. Deng, Y.-M. Gao, et al. 2020. "Activation of α7 nAChR by PNU-282987 Improves Synaptic and Cognitive Functions Through Restoring the Expression of Synaptic-Associated Proteins and the CaM-CaMKII-CREB Signaling Pathway." *Aging* 12: 543–570.

Wang, Z., Z. Chen, J. Yang, et al. 2019. "Treatment of Secondary Brain Injury by Perturbing Postsynaptic Density Protein-95-NMDA Receptor Interaction After Intracerebral Haemorrhage in Rats." *Journal of Cerebral Blood Flow & Metabolism* 39, no. 8: 1588–1601. https://doi.org/10.1177/0271678X18762637.

Weissgerber, T. L., N. M. Milic, S. J. Winham, and V. D. Garovic. 2015. "Beyond Bar and Line Graphs: Time for a New Data Presentation Paradigm." *PLoS Biology* 13: e1002128.

Xian, P., Y. Hei, R. Wang, et al. 2019. "Mesenchymal Stem Cell-Derived Exosomes as a Nanotherapeutic Agent for Amelioration of Inflammation-Induced Astrocyte Alterations in Mice." *Theranostics* 9: 5956–5975.

Xie, L., F. Shi, Z. Tan, Y. Li, A. M. Bode, and Y. Cao. 2018. "Mitochondrial Network Structure Homeostasis and Cell Death." *Cancer Science* 109: 3686–3694.

Yang, C., T. Li, H. Xue, et al. 2019. "Inhibition of Necroptosis Rescues SAH-Induced Synaptic Impairments in Hippocampus via CREB-BDNF Pathway." *Frontiers in Neuroscience* 12: 990. https://doi.org/10.3389/fnins.2018.00990.

Ye, C. Y., Y. Lei, X. C. Tang, and H. Y. Zhang. 2015. "Donepezil Attenuates Aβ-Associated Mitochondrial Dysfunction and Reduces Mitochondrial Aβ Accumulation In Vivo and In Vitro." *Neuropharmacology* 95: 29–36.

Zhang, T., P. Wu, E. Budbazar, et al. 2019. "Mitophagy Reduces Oxidative Stress via Keap1 (Kelch-Like Epichlorohydrin-Associated Protein 1)/Nrf2 (Nuclear Factor-E2-Related Factor 2)/PHB2 (Prohibitin 2) Pathway After Subarachnoid Haemorrhage in Rats." *Stroke* 50, no. 4: 978–988. https://doi.org/10.1161/STROKEAHA.118.021590.

Zhang, X., Q. Wu, Q. Zhang, et al. 2017. "Resveratrol Attenuates Early Brain Injury After Experimental Subarachnoid Haemorrhage via Inhibition of NLRP3 Inflammasome Activation." *Frontiers in Neuroscience* 11: 611.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

14714159, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/jnc.70063 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Plasma Physics and
Controlled Fusion

PURPOSE-LED
PUBLISHING™

PAPER

Effect of Alfvén wave dispersion on localization and current sheet formation in solar wind near reconnection region

To cite this article: Ritu *et al* 2025 *Plasma Phys. Control. Fusion* **67** 055015

View the article online for updates and enhancements.

You may also like

- Improved scaling of the scrape-off layer particle flux width by the Bayes' theorem on EAST
D C Liu, X Liu, L Wang et al.
- Fast physics-based launcher optimization for electron cyclotron current drive
N A Lopez, A Alieva, S A M McNamara et al.
- Semi-analytical model of ion cyclotron resonance heating antenna-plasma coupling and wave propagation in hot magnetized plasma
Claudia Salvia, Alessandro Cardinali, Silvio Ceccuzzi et al.

This content was downloaded from IP address 14.139.251.98 on 01/05/2025 at 07:21

IOP Publishing Plasma Physics and Controlled Fusion
Plasma Phys. Control. Fusion **67** (2025) 055015 (11pp) https://doi.org/10.1088/1361-6587/adcae9

Effect of Alfvén wave dispersion on localization and current sheet formation in solar wind near reconnection region

Ritu[1], Jyoti[1], Suresh C Sharma[1,*] and R P Sharma[2]

[1] Department of Applied Physics, Delhi Technological University, Shahbad Daulatpur, Bawana Road, Delhi 110042, India
[2] Department of Energy Science and Engineering, Indian Institute of Technology, Delhi, Hauz Khas, New Delhi 110016, India

E-mail: suresh321sharma@gmail.com

Received 11 September 2024, revised 25 March 2025
Accepted for publication 9 April 2025
Published 23 April 2025

Abstract

Observations show that dispersive Alfvén waves (DAWs) play a significant role in the heating and generation of turbulence at reconnection sites. The present work investigates the nonlinear evolution of DAWs and the related turbulence with pre-existing magnetic islands in solar wind. The turbulence is driven by nonlinear interactions, including wave–wave coupling and ponderomotive force, in the presence of pre-existing magnetic islands. Using computer modeling and numerical techniques, the resulting dynamical equation is solved. Space integration is performed using a pseudo-spectral method, while time integration is carried out using a finite difference scheme. The simulation findings validate the presence of turbulence and demonstrate the spatiotemporal evolution of the DAW localized structures and current sheets. The scale sizes of the localized structures have been calculated using a semi-analytic model and it has been demonstrated that these scale sizes are dependent on the characteristics of magnetic islands and ponderomotive nonlinearity. Thermal tail formation in energetic ions has been investigated based on the power-law scaling of turbulence generation.

Keywords: magnetic reconnection, turbulence, localization, reconnection rate

1. Introduction

Many spacecraft observations have revealed the turbulent state of electromagnetic fluctuations in the solar wind regime. Additionally, the solar wind may be utilized as an endless, natural laboratory for investigating a variety of scientific phenomena occurring in plasma turbulence that are difficult to replicate in a laboratory [1–5]. Smaller scales of energy exchange are facilitated by the wave energy cascade created by the turbulence in the solar wind. This cascading of energy is thought to play a significant role in the creation of solar winds and the acceleration of high-energy particles. Solar wind turbulence exhibits $k^{-5/3}$ scaling in the inertial range, regardless of energy injection into the system or energy dissipation at small scales. Energy is dissipated until the fluctuation-scale size approaches the order of the proton inertial length (λ_i) or the ion gyro-radius (ρ_s) after being injected into this region from the energy injection region (corresponding scaling k^{-1}). This region is regarded as a large-scale region where the description of magnetohydrodynamics (MHD) may be fully described [3, 6].

When Landau damping and other phenomena start to dominate, the MHD description of turbulence becomes invalid and changes in turbulence characteristics can be observed. The spectrum also steepens in comparison with inertial scaling. Moreover, the turbulent spectrum at the kinetic scale ($k\lambda_i > 1$) can be described using two models. One possible approach is

* Author to whom any correspondence should be addressed.

© 2025 IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

the cascading of whistler waves or kinetic Alfvén waves [7–9]. Another approach is the localization of parallel propagating dispersive Alfvén waves (DAWs). Alfvén waves are usually non-dispersive in nature (i.e. the speed of these waves remains the same as the change in wavelength) but when the frequency of Alfvén waves becomes comparable to the ion cyclotron frequency, then the speed of Alfvén waves also depends on their wavelength, showing dispersive behavior. Many researchers have shown the dispersion of Alfvén waves with the insertion of the Hall term [10–13].

Furthermore, Alfvén waves are the most prominent phenomenon in magnetized plasmas. Specifically, nonlinear Alfvén waves have a significant impact on a range of plasma processes both in the laboratory and in space, and these processes include turbulence, reconnection and heating [14–18]. In addition to the different space phenomena, it has been determined that magnetic reconnection is essential to the dynamics of turbulence [19, 20]. An inherent plasma characteristic, also present in solar wind plasmas, is the development of magnetic islands brought about by the dynamic processes of magnetic reconnection and turbulence [21, 22]. Magnetic islands can take on different shapes, such as spherical or elongated forms. The acceleration of particles in solar wind plasmas might be significantly aided by these magnetic islands. The energy transmission between various length scales may be impacted by the influence of magnetic islands on the wave dynamics [23–25].

DAWs are widely recognized in the literature for their involvement in turbulence generation and are highly promising for magnetic reconnection [26–28]. Consequently, it is possible for turbulence and magnetic reconnection, two basic phenomena, to coexist and have an impact on one another. We examine the two phenomena (i.e. turbulence and reconnection) in the framework of DAWs, which acquire the properties of dispersion due to the wave's finite frequency correction [27–31]. The finite frequency refers to the consideration of effects that arise when the frequency of Alfvén waves is not negligibly small compared to the ion cyclotron frequency. To fully comprehend space plasma dynamics, it becomes necessary to study the relationship between turbulent dissipation and reconnection. The formation of coherent structures and current sheets has been studied by Perri *et al* [32, 33] in the dissipation range (i.e. at smaller length scales). Their findings support the presence of current sheets and localized structures at sub-proton scale sizes, where energy exchange over the dissipation range might occur via magnetic reconnection. Hence, it becomes necessary to further study whether dissipation over small length scales occurs as a result of reconnection or other factors such as instabilities which cause the generation of localized/coherent or both. The purpose of the current analysis is to explore this question given the aforementioned context. Using WIND spacecraft data at 1 AU, Voros *et al* [34] showed that turbulence generation in solar wind is due to reconnection outflows. The present study is motivated by recent observations [32, 34–36] where the Hall effect, pressure tensor and other physical conditions affect the reconnecting electric field, current sheet formation and their scale size. In this work we consider a simple model where the effect of the Hall term is studied on the localized structures and current sheet formation by numerical simulations and analytical models. Pseudo-spectral and finite difference methods are used for space integration and time domain integration, respectively [37]. For this, we consider a situation where Alfvén waves are present at the reconnection site. Various types of waves, including right- and left-handed polarized Alfvén waves, kinetic Alfvén waves, and torsional Alfvén waves, have been observed at the reconnection site [38–40], and due to the Hall effect, they become dispersive. In solar wind turbulence, right-handed waves can also survive the shift into the intermediate dispersive region at higher wavenumbers [10, 11, 13, 41]. For this reason, the main focus of the present study will be on right-handed polarized DAWs. It is also discovered that the Hall term, due to the notable rise in the reconnection rate, has a significant impact on collisionless plasmas [42–44]. The turbulence in the solar wind and its related consequences are diversified due to the wide range of frequencies and amplitudes of Alfvén waves.

The present work investigates the temporal phase of nonlinear evolution of right-handed DAWs at reconnection sites that cause current sheet formation, localization and turbulence. The objective of this study is to analyze the dynamics of DAWs, the reconnection rate and the evolution of current sheets for two cases, as follows: case I, when the Hall term (i.e. ω/ω_{ci}) is present; and case II, when the Hall term is absent (i.e. $\omega/\omega_{ci} \to 0$) in the presence of a fully developed chain of magnetic islands and ponderomotive nonlinearity. The obtained results demonstrate consistency with earlier reported observations [32, 34, 35]. Both models investigate structures and dynamics in solar wind turbulence, emphasizing the role of coherent structures such as magnetic islands and current sheets.

This paper is arranged as follows: The dynamics of DAWs and ion acoustic waves (IAWs) are covered in section 2. The numerical simulation methods and their outcomes are examined in section 3. In addition, we examine the semi-analytical model in section 4 to comprehend the fundamental physics behind the localization. Section 5 provides a summary and conclusion of the present study.

2. Formulation of wave equations

Let us consider the dynamics of a DAW having finite frequency and amplitude propagating parallel to the magnetic field. Alfvén waves are treated as non-dispersive, meaning their velocity is constant and does not depend on wavelength. However, when the wave frequency approaches a significant fraction of the ion cyclotron frequency, dispersive effects come into play. We also consider a medium having an ambient magnetic field (B_{0z}) along the z-axis and the magnetic islands present at the reconnection sites owing to the perturbation in the background magnetic field by δB_y. The dependence of the

magnetic field of the islands in the y-direction is considered to be much larger than their dependence in the x-direction, i.e. $\delta B_y \gg \delta B_x$. The total magnetic field (B) under the influence of which the DAW is propagating becomes $\vec{B} = B_{0z}\hat{k} + \delta B_y \hat{j}$. The DAW also exhibits transverse instability [45].

By using Faraday and Ampere laws, one may obtain [46]

$$\nabla^2 \vec{E} - \nabla\left(\nabla.\vec{E}\right) = \frac{1}{c^2}\frac{\partial^2 \vec{D}}{\partial t^2}, \tag{1}$$

where c is the speed of light in vacuum. Using the well-known relation of the dielectric tensor, $\tilde{\varepsilon}$, and electric field, $\vec{E}$, with $\vec{D}$, i.e. $\vec{D} = \tilde{\varepsilon}\vec{E}$, taking $\vec{\nabla}.\vec{D} = 0$, and further taking field variation as $e^{-i\omega t}$, equation (1) in its component form can be written as follows:

$$\frac{\partial^2 E_x}{\partial y^2} + \frac{\partial^2 E_x}{\partial z^2} - \frac{\partial^2 E_z}{\partial x \partial z} - \frac{\partial^2 E_y}{\partial x \partial y} = \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\left(\tilde{\varepsilon}_{xx}E_x + \tilde{\varepsilon}_{xy}E_y\right) \tag{2}$$

$$\frac{\partial^2 E_y}{\partial x^2} + \frac{\partial^2 E_y}{\partial z^2} - \frac{\partial^2 E_x}{\partial y \partial x} - \frac{\partial^2 E_z}{\partial y \partial z} = \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\left(\tilde{\varepsilon}_{yx}E_x + \tilde{\varepsilon}_{yy}E_y\right) \tag{3}$$

$$\frac{\partial E_z}{\partial z} = -\frac{1}{\tilde{\varepsilon}_{zz}}\left[\tilde{\varepsilon}_{xx}\frac{\partial E_x}{\partial x} + \tilde{\varepsilon}_{xy}\frac{\partial E_y}{\partial x} - \tilde{\varepsilon}_{xy}\frac{\partial E_x}{\partial y} + \tilde{\varepsilon}_{yy}\frac{\partial E_y}{\partial y}\right]. \tag{4}$$

Defining $E_- = E_x - iE_y$ and $E_+ = E_x + iE_y$ as the right and left modes of a circularly polarized wave. The following coupled wave equation can be obtained by using equations (2)–(4) as

$$\begin{aligned}
&\frac{\partial^2 E_+}{\partial z^2} + \frac{1}{2}\left(1 + \frac{\tilde{\varepsilon}_{+00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial^2 E_+}{\partial x^2} + \frac{\partial^2 E_+}{\partial y^2}\right) + \frac{1}{2}\left(-1 + \frac{\tilde{\varepsilon}_{-00}}{\varepsilon_{0zz}}\right) \\
&\quad \times \left(\frac{\partial^2 E_-}{\partial x^2} + \frac{\partial^2 E_-}{\partial y^2}\right) - i\left(1 - \frac{\tilde{\varepsilon}_{-00}}{\varepsilon_{0zz}}\right) \times \frac{\partial^2 E_-}{\partial x \partial y} \\
&= \frac{\tilde{\varepsilon}_{+0}}{c^2}\frac{\partial^2}{\partial t^2}E_+ . \frac{\partial^2 E_-}{\partial z^2} + \frac{1}{2}\left(1 + \frac{\tilde{\varepsilon}_{-00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial^2 E_-}{\partial x^2} + \frac{\partial^2 E_-}{\partial y^2}\right) \\
&\quad + \frac{1}{2}\left(-1 + \frac{\tilde{\varepsilon}_{+00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial^2 E_+}{\partial x^2} + \frac{\partial^2 E_+}{\partial y^2}\right) - i\left(1 - \frac{\tilde{\varepsilon}_{+00}}{\varepsilon_{0zz}}\right) \\
&\quad \times \frac{\partial^2 E_+}{\partial x \partial y} = \frac{\tilde{\varepsilon}_{-0}}{c^2}\frac{\partial^2}{\partial t}E_- .
\end{aligned} \tag{5}$$

where $\tilde{\varepsilon}_{+00} = \frac{\omega_{pe}^2}{\omega_{ce}(\omega - \omega_{ci})}$ and $\tilde{\varepsilon}_{-00} = \frac{\omega_{pe}^2}{\omega_{ce}(\omega + \omega_{ci})}$ represent the linear parts of the dielectric, $\tilde{\varepsilon}_{+0} = \frac{\omega_{pe}^2}{\omega_{ce}(\omega - \omega_{ci})}\left(1 - 2\frac{\delta B_y}{B_{0z}} + \frac{\delta n}{n_0}\right)$, $\tilde{\varepsilon}_{-0} = \frac{\omega_{pe}^2}{\omega_{ce}(\omega + \omega_{ci})}\left(1 - 2\frac{\delta B_y}{B_{0z}} + \frac{\delta n}{n_0}\right)$ and $\varepsilon_{0zz} = 1 - \frac{\omega_{pe}^2}{\omega^2}$. ω_{cf} denotes the cyclotron frequency of species f (i for ion and e for electron), $\omega_{cf} = eB_{0z}/m_f c$. Further, $\omega_{pf} = \sqrt{4\pi n_0 e^2/m_f}$ is the plasma frequency and n_0 and δn represent the background density and perturbation in the background density, respectively.

Verscharen *et al* [41] observed that the right- and left-handed modes, having frequencies of the order of the ion cyclotron frequency (i.e. $\omega \approx \omega_{ci}$) or wavelengths of the order of $1/\lambda_i$ (where $\lambda_i = c/\omega_{pi}$), can co-exist in the solar wind. At higher wave numbers, left modes undergo cyclotron resonance [47, 48] and make the system complex. Only right-handed waves survive and are able to reach small wavenumbers ($k\lambda_i > 1$). Therefore, we will focus only on the right-handed mode for a better understanding of the system.

For the right-handed mode of a circularly polarized DAW, $E_+ = 0$, and using the relation $\vec{B} = \nabla \times \vec{A}$along with Faraday's law, the following equation is obtained

$$\begin{aligned}
&\frac{\partial^2 A_-}{\partial z^2} + \frac{1}{2}\left(1 + \frac{\tilde{\varepsilon}_{-00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial^2 A_-}{\partial x^2} + \frac{\partial^2 A_-}{\partial y^2}\right) \\
&= \frac{1}{c^2}\frac{\omega_{pe}^2}{\omega_{ce}(\omega + \omega_{ci})}\left(1 - 2\frac{\delta A_z}{A_{0y}} + \frac{\delta n}{n_0}\right)\frac{\partial^2}{\partial t^2}A_- .
\end{aligned} \tag{6}$$

$\frac{\delta B_y}{B_{0z}} = \frac{\delta A_z}{A_{0y}}$ has been obtained by using the above given relation between magnetic field and vector potential and A_- is the magnetic vector potential of the DAW.

2.1. IAW dynamics

Consider an IAW having low frequency and propagation parallel to the magnetic field present in the background. The fundamental equations governing the dynamics of the wave are as follows:

(i) Equation of motion

$$\frac{\partial v_f}{\partial t} = \frac{q_f \vec{E}}{m_f} + \frac{q_f}{c m_f}\left(\vec{v}_f \times \vec{B}\right) + \frac{\vec{F}_f}{m_f} - \frac{T_f}{m_f}\nabla\frac{n_f}{n_0} \tag{7}$$

(ii) Equation of continuity

$$\frac{\partial n_f}{\partial t} + \nabla\left(n_f \vec{v}_f\right) = 0 \tag{8}$$

(iii) Ponderomotive force

$$\vec{F}_f = \frac{q_f}{c}\left(\vec{v}_f^{\,*} \times \vec{B}\right) - m_f\left(\vec{v}_f^{\,*}.\vec{\nabla}\right)\vec{v}_f \tag{9}$$

where n_f is the density, m_f is the mass, $\vec{v}_f$is the velocity and q_f is the charge of the species 'f' of plasma ($f = e$ for electrons and i for ions). Furthermore, in terms of the scalar potential, the electric field is defined as $\vec{E} = -\vec{\nabla}\phi$. The thermal velocity of an electron has been considered to be very much larger than the phase velocity of a wave. The scalar potential can be written as [18]

$$\phi = \frac{T_e \delta n}{e n_0} - \frac{F_{ez}}{ik_z e}. \tag{10}$$

Upon using equation (10) and the components of equation of motion, we will obtain the following form

$$\frac{\partial v_{ix}}{\partial t} = \frac{\Omega^2 \rho_s^2}{n_0}\frac{\partial n}{\partial x} - \frac{\Omega^2 k_x F_{ez}}{\omega_{ci}^2 m_i k_z} - \frac{\Omega^2 F_{ix}}{\omega_{ci}^2 m_i} - \frac{i\Omega F_{iy}}{\omega_{ci} m_i} \tag{11}$$

$$\frac{\partial v_{iz}}{\partial t} = -\frac{c_s^2}{n_0}\frac{\partial n}{\partial z} + \frac{(F_{ez} + F_{iz})}{m_i}. \tag{12}$$

Now, using equations (11) and (12), the equation of continuitytakes the following form

$$\frac{\partial^2 n}{\partial t^2} - \rho_s^2 \frac{\partial^2}{\partial t^2}\frac{\partial^2 n}{\partial x^2} - c_s^2 \frac{\partial^2 n}{\partial z^2} = \frac{\Omega^2 n_0}{\omega_{ci}^2 m_i}\frac{\partial F_{ix}}{\partial x} + \frac{i\Omega n_0}{m_i \omega_{ci}}\frac{\partial F_{iy}}{\partial x}$$

$$+\frac{\Omega^2 n_0 k_x}{\omega_{ci}^2 m_i k_z}\frac{\partial F_{ez}}{\partial x}-\frac{n_0}{m_i}\frac{\partial\left(F_{ez}+F_{iz}\right)}{\partial z}. \tag{13}$$

Here, $n=n_0+\delta n$, where n and δn represent the modified number density and modified density, respectively. Ω is the frequency of the IAW and its value is much less the frequency of the DAW. The constants are defined as follows: $c_s=\sqrt{(T_e+T_i)/m_i}$ is the ion sound speed and $\rho_s=c_s/\omega_{ci}$ is the ion acoustic gyro-radius (here, 's' in subscript refers to the sound scale or sound speed). The ponderomotive force components (i.e. F_{ix},F_{iy},F_{iz} and F_{ez}) are obtained using equation (9) and have the form:

$$F_{ix}=-\frac{e^2\omega^2}{8m_i\left(\omega^2-\omega_{ci}^2\right)^2}\left[2-\frac{\omega_{ci}}{\Omega}\left(3-\frac{\omega_{ci}^2}{\Omega^2}\right)\right]\frac{\partial|A_-|^2}{\partial x} \tag{14}$$

$$F_{iy}=-\frac{ie^2}{16m_i\left(\omega^2-\omega_{ci}^2\right)^2}\left[\left(\omega^2-\omega_{ci}^2\right)^2+\left(1+\frac{\omega_{ci}}{\omega}\right)\left(\omega^2-\omega_{ci}^2\right)\right]\times\frac{\partial|A_-|^2}{\partial z} \tag{15}$$

$$F_{iz}=-\frac{e^2}{8m_i\left(\omega^2-\omega_{ci}^2\right)}\left(1-\frac{\omega_{ci}}{\omega}\right)\frac{\partial|A_-|^2}{\partial z} \tag{16}$$

and

$$F_{ez}=-\frac{e^2}{8m_e\left(\omega^2-\omega_{ce}^2\right)}\left(1-\frac{\omega_{ce}}{\omega}\right)\frac{\partial|A_-|^2}{\partial z}. \tag{17}$$

Using ponderomotive force components (i.e. equations (14)–(17) in equation (13), this equation will take the following form $\frac{\partial^2 n}{\partial t^2}-\rho_s^2\frac{\partial^2}{\partial t^2}\frac{\partial^2 n}{\partial x^2}-c_s^2\frac{\partial^2 n}{\partial z^2}=-\frac{e^2\Omega^2}{16m_i^2\omega_{ci}^2}\left[\frac{2(1+\omega_{ce}/\omega)m_i}{(\omega^2-\omega_{ce}^2)m_e}+\frac{\omega^2}{(\omega^2-\omega_{ci}^2)^2}\left\{2-\frac{\omega_{ci}}{\omega}\left(3-\frac{\omega_{ci}^2}{\omega^2}\right)\right\}\right]\frac{\partial^2|A_-|^2}{\partial x^2}$

$$-\frac{e^2\left(1-\omega/\omega_{ci}\right)}{8m_i^2\left(\omega_{ci}^2-\omega^2\right)}\frac{\partial^2|A_-|^2}{\partial z^2}. \tag{18}$$

The above equation represents the time-dependent variation of the acoustic density perturbation for the effect of ponderomotive nonlinearity, which is determined by the DAW.

Now, by taking the steady-state approximation, density perturbation takes the form

$$\frac{\delta n}{n_0}=\exp\left(-\alpha_0|A_-|^2\right)-1 \tag{19}$$

where $\alpha_0=\frac{\omega_{pi}^2}{32\pi n_0 T(\omega_{ci}^2-\omega^2)}\left(1-\frac{\omega}{\omega_{ci}}\right)$.

Further, equation (6) is solved semi-analytically by considering the wave solution [49] as follows

$$A_-=A\left(x,y,z,t\right)\mathrm{e}^{-\mathrm{i}(\omega t-k_- z)}. \tag{20}$$

where A is the complex amplitude of the DAW and $k_-=\frac{\omega}{c}\tilde{\varepsilon}_{-00}^{1/2}$ is the wave number. Substituting equations (19) and (20) in equation (6) and further assuming $\partial_z A\ll k_+ A$, the following equation is obtained

$$\frac{2\mathrm{i}}{\omega}\frac{\partial A}{\partial t}+\frac{2\mathrm{i}}{k_-}\frac{\partial A}{\partial z}+\frac{1}{2k_-^2}\left(1+\frac{\tilde{\varepsilon}_{-00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial^2 A}{\partial x^2}+\frac{\partial^2 A}{\partial y^2}\right)-2\frac{\delta A_z}{A_{0y}}A+\left(\exp\left(\alpha_0|A|^2-1\right)\right)A=0. \tag{21}$$

The dimensionless form of equation (21) can be written as follows

$$\mathrm{i}\frac{\partial A}{\partial t}+\mathrm{i}C_1\frac{\partial A}{\partial z}+C_2\left(\frac{\partial^2 A}{\partial x^2}+\frac{\partial^2 A}{\partial y^2}\right)+\left(2\frac{\delta A_z}{A_0}-\frac{\delta n}{n_0}\right)A=0 \tag{22}$$

where C_1 and C_2 are constants. They are defined as $C_1=\frac{\omega t_n}{k_- z_n}$ and $C_2=\frac{\omega t_n}{4k_-^2}\left(1+\frac{\varepsilon_{-00}}{\varepsilon_{0zz}}\right)$. $x_n=\lambda_i, n_n=n_0, z_n=\frac{2}{k_-}, t_n=\frac{2}{\omega}$ and $A_n=\left(\frac{8m_i^2\omega^2(\omega_{ci}^2-\omega^2)}{k_-^2 e^2(1-\omega/\omega_{ci})}\right)^{1/2}$ are normalization parameters.

Weak and slowly changing field fluctuations are caused by the previously existing chain of magnetic islands. Magnetic reconnection has been considered within the framework of magnetic islands. The magnetic island profile [50] is considered as

$$\frac{\delta A_z}{A_0}=-\frac{x^2}{2L^2}+0.5\cos ky. \tag{23}$$

Using equation (23) in equation (22), the dynamical equation of the DAW is obtained. This will be useful for a basic understanding of localized structures and current sheet formation under the influence of magnetic islands at the reconnection sites. Finally, we have

$$\mathrm{i}\frac{\partial A}{\partial t}+\mathrm{i}C_1\frac{\partial A}{\partial z}+C_2\left(\frac{\partial^2 A}{\partial x^2}+\frac{\partial^2 A}{\partial y^2}\right)+\left(2\left(-\frac{x^2}{2L^2}+0.5\cos ky\right)-\frac{\delta n}{n_0}\right)A=0. \tag{24}$$

Equation (24) represents the nonlinear dynamical equation of the DAW which will be solved further by numerical simulations. The dynamical equations of the DAW will be solved further by using solar wind parameters.

It should also be emphasized that the values of ion inertial length and ion gyroradius ($\rho_s=c_s/\omega_{ci}$) are approximately equal [51].

3. Numerical simulation

A modified nonlinear Schrödinger model is developed for solving equation (24). Equation (24) is solved numerically, in a 3D periodic spatial domain of size $(2\pi/\alpha_x)\times(2\pi/\alpha_y)\times(2\pi/\alpha_z)$ and grid points $(128\times128\times128)$ with step size of (5×10^{-5}). The perturbations in the wave number of the DAW are $\alpha_x=\alpha_y=\alpha_z=0.2$. The system resolution is $\Delta x=\Delta y=\Delta z=\frac{2\pi}{\alpha n'}$, where n' represents the grid points and $\alpha=0.2$ is a dimensionless quantity. A numerical methodology is employed in which time evolution is handled using the finite difference method combined with a predictor–corrector scheme, while spatial integration is performed using a pseudo-spectral method. The simplicity of the finite difference method over other approaches, including the finite element and finite volume methods, make it significantly beneficial. Finite difference methods take less time for a given grid. Computationally,

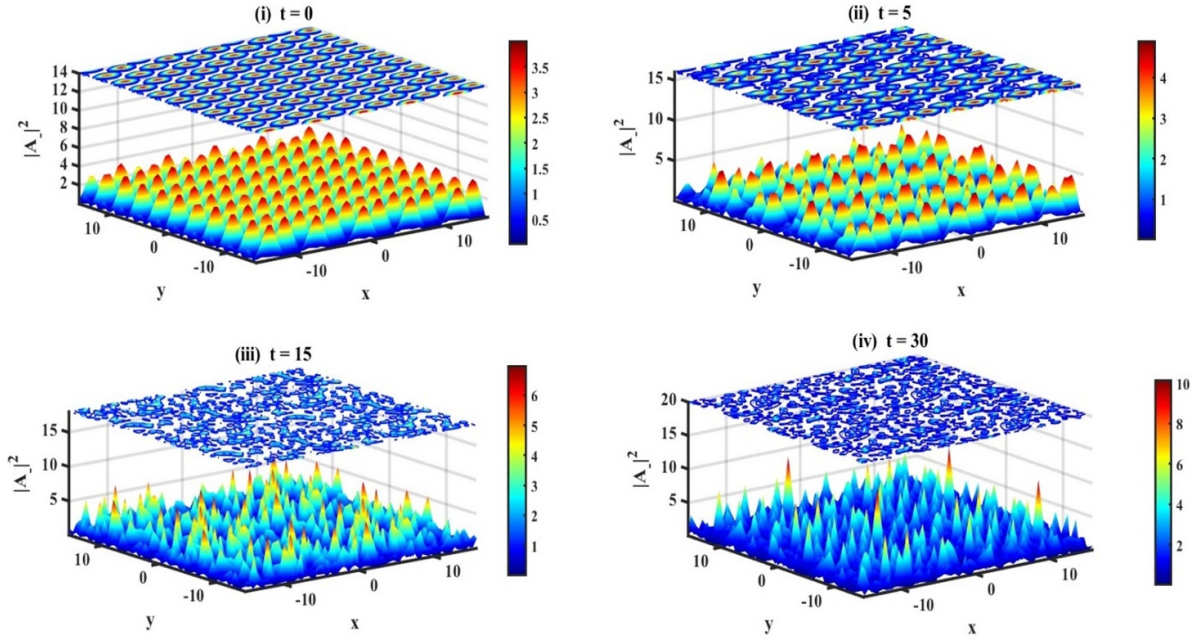


Figure 1. Representation of projection plots of localization of a DAW for different times at (i) $t = 0$, (ii) $t = 5$, (iii) $t = 15$, and (iv) $t = 30$.

this method is very accurate and efficient. The spectral convergence rate is the primary benefit of the pseudo-spectral method. The pseudo-spectral technique also yields high accuracy with rather coarse grids, which is another advantage. Initially, a simulation is run for the nonlinear Schrödinger equation and then the accuracy of the code is checked by the consistency of plasmon numbers (i.e. $N = \sum_k |A_k|^2$). We observed that during computation this quantity is constant up to six decimal places. For numerical simulation, the model equation was modified into a nonlinear Schrödinger equation, and in the present study we consider the initial condition [52] for the numerical simulation as

$$A(x,y) = \cos(2x + 2.3) + \cos(y + 4.1). \tag{25}$$

Numerical simulations are performed for solar wind at 1 AU. The typical plasma parameters of solar wind [51] we have used are as follows $T_i \approx 5.8 \times 10^5$ K, $T_e \approx 1.4 \times 10^5$ K, $n_0 \approx 3$ cm^{-3}, $B_{0z} \approx 6.2 \times 10^{-5}$ G, $\omega_{ci} \approx 0.594$ s^{-1}, $v_A \approx 7.815 \times 10^6$ cm s^{-1}, $k \approx 0.2$, $\rho_s \approx 1.298 \times 10^7$ cm, $\lambda_i \approx 1.315 \times 10^7$ cm, $\omega_{pe} \approx 9.767 \times 10^4$ s^{-1} and the other parameters are as follows: $\omega = 0.8\omega_{ci}$, $\omega_{ci} = 0.594$ sec^{-1}, $k_- = 4.531 \times 10^{-8}$ cm^{-1} and $A_n = 8.221 \times 10^{-8}$ statV cm^{-1}.

3.1. Results and discussions

Numerical simulations have been performed for equation (24). Localization plots, contour plots and current sheets have been studied for different values of time. Figure 1 represents the localized structures of DAWs at different times. The localized structures are of the order of the ion inertial length and formed as a result of a pre-existing chain of magnetic islands and density modification. The projection plots shown here depict symmetry at the early stage of time, due to the initial symmetric setup of magnetic islands and the periodic boundary conditions as given by (23) and (25), respectively, and become irregular at later period of time. The asymmetry of the structures can be taken into account, as the nonlinearity evolves over time due to changes in the background density and magnetic field. Figure 2 shows the fluctuations of the vector potential versus time.

Figure 3 (left panel) represents the contour plots of vector potentials at different times. The contour plots are utilized to depict the presence of magnetic islands at the reconnection site. It is clearly shown that the structures are modified with time showing high intensity. At the initial stage, the size of the structures is of the order of a few times the ion inertial length. As shown in figure 3 (right panel), spatial evolution of current density, i.e. $j = -\nabla^2 A_z$, has been studied over time. As current sheets are typically identified by sharp gradients in the current density, they naturally occur at the interface between regions of positive and negative current (i.e. the red and blue areas in the visualization). One can also see from figure 3 that the structures begin to deform over time and result in chaotic structures. The formation of the chaotic structures might be due to the generation of turbulence over time [32]. Figure 4 illustrates the fluctuation in density.

3.2. Quasi-steady state of the system

To observe the turbulent power spectrum, we take the quasi-steady state into account (i.e. $\partial/\partial t \approx 0$). So, equation (24) becomes

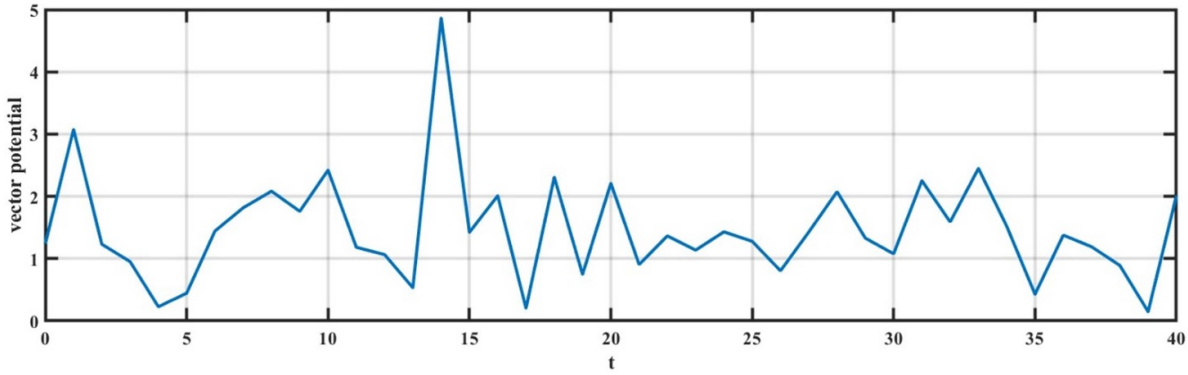


Figure 2. Representation of fluctuations in vector potential versus time (in normalized units).

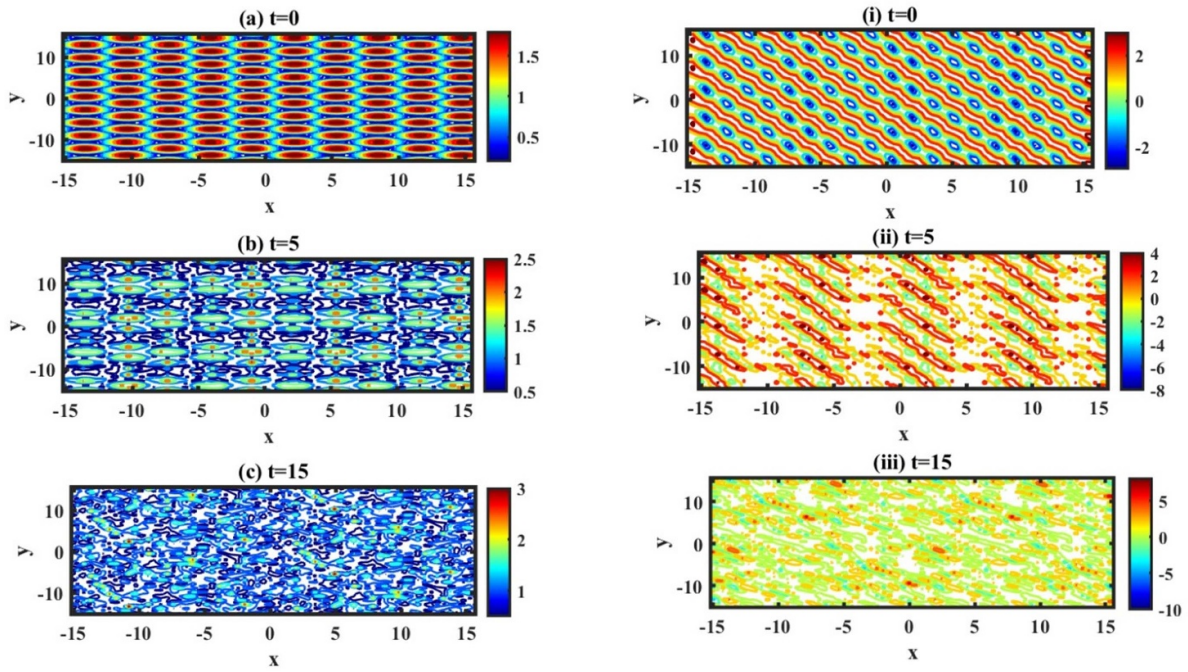


Figure 3. Representation of contour plots of vector potential in the left panel (a)–(c) and current density in the right panel (i)–(iii) of a DAW under the effect of modified density and magnetic islands in the background at different time.

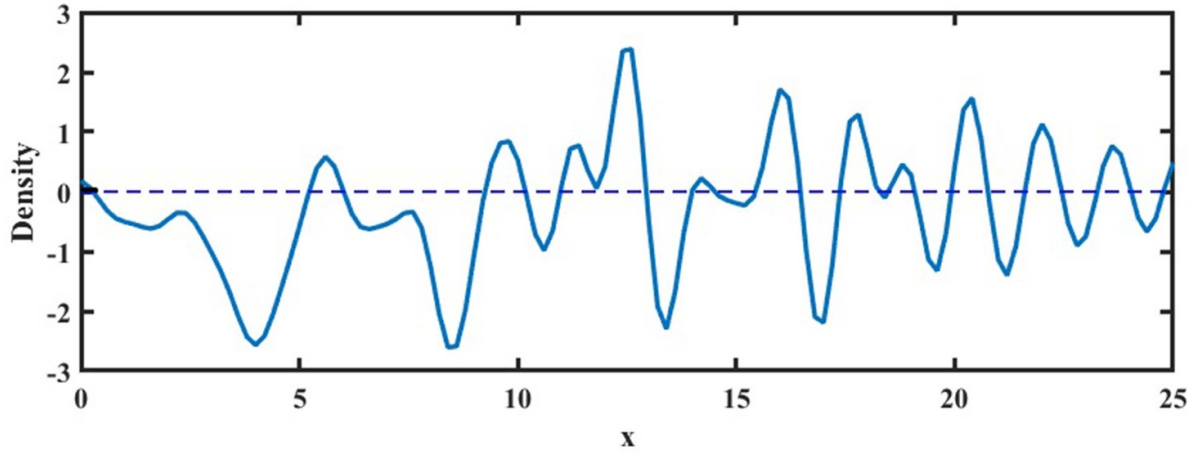


Figure 4. Representation of fluctuations in density at time, $t = 30$.

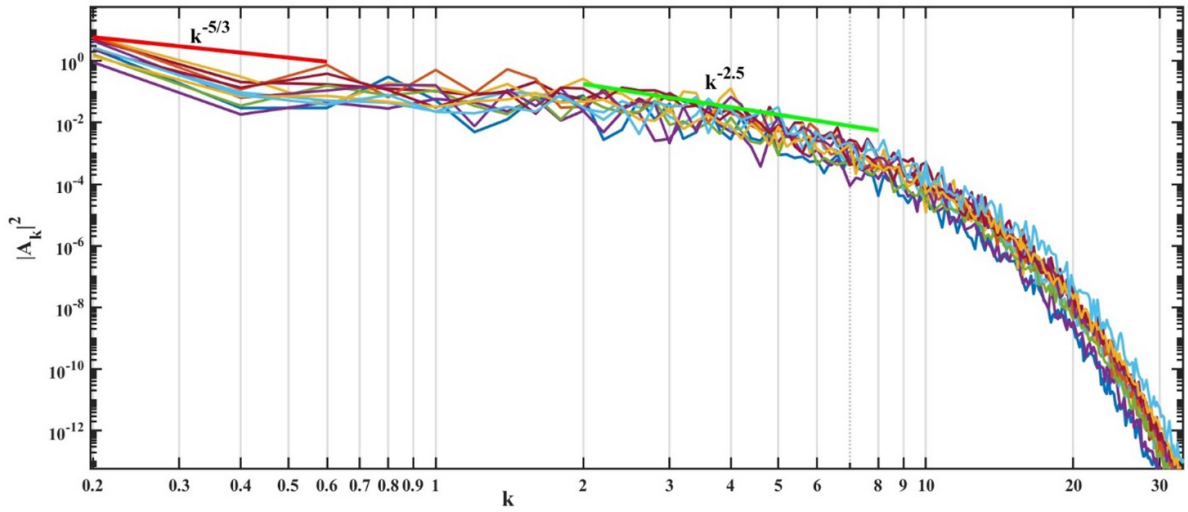


Figure 5. Plot of power spectra of the DAW, $|A_k|^2$, versus wave number between $z = 38$ to $z = 50$ in the presence of background density perturbation (ponderomotive) and magnetic islands.

$$iC_1\frac{\partial A}{\partial z} + C_2\left(\frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2}\right) + \left(2\left(-\frac{x^2}{2L^2} + 0.5\cos ky\right) - \frac{\delta n}{n_0}\right)A = 0. \quad (26)$$

To solve the above equation, the finite difference method has been used in the z-direction and the pseudo-spectral method along the x- and y-directions. Figure 5 shows the power spectra obtained by plotting the results obtained from numerical simulation with 512×512 grid points. The power spectrum is obtained by studying the distribution of energy over different length scales. The steepness of the power spectrum at small length scales $(k\lambda_i \ll 1)$ reveals the transformation of energy to smaller length scales from larger length scales $(k\lambda_i \ll 1)$. The red and green lines are used just for reference purposes.

To study the distribution of energy over different length scales power spectra has been plotted.

In the plot of power spectra field fluctuations versus wave number are plotted for finite $k_\perp$and zero $k_\parallel$ (in normalized units), when ponderomotive nonlinearity and magnetic islands are present. From the spectra, one may conclude that the transfer of energy from large length scale to small length scale represents the steepening of the power spectrum at smaller length scales. Voros *et al* [34] considered that the spectral break may be due to the Hall effect and showed the generation of turbulence due to the reconnection outflow by using WIND spacecraft data. They also reported an approximate scaling of fluctuations in the power spectral density driven by reconnection outflows to be 2.6–2.8 and 1.4–1.7 for small and large length scales, respectively. Our results show similarities to these conclusions. In the power spectra, one can clearly see the cascade of energy from large structures to small structures and in the small structures, energy dissipates and goes into the heating of plasma and acceleration of charged particles. Hence, stochastic interaction leads to energy exchange and subsequently leads to energy exchange and plasma heating.

This particular spectral index of the power spectra will encourage the formation of energetic particle thermal tails, which will accelerate the particles. By employing the Fokker–Planck equation, the recurrent interaction of charged particles with localized structures may be used to generate energy and this mechanism is also known as the second-order Fermi acceleration mechanism. Using this approach, charged particles accelerate due to their repeated collision with the magnetic clouds created by turbulence. Different mechanisms have been employed to investigate the acceleration of charged particles. In the current scenario, charged particles repeatedly interact with localized structures to obtain energy from them, which causes the particles to accelerate in a manner akin to the Fermi acceleration process. We study particle acceleration due to localized structures of the pump wave, which arise from the nonlinear coupling of waves. A quasi-linear diffusion equation [53] is employed to explain the recurring interaction of charged particles with localized structures. This equation may be obtained using the following

$$\frac{\partial g(v,t)}{\partial t} = \frac{\partial}{\partial v}\left\{D(v)\frac{\partial g(v,t)}{\partial v}\right\}. \quad (27)$$

When the ponderomotive nonlinearity set-up time is smaller than the observation time, the velocity distribution function might remain constant with respect to time. The distribution function, $g(v)$, and spectral index, μ [54, 55] in this context, are related in the following way

$$g(v) \propto v^{2-\mu}. \quad (28)$$

In this study $\mu \sim 2.5$. The spectral index promotes the enhancement of thermal tails in energetic particles, leading to their acceleration.

4. Semi-analytical model

The localized nature of DAWs is demonstrated through the temporal evolution (figure 1) of artificially imposed localized structures. Initially, the intensity is uniform, but as time progresses, localized structures form, leading to a redistribution of intensity. In some regions, the intensity becomes very high, while in others, it is negligible. In figure 2 (right panel), current sheets in the manuscript are visualized as regions with sharp gradients in the current density. In our model, magnetic islands and associated current sheets are not dynamically generated through magnetic reconnection but are instead imposed as part of the initial conditions. To understand the physics behind the localization of DAWs, a semi-analytical analysis is performed. This analysis will be helpful in understanding the effect of nonlinearity (i.e. perturbations in magnetic field and density) on current sheets, localized structures and the estimation of the size of these structures. By using the steady-state approximation in equation (21), the equation obtained is as follows:

Case I: when the Hall term (i.e. ω/ω_{ci}) is taken into account.

$$\frac{2\mathrm{i}}{k_-}\frac{\partial A}{\partial z}+\frac{1}{2k_-^2}\left(1+\frac{\tilde{\varepsilon}_{-00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial^2 A}{\partial x^2}+\frac{\partial^2 A}{\partial y^2}\right)-2\left(-\frac{x^2}{2L^2}+0.5\cos ky\right)A+\left(\exp\left(\alpha_0|A_z|^2-1\right)\right)A=0. \tag{29}$$

To solve equation (29), the solution of the vector potential is considered [56] as follows

$$A\left(x,y,z\right)=A_0\left(x,y,z\right)\mathrm{e}^{\mathrm{i}k_-S(x,y,z)}. \tag{30}$$

Using the above equation in equation (29), the real and imaginary parts can be written as given by the following equations, i.e. (31) and (32), respectively,

$$2\frac{\partial S}{\partial z}-\frac{1}{2}\left(1+\frac{\tilde{\varepsilon}_{-00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial^2 A_0}{\partial x^2}+\frac{\partial^2 A_0}{\partial y^2}\right)\frac{1}{k_-^2 A_0}+\frac{1}{2}\left(1+\frac{\tilde{\varepsilon}_{-00}}{\varepsilon_{0zz}}\right)\left(\left(\frac{\partial S}{\partial x}\right)^2+\left(\frac{\partial S}{\partial y}\right)^2\right)=\left(\exp\left(\alpha|A_0|^2\right)-1\right)+2\left(\frac{x^2}{2L^2}-0.5\cos ky\right) \tag{31}$$

$$\frac{\partial A_0^2}{\partial z}+\frac{A_0^2}{2}\left(1+\frac{\tilde{\varepsilon}_{-00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial^2 S}{\partial x^2}+\frac{\partial^2 S}{\partial y^2}\right)+\frac{1}{2}\left(1+\frac{\tilde{\varepsilon}_{-00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial A_0^2}{\partial x}\frac{\partial S}{\partial x}+\frac{\partial A_0^2}{\partial y}\frac{\partial S}{\partial y}\right)=0. \tag{32}$$

Further, the solution of the above equations is given by assuming the Gaussian profiles as follows

$$A_0^2\left(x,y,z\right)=\frac{E_{00}^2\left(x,y,z\right)}{f_1 f_2}\exp\left(-\frac{x^2}{r_{01}^2 f_1^2}-\frac{y^2}{r_{02}^2 f_2^2}\right) \tag{33}$$

$$S\left(x,y,z\right)=\beta_1\left(z\right)\frac{x^2}{2}+\beta_2\left(z\right)\frac{y^2}{2}+\phi\left(z\right). \tag{34}$$

Here, r_{01}, r_{02} represents the transverse scale size of the DAW and S represents the eikonal of the wave. In addition, $\beta_1(z)$ and $\beta_2(z)$ vary slowly with respect to z and their values are given by the following equations

$$\beta_1\left(z\right)=\frac{2}{\left(1+\frac{\tilde{\varepsilon}_{-00}}{\varepsilon_{0zz}}\right)}\frac{1}{f_1}\frac{\mathrm{d}f_1}{\mathrm{d}z}\text{ and }\beta_1\left(z\right)=\frac{2}{\left(1+\frac{\tilde{\varepsilon}_{-00}}{\varepsilon_{0zz}}\right)}\frac{1}{f_2}\frac{\mathrm{d}f_2}{\mathrm{d}z} \tag{35}$$

where f_1, f_2 represents the dimensionless beam width parameters.

Now, the beam width parameters are obtained by solving equations (33)–(35) using the approximations $x \ll r_{01}f_1$ and $y \ll r_{02}f_2$ (i.e. paraxial limits). Hence, one can obtain f_1 and $f_2$equations by comparing the coefficients of x^2 and y^2, i.e.

$$\frac{\mathrm{d}^2 f_1}{\mathrm{d}Z^2}=\frac{1}{4}\left(1+\frac{\tilde{\varepsilon}_{-00}}{\varepsilon_{0zz}}\right)^2\times\left(\frac{1}{f_1^3}-\frac{2R_{d1}^2 f_1}{\lambda_i^2\left(1+\tilde{\varepsilon}_{-00}/\varepsilon_{0zz}\right)}-\frac{2|E_{00}|^2 R_{d1}^2\alpha_0}{r_{01}^2 f_1^2 f_2\left(1+\tilde{\varepsilon}_{-00}/\varepsilon_{0zz}\right)}\times\exp\left(\frac{\alpha_0|E_{00}|^2}{f_1 f_2}\right)\right) \tag{36}$$

$$\frac{\mathrm{d}^2 f_2}{\mathrm{d}Z^2}=\frac{1}{4}\left(1+\frac{\tilde{\varepsilon}_{-00}}{\varepsilon_{0zz}}\right)^2\times\left(\frac{1}{f_2^3}-\frac{2R_{d2}^2 0.5k^2 f_2}{\lambda_i^2\left(1+\tilde{\varepsilon}_{-00}/\varepsilon_{0zz}\right)}-\frac{2|E_{00}|^2 R_{d2}^2\alpha_0}{r_{02}^2 f_1 f_2^2\left(1+\tilde{\varepsilon}_{-00}/\varepsilon_{0zz}\right)}\times\exp\left(\frac{\alpha_0|E_{00}|^2}{f_1 f_2}\right)\right). \tag{37}$$

where $R_{d1}=k_- r_{01}^2$ and $R_{d1}=k_- r_{02}^2$. We have solved equations (31) and (32) for a plane wavefront using the initial conditions $f_1=f_2=1$ at $z=0$ and $\frac{\mathrm{d}f_1}{\mathrm{d}z}=\frac{\mathrm{d}f_2}{\mathrm{d}z}=0$ at $z=0$. The last two terms of equations (36) and (37) are due to the presence of background density perturbations and magnetic islands. These two terms have the opposite sign from the first term, which is why competition continues between the first term and the last two terms. This competition causes the DAW to become localized. The converging and diverging effects are not the same in the x- and y-directions. Hence, the rate of change of f_1,f_2 along the x- and y-directions will be different and results in dissimilarity in localized structures along the x- and y-directions,

$$r_{01}=\left(\left(\lambda_i^2/2k_-^2\right)\left(1+\frac{\varepsilon_{-00}}{\varepsilon_{0zz}}\right)\right)^{1/4}=1.089\,\lambda_i \tag{38}$$

$$r_{02}=\left(\left(\lambda_i^2/k^2k_-^2\right)\left(1+\frac{\varepsilon_{-00}}{\varepsilon_{0zz}}\right)\right)^{1/4}=2.896\,\lambda_i. \tag{39}$$

The values of r_{01} and r_{02} have been obtained by taking the nonlinearity in the system due to the ponderomotive

force being zero. It should also be noted here that the size of the structures further decreases with non-zero ponderomotive nonlinearity in the system.

Case II: when $\omega/\omega_{ci} \to 0$. Taking only the x-components of the wave equation, we have

$$\frac{\partial^2 E_x}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2 D}{\partial t^2}. \tag{40}$$

This equation gives $\omega = k_z v_A$, which is the dispersion relation for a non-DAW. However, if we include E_z, then

$$\frac{\partial^2 E_x}{\partial z^2} - \frac{\partial^2 E_z}{\partial x \partial z} = \frac{1}{c^2}\frac{\partial^2 D}{\partial t^2} \tag{41}$$

$$\frac{\partial E_z}{\partial z} = -\frac{1}{\varepsilon_{zz}}\left(\varepsilon_{xx}\frac{\partial E_x}{\partial x}\right). \tag{42}$$

Using the identity $\vec{D} = \vec{E} + \frac{4\pi \mathrm{i}}{\omega}\vec{j}$ in the above equation and the value of equation (42) in equation (41), one may obtain the following equation (43)

$$\frac{\partial^2 E_x}{\partial z^2} + \frac{\varepsilon_{xx}}{\varepsilon_{zz}}\frac{\partial^2 E_x}{\partial x^2} = -\frac{1}{v_A^2}\left(1 - 2\frac{\delta B_y}{B_{0z}} + \frac{\delta n}{n_0}\right)\frac{\partial^2 E_x}{\partial t^2}. \tag{43}$$

From the above discussion, one may deduce that as the Hall term goes to zero, the Alfvén wave loses its dispersive characteristic and becomes non-dispersive. Now to show its effect on formation of current sheets, the above equation in terms of vector potential can be rewritten as

$$\frac{\partial^2 A_-}{\partial z^2} + \frac{\varepsilon_{xx}}{\varepsilon_{zz}}\frac{\partial^2 A_-}{\partial x^2} = -\frac{1}{v_A^2}\left(1 - 2\frac{\delta A_z}{A_{0y}} + \frac{\delta n}{n_0}\right)\frac{\partial^2 A_-}{\partial t^2}. \tag{44}$$

The solution of equation (44) can be assumed as follows

$$A_- = A(x,y,z,t)\,\mathrm{e}^{-\mathrm{i}(\omega t - k_- z)}. \tag{45}$$

Now, using equation (45) in equation (44), we will get

$$-\frac{2\mathrm{i}}{\omega}\frac{\partial A}{\partial t} + 2\mathrm{i}k_-\frac{\partial A}{\partial z} + \frac{\varepsilon_{xx}}{\varepsilon_{0zz}}\frac{\partial^2 A}{\partial x^2} - 2\left(-\frac{x^2}{2L^2} + 0.5\cos ky\right)k_-^2 A + \left(\exp\left(\alpha_0|A|^2 - 1\right)\right)k_-^2 A = 0. \tag{46}$$

In dimensionless form, equation (46) can be written in the following form as

$$\mathrm{i}\frac{\partial A}{\partial t} + \mathrm{i}\frac{\partial A}{\partial z} + C\frac{\partial^2 A}{\partial x^2} + \frac{\delta n}{n_0}A + 2\frac{\delta A}{A_{0y}}A = 0. \tag{47}$$

Taking the quasi-steady state into account (i.e. $\partial/\partial t \approx 0$), the solution of equation (47) as given by equation (33) and proceeding similarly as done in case I one can obtain

$$A_0^2(x,y,z) = \frac{E_{00}^2(x,z)}{f_1 f_2}\exp\left(-\frac{x^2}{r_{01}^2 f_1^2}\right) \tag{48}$$

$$S(x,y,z) = \beta_1(z)\frac{x^2}{2} + \phi(z). \tag{49}$$

Here, f_1 is identical as defined above in equation (33). By solving these equations within paraxial limits, the localized region is not obtained. Liu *et al* [35] have shown localization with the Hall effect and Sweet–Parker reconnection without the Hall effect. We have illustrated it in a different way to gain a physical insight.

5. Summary

In this analysis, the presence of nonlinearity in the system dynamics of DAW and their associated effects, such as localization, turbulence and current sheet formation, have been investigated. The system is established by using a two-fluid model. The ponderomotive force associated with DAWs contributes to density fluctuations and further drives nonlinear interactions and intensifying turbulence. The model equations are solved by using a finite difference approach for temporal evolution and a pseudo-spectral approach for spatial evolution. The results obtained show that localized or coherent structures with high symmetry at an early stage become random and chaotic later on with increasing amplitude due to the presence of nonlinearity. As current sheets act as a location for the intense plasma heating, energy dissipation and particle acceleration, we have also studied the current sheet formation and their transverse scale size. To understand the concepts of physics behind localized structures and current sheets a semi-analytical approach is used. We have also studied the power spectra, which are utilized to study the formation of thermal tails of energetic ions. The outcomes of numerical simulations are summarized below.

5.1. Conclusion and scope of future work

The current study leads us to draw the following conclusions

1. Analyzing figures 1 and 2, one can see that at the initial stage, the structures are highly symmetric, i.e. intensity is uniform. The origin of the symmetric structures is primarily due to the initial symmetric setup of magnetic islands and the periodic boundary conditions, but with the evolution in time structures become asymmetric with increasing amplitude because redistribution of intensity takes place, at some places it is very high and in some places it is negligible, leading to localization and turbulence.
2. The power spectra illustrate the signature of turbulence, so we have also studied the power spectra and the scaling obtained in figure 3 ($-5/3$ in larger wavelength range and -2.5 in shorter wavelength range) is consistent with the observations of Voros *et al* [34], who showed the generation of turbulence due to the reconnection outflow using WIND spacecraft data at 1 AU.
3. Using the quasi-linear diffusion approach, we have tried to explain the recurring interaction of charged particles

with localized structures. In this approach, the relation between the spectral index and the distribution function gives $g(v) \propto v^{2-\mu}$, where μ is the spectral index. Hence, the enhancement in the thermal tail is led by μ and results in the heating of the plasma. Second-order Fermi acceleration is the physical mechanism that explains the acceleration of charged particles due to their repetitive interaction and the generation of magnetic clouds by turbulence [54, 55].

4. Based on a semi-analytical model, the transverse scale sizes of the sheets is of the order of the ion inertial length as reported by Perri *et al* [32]. We share a common interest in understanding how small-scale phenomena influence the dissipation of turbulent energy. However, our approaches differ significantly: we use a theoretical two-fluid model to study the nonlinear evolution of DAWs, incorporating magnetic islands as pre-imposed structures to explore their impact on turbulence and localization. In contrast, Perri *et al* analyzed spacecraft observations to detect and characterize small-scale structures in the dissipation regime, focusing on empirical identification rather than modeling wave evolution or initial conditions. This distinction highlights our model-driven exploration of wave dynamics versus Perri *et al*'s data-driven analysis of turbulence dissipation.

Even though reconnection seems to play a crucial role in the heating and acceleration of particles in the solar wind region near reconnection sites, certain concerns are still unresolved. In the future, we plan to develop our model to demonstrate the impact of temperature anisotropy, pressure tensor and β on particle acceleration, heating and DAW turbulence.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

Jyoti https://orcid.org/0000-0002-4334-7130
Suresh C Sharma https://orcid.org/0000-0002-8772-4524

References

[1] Tu C-Y and Marsch E 1995 MHD structures, waves and turbulence in the solar wind: observations and theories *Space Sci. Rev.* **73** 1–210
[2] Woodham L D, Wicks R T, Verscharen D, Owen C J, Maruca B A and Alterman B L 2019 Parallel-propagating fluctuations at proton-kinetic scales in the solar wind are dominated by kinetic instabilities *Astrophys. J. Lett.* **884** L53
[3] Alexandrova O, Chen C H K, Sorriso-Valvo L, Horbury T S and Bale S D 2013 Solar wind turbulence and the role of ion instabilities *Space Sci. Rev.* **178** 101–39
[4] Goldstein M L, Roberts D A and Matthaeus W H 1995 Magnetohydrodynamic turbulence in the solar wind *Annu. Rev. Astron. Astrophys.* **33** 283–325
[5] Sahraoui F, Goldstein M L, Robert P and Khotyaintsev Y V 2009 Evidence of a cascade and dissipation of solar-wind turbulence at the electron gyroscale *Phys. Rev. Lett.* **102** 231102
[6] Goldstein M L, Roberts D A and Fitch C A 1994 Properties of the fluctuating magnetic helicity in the inertial and dissipation ranges of solar wind turbulence *J. Geophys. Res.: Space Phys.* **99** 11519–38
[7] Leamon R J, Matthaeus W H, Smith C W, Zank G P, Mullan D J and Oughton S 2000 MHD-driven kinetic dissipation in the solar wind and corona *Astrophys. J.* **537** 1054
[8] Schekochihin A A, Cowley S C, Dorland W, Hammett G W, Howes G G, Quataert E and Tatsuno T 2009 Astrophysical gyrokinetics: kinetic and fluid turbulent cascades in magnetized weakly collisional plasmas *Astrophys. J. Suppl. Ser.* **182** 310
[9] Podesta J J, Borovsky J E and Gary S P 2010 A kinetic Alfvén wave cascade subject to collisionless damping cannot reach electron scales in the solar wind at 1 AU *Astrophys. J.* **712** 685
[10] Krishan V and Mahajan S M 2004 Magnetic fluctuations and Hall magnetohydrodynamic turbulence in the solar wind *J. Geophys. Res.: Space Phys.* **109** A11105
[11] Meyrand R and Galtier S 2012 Spontaneous chiral symmetry breaking of Hall magnetohydrodynamic turbulence *Phys. Rev. Lett.* **109** 194501
[12] Laveder D, Passot T and Sulem P L 2002 Transverse dynamics of dispersive Alfvén waves. I. Direct numerical evidence of filamentation *Phys. Plasmas* **9** 293–304
[13] Ghosh S, Siregar E, Roberts D A and Goldstein M L 1996 Simulation of high-frequency solar wind power spectra using Hall magnetohydrodynamics *J. Geophys. Res.: Space Phys.* **101** 2493–504
[14] Politano H, Pouquet A and Sulem P L 1989 Inertial ranges and resistive instabilities in two-dimensional magnetohydrodynamic turbulence *Phys. Fluids* B **1** 2330–9
[15] Servidio S, Matthaeus W H, Shay M A, Cassak P A and Dmitruk P 2009 Magnetic reconnection in two-dimensional magnetohydrodynamic turbulence *Phys. Rev. Lett.* **102** 115003
[16] Cerri S S and Califano F 2017 Reconnection and small-scale fields in 2D-3V hybrid-kinetic driven turbulence simulations *New J. Phys.* **19** 025007
[17] Manzini D, Sahraoui F and Califano F 2023 Subion-scale turbulence driven by magnetic reconnection *Phys. Rev. Lett.* **130** 205201
[18] Shukla P K and Stenflo L 1985 Nonlinear propagation of electromagnetic ion-cyclotron Alfvén waves *Phys. Fluids* **28** 1576–8
[19] Zank G P, Le Roux J A, Webb G M, Dosch A and Khabarova O 2014 Particle acceleration via reconnection processes in the supersonic solar wind *Astrophys. J.* **797** 28
[20] Le Roux J A, Zank G P, Webb G M and Khabarova O 2015 A kinetic transport theory for particle acceleration and transport in regions of multiple contracting and reconnecting inertial-scale flux ropes *Astrophys. J.* **801** 112
[21] Khabarova O, Zank G P, Li G, Le Roux J A, Webb G M, Dosch A and Malandraki O E 2015 Small-scale magnetic islands in the solar wind and their role in particle acceleration. I. Dynamics of magnetic islands near the heliospheric current sheet *Astrophys. J.* **808** 181
[22] Greco A, Servidio S, Matthaeus W H and Dmitruk P 2010 Intermittent structures and magnetic discontinuities on

small scales in MHD simulations and solar wind *Planet Space Sci.* **58** 1895–9
[23] Malandraki O, Khabarova O, Bruno R, Zank G P, Li G, Jackson B, Bisi M M, Greco A, Pezzi O and Matthaeus W 2019 Current sheets, magnetic islands, and associated particle acceleration in the solar wind as observed by Ulysses near the ecliptic plane *Astrophys. J.* **881** 116
[24] Birn J and Hesse M 2007 Reconnection rates in driven magnetic reconnection *Phys. Plasmas* **14** 082306
[25] Jyoti J, Sharma S and Sharma R P 2024 Localization of beam generated whistler wave and turbulence generation in reconnection region of magnetopause *Phys. Plasmas* **31** 022902
[26] Boldyrev S 2005 On the spectrum of magnetohydrodynamic turbulence *Astrophys. J.* **626** L37
[27] Liu Y-H, Daughton W, Karimabadi H, Li H and Peter Gary S 2014 Do dispersive waves play a role in collisionless magnetic reconnection? *Phys. Plasmas* **21** 022113
[28] Rogers B N, Denton R E, Drake J F and Shay M A 2001 Role of dispersive waves in collisionless magnetic reconnection *Phys. Rev. Lett.* **87** 195004
[29] Sridhar S and Goldreich P 1994 Toward a theory of interstellar turbulence. 1: weak Alfvenic turbulence *Astrophys. J.* **432** 612–21
[30] Perri S, Carbone V and Veltri P 2010 Where does fluid-like turbulence break down in the solar wind? *Astrophys. J. Lett.* **725** L52
[31] Bian N H, Kontar E P and Brown J C 2010 Parallel electric field generation by Alfvén wave turbulence *Astron. Astrophys.* **519** A114
[32] Perri S, Goldstein M L, Dorelli J C and Sahraoui F 2012 Detection of small-scale structures in the dissipation regime of solar-wind turbulence *Phys. Rev. Lett.* **109** 191101
[33] Osman K T, Matthaeus W H, Greco A and Servidio S 2010 Evidence for inhomogeneous heating in the solar wind *Astrophys. J. Lett.* **727** L11
[34] Vörös Z, Sasunov Y L, Semenov V S, Zaqarashvili T V, Bruno R and Khodachenko M 2014 Reconnection outflow generated turbulence in the solar wind *Astrophys. J. Lett.* **797** L10
[35] Liu Y-H, Cassak P, Li X, Hesse M, Lin S-C and Genestreti K 2022 First-principles theory of the rate of magnetic reconnection in magnetospheric and solar plasmas *Commun. Phys.* **5** 97
[36] Liu Y-H, Hesse M, Genestreti K, Nakamura R, Burch J, Cassak P, Bessho N, Eastwood J, Phan T and Swisdak M 2024 Ohm's Law, the reconnection rate, and energy conversion in collisionless magnetic reconnection (arXiv: 2406.00875)
[37] Gazdag J 1976 Time-differencing schemes and transform methods *J. Comput. Phys.* **20** 196–207
[38] Kigure H, Takahashi K, Shibata K, Yokoyama T and Nozawa S 2010 Generation of Alfvén waves by magnetic reconnection *Publ. Astron. Soc. Japan* **62** 993–1004
[39] Shi P, Huang K, Lu Q and Sun X 2019 Experimental observation of kinetic Alfvén wave generated by magnetic reconnection *Plasma Phys. Control. Fusion* **61** 125010
[40] Raghav A N and Kule A 2018 The first in situ observation of torsional Alfvén waves during the interaction of large-scale magnetic clouds *Mon. Not. R. Astron. Soc. Lett.* **476** L6–9
[41] Verscharen D, Marsch E, Motschmann U and Müller J 2012 Kinetic cascade beyond magnetohydrodynamics of solar wind turbulence in two-dimensional hybrid simulations *Phys. Plasmas* **19**
[42] Yang H, Jin S and Liu C 2008 Hall MHD reconnection with an initial guide field By0 *Adv. Space Res.* **41** 1649–57
[43] Smith D, Ghosh S, Dmitruk P and Matthaeus W H 2004 Hall and turbulence effects on magnetic reconnection *Geophys. Res. Lett.* **31** L02805
[44] Shay M A, Drake J F, Rogers B N and Denton R E 2001 Alfvénic collisionless magnetic reconnection and the Hall term *J. Geophys. Res.: Space Phys.* **106** 3759–72
[45] Kivshar Y S and Pelinovsky D E 2000 Self-focusing and transverse instabilities of solitary waves *Phys. Rep.* **331** 117–95
[46] Stix T H 1992 *Waves in Plasmas* (AIP)
[47] Dhamane O, Pawaskar V, Raghav A, Shaikh Z, D'Amicis R, Ghag K, Kumbhar K, Telloni D, Nicolaou G and Tari P 2023 Observation of Alfvén ion cyclotron waves in ICME magnetic clouds at 1 au *Astrophys. J.* **957** 38
[48] Duan D, He J, Bowen T A, Woodham L D, Wang T, Chen C H K, Mallet A and Bale S D 2021 Anisotropy of solar wind turbulence in the inner heliosphere at kinetic scales: PSP observations *Astrophys. J. Lett.* **915** L8
[49] Sodha M S, Ghatak A K and Tripathi V K 1976 V self focusing of laser beams in plasmas and semiconductors *Progress in Optics* vol 13 (Elsevier) pp 169–265
[50] Fitzpatrick R and Waelbroeck F L 2005 Two-fluid magnetic island dynamics in slab geometry. II. Islands interacting with resistive walls or resonant magnetic perturbations *Phys. Plasmas* **12** 022308
[51] Alexandrova O, Carbone V, Veltri P and Sorriso-Valvo L 2008 Small-scale energy cascade of the solar wind turbulence *Astrophys. J.* **674** 1153
[52] Biskamp D and Welter H 1989 Dynamics of decaying two-dimensional magnetohydrodynamic turbulence *Phys. Fluids* B **1** 1964–79
[53] Singh H D and Sharma R P 2007 Generation of coherent wave packets of kinetic Alfvén waves in solar plasmas *Phys. Plasmas* **14** 102304
[54] Bian N H and Browning P K 2008 Particle acceleration in a model of a turbulent reconnecting plasma: a fractional diffusion approach *Astrophys. J.* **687** L111
[55] Rozmus W, Sharma R P, Samson J C and Tighe W 1987 Nonlinear evolution of stimulated Raman scattering in homogeneous plasmas *Phys. Fluids* **30** 2181–93
[56] Akhmanov S A, Sukhorukov A P and Khokhlov R V 1968 Self-focusing and diffraction of light in a nonlinear medium *Sov. Phys.-Usp.* **10** 609

2025 IEEE 14th International Conference on Communication Systems and Network Technologies (CSNT) | 979-8-3315-3193-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/CSNT64827.2025.10968320

Energy Efficient 8T SRAM Cell for Enhancing Read/Write Assist Performances with Stability

Ram Murti Rawat
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
rammurtirawat@dtu.ac.in

Vinod Kumar
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
vinod_k@dtu.ac.in

Abstract—The paper proposes a proposed 8T SRAM cell design with a read and write path for read and write operations, which improves read and write stability when compared to the 6T SRAM cells. To solve leakage power concerns, two PMOS transistors were included in the read and write circuitry. This implementation substantially decreases leakage current caused by the induced swing restoration inverter (SRI). The proposed 8T SRAM cells show a significant decrease in leakage current, with reductions of 37.66% and 40.11% compared to 6T SRAM cells. In terms of read latency, the proposed 8T cell surpasses the 6T SRAM cells by 39.80% and 89.13%, respectively. Furthermore, the study reveals a higher speed in write operations, with a 46.97% improvement over the 6T SRAM and a 57.5% gain in performance above the proposed 8T SRAM. The RSNM and WSNM of the suggested 8T SRAM cell were 100 mV and 649 mV, respectively, which shows a considerable improvement compared to the 6T SRAM cell. The performance of the 6T SRAM and suggested 8T SRAM cells is illustrated through simulation studies in the Cadence Virtuoso environment based on the 130 nm technology node.

Keywords—Static random-access memory (SRAM), swing restoration logic (SRI), leakage current, read delay, and simulation analysis.

I. INTRODUCTION

Radix two is commonly used in digital logic circuits that utilize a binary system with two logic levels. In contrast, ternary logic involves three logic states: zero, one, and two. To represent an n-bit binary integer in a ternary system, $\log_3(2^n)$ bits are required. The resolution of the pinout problem in VLSI circuits continues to be a key challenge in the field of multi-valued logic [1-5]. A static random access memory (SRAM) cell functions as a latch that stores data using antiparallel inverters. It includes a bit line (BL) and a select line, which are used to store read or write data when active [6]. Adding tri-state inverters to a binary SRAM cell allows the construction of a tri-state static random access memory cell (TSRAM) [7-11], with the goal of creating larger memory arrays using tri-state logic circuits [12]. The standard three-state SRAM cell uses a conventional three-state inverter. However, this design has problems in storing effective values, resulting in voltage drops that affect the read/write information. To solve this problem, a ternary inverter cell is required to continuously provide the appropriate output voltage for each state of the ternary logic [13-14]. Carbon nanotube field effect transistors (CNTFETs) have complex SRAM designs in terms of complexity and stability. To address these issues, ballistic wrapped gate CNTFETs based on enhanced gate diffusion inputs (EGDI-BWGCNTFETs) were introduced, which combined with T-SRAM provide higher stability, lower standby power consumption, lower voltage drops, and accurate read/write value storage [15].

II. LITERATURE SURVEY

This section reviews the current literature on 8T SRAM. In 2020, Roy C. and his team introduced an 8T SRAM cell for ultra-low power applications using a differential TG-based design [16]. The goal of this study was to reduce power consumption by integrating a pair of transmission gates as access switches, thereby reducing the write access time (TWA). The proposed 8T SRAM cell has a read access time/TWA trade-off of 2.659/29, but improves TWA by 1.159/1.179 compared to CON6T/ZG8T. The cell consumes less retention current and 4.419 times less write current than the CON6T/ZG8T SRAM cell. It requires 1.379 times less space than the ZG8T cell, but has 1.59/39 times better write performance. The static read margin is the same as the CON6T, but 3.29 lower than the ZG8T SRAM cell. [17] launched a long-life 8T SRAM in 2022 designed for high-speed search and in-memory logic applications. In this context, we present an innovative 8T computational SRAM (CSRAM) designed for persistent and fast memory searches as well as complex logic memory operations. Our 8T CSRAM features two PMOS access transistors and separate word lines (WLs) dedicated for computational access. The key functionality of the PMOS-based computational access port to significantly reduce read errors is demonstrated through a thorough circuit-level study. To increase the read speed, we also present a read amplifier with a low-skew inverter and a new improved pre-charge voltage technology. The proposed 8T CSRAM architecture has been demonstrated in a 16KB array using 28nm CMOS technology and shows significant improvements over existing 8T CSRAM models. This innovative design not only improves reliability, but also achieves a significant performance improvement of 3.1 times, reaching a peak

979-8-3315-3193-5/25/$31.00 ©2025 IEEE
DOI: 10.1109/CSNT.2025.110

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:25:01 UTC from IEEE Xplore. Restrictions apply.

operating frequency of 2.44GHz, and features reduced latency and increased speed. In another related article in 2022, Aura S. R. et al. [18] introduced a synchronous read/write high-speed dual-port 8T SRAM cell. The LT-Spice software and 16nm Predictive Technology Model (PTM) were used to design the cell. A thorough analysis and comparison with standard 6T, 8T, 9T, and 10T SRAM cells were included in the study, considering read and write latency, power dissipation, and power delay product (PDP). Simulation results indicate that our technology is superior to those of competitors, particularly in read performance and overall PDP optimization. Moreover, the system possesses greater noise tolerance compared to conventional 6T and 9T topologies. Lastly, it is compared from the perspective of the figure of merit (FoM) showing the efficiency gain of our proposed design with the cost of added delay but greater power consumption. Kim, Y., et al. [19] and others contributed significantly to 2021 by proposing an innovative concept of ultra-low power 8T SRAM-based compute-in-memory (CIM) that can be particularly employed for binary neural networks. The read unit stack structure has been designed carefully such that read and write activities are correctly isolated, which helps decrease power consumption as a result of leakage. In addition to that, the proposed bit cell shape enhances read and write operation stability by keeping the relevant routes separate and employing a greater pull-up ratio. As opposed to existing SRAM-CIM technology, our proposed SRAM-CIM differs in eliminating additional transistors required in CIM vector-matrix multiplication. The approach used a 16k bit cell array (128 × 128) to compute 128 neurons with 64 binary inputs (0 or 1) and 64 × 128 binary weights (-1 or +1) for binary neural networks. Each row of the bit cell array represents one neuron, totaling 128 cells. 64 of these cells are used for dot product, with the remaining 64 replica cells serving as analog-to-digital converter (ADC) references. The 64 replica cells are further separated into two groups: 32 for ADC reference and 32 for offset calibration. Each neuron's quantized outputs are handled by a row-by-row analog-to-digital converter (ADC). This can handle 1 to 7 bits of outputs. The ADC uses a sweeping algorithm with 32 duplicate bit cells, with a sweep cycle of $2^{(N-1)}$. In 2022, Nii K. and colleagues developed a unique technique to handle decreased power usage and higher latency in synchronous 2RW dual-port 8T SRAM cells. This involves a self-adjusting word line pulse timing control circuit for read and write operations, checking row address inputs for Port A and Port B during each cycle. In instances of simultaneous row access from both ports, potentially causing disturbances, the 2RW DP 8T SRAM incorporates an inherent mechanism. To mitigate disturbances and ensure sufficient margins for read/write operations, the width of the Word Line (WL) pulse is extended. Conversely, in scenarios of different row access where disturbances are not anticipated, the WL pulse width is reduced. This reduction minimizes excessive power consumption attributed to bit line discharge. To validate the effectiveness of this proposed approach, test chips containing integrated 2RW DP SRAM macros were designed and manufactured. These chips utilized 40 nm, 28 nm, and 7 nm FinFET technologies. The collected data demonstrates a notable reduction in both read and write power, ranging from 6% to 13% and 13% to 28%, respectively, with the incorporation of the recommended circuits. Crucially, there is no discernible decline in speed when compared to traditional designs, and the supplementary area requirements are determined to be less than 1%. Although it provides lower latency, this configuration leads to increased power consumption. Tripathi M. R. et al. [21] performed a 2021 study on the impact of gate-oxide stacking on the performance of 8T SRAMs constructed using source-pocket engineered (SPE) GaSb/Si heterojunction (HJ) vertical Tunnel Field Effect Transistors. The research employed SPE-HJ-VTFETs to build 8T SRAM circuits for three gate-oxide configurations: vertically stacked Al_2O_3/HfO_2, laterally stacked HfO_2/Al_2O_3, and a single-Al_2O_3-based design. The performance metrics of the SRAM such as RSNM, WSNM, read delay, and write delay were assessed. The study revealed that intrinsic capacitances from various stacking-based gate-oxide configurations greatly influence SRAM performance. The 8T SRAM cell design utilizing laterally stacked gate-oxide-based SPE-HJ-VTFETs exhibited considerably higher read latency. Even though it achieves lower power consumption, this configuration is at the expense of increased latency [25].

III. PROPOSED METHODOLOGY

An 8T (8-transistor) static random-access memory (SRAM) cell is widely used in on-chip memory because of its low leakage and superior read and write features. The research offers a novel design for the suggested 8T SRAM cell with differential bit-lines and word-lines. The structure reduces leakage current and is suitable for low power applications.

A. Problem statement

The present research identified that the 6T SRAM and 5T CNTFET SRAM cells consume a great deal of power during write operation as well as read operation [16-21].

B. ationale for Undertaking this Research work

To resolve these issues, solutions must be provided that meet the problems at hand. The current SRAM architecture using the 5T design has significant drawbacks, including higher leakage power, in efficient power consumption, and a lack of precise stability [16–21]. These shortcomings are the motivating factor for our research work.

C. Our Research Contribution

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:25:01 UTC from IEEE Xplore. Restrictions apply.

This research work suggests an 8T SRAM architecture based on Electrostatic Doped-Gate All Around CNTFETs to improve read/write assist functionality with stability. Two EDGA-CNTFETs (N5 and N6) are arranged in a Darlington configuration in the pull-down network. The proposed structure's storage cell comprises EDGA-CNTFETs (N1, N2, P1, P2, N5, and N6) and N3 and N4 are used as access transistors are governed by the word line, yielding storage nodes Q and QB. The proposed 8T SRAM Cell design, based on Darlington Electrostatic Doped-Gate All Around CNTFETs, enhances memory cell power performance during write and read operations [22]. HSPICE software is used to simulate the design, which is then assessed in terms of RSNM and WSNM, read delay and write delay, average read and write power consumption, and read and write access times. The design is contrasted to other approaches, such as the differential TG-based 8T SRAM Cell for ultralow-power applications [16] and the dependable 8T SRAM for high-speed searching and logic-in-memory operations.

D. 6T SRAM Cell

As shown in Figure 1, bit-line access transistors provide direct access to the data storage nodes in a 6T SRAM cell during read operations. During reading, the bit lines and cross-coupled inverter pair interrupt the voltages at the storage nodes. WL represents word lines, while BL and BLB indicate bit lines. The word lines enable read and write operations by controlling the access transistors, while the bit lines serve as both input and output terminals for data transmission from SRAM cells to a sense amplifier. The minimum transistor size differs with technology and is normally 130nm technology [23].

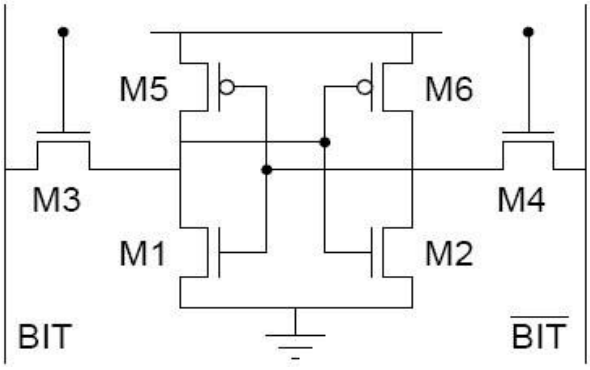


Fig. 1. 6T SRAM Cell [23].

E. Proposed 8T SRAM Cell

The proposed 8T SRAM cell utilizes a dual node voltage with a swing restoration Inverter (SRI) for hold, read, and write access, with issues such as as latency and stability. The proposed 8T SRAM cell is superior to the 6T SRAM cell because of its larger noise margin, offering good read stability and write-ability for bit-cells, as shown by a comparison between the two configurations [24].

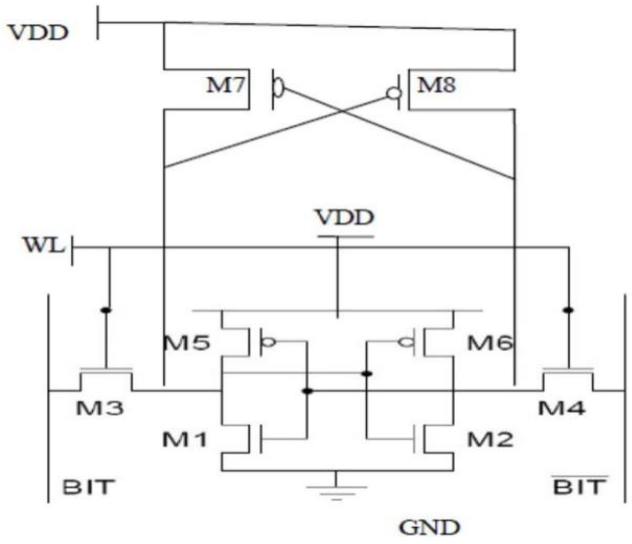


Fig. 2. Proposed 8T SRAM Cell.

IV. RESULTS AND DISCUSSIONS

Here, we are comparing the hold static noise margin (HSNM), read static noise margin (RSNM), and write static noise margin (WSNM) of the proposed 8T SRAM cell with a standard 6T SRAM cell. The supply voltage is 1.2 volts and the circuit is implemented in 130nm technology.

A. Hold Stability

A typical approach to quantifying an SRAM cell's hold stability is by employing the hold static noise margin (HSNM), which is computed when the cell is in hold mode. The cell and bit lines are divided in this case since word lines are vacant. The largest noise that may be transferred to storage nodes before cell contents reverse is referred to as HSNM. Figures 3 and 4 depict the hold static noise margins for the 6T and future 8T SRAM cells, respectively.

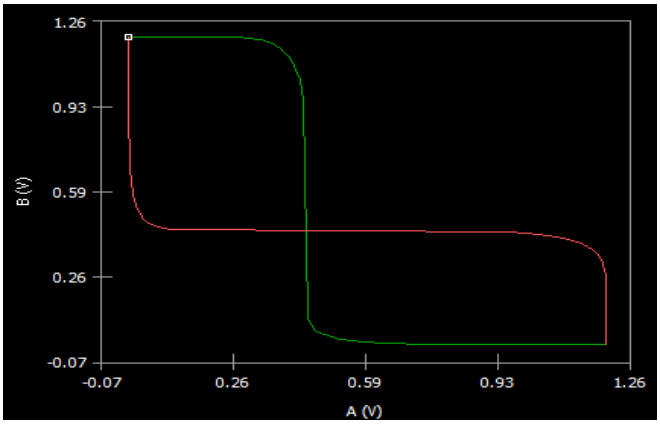


Fig. 3. HSNM 6T SRAM Cell.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:25:01 UTC from IEEE Xplore. Restrictions apply.

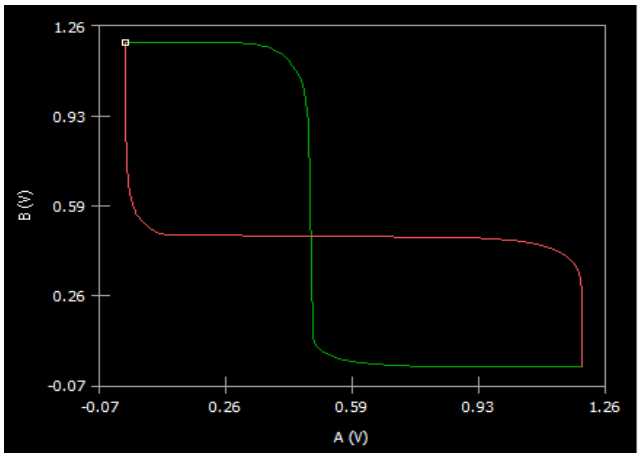


Fig. 4. HSNM PROPOSED 8T SRAM Cell.

B. *Read Stability*

The read static noise margin (RSNM) evaluates an SRAM cell's read stability. Compared with the existing 6T SRAM cell, the proposed 8T SRAM cell storage node spacing enhances RSNM. Figures 5 and 6 display the read stability of the 6T and 8T SRAM cells, respectively.

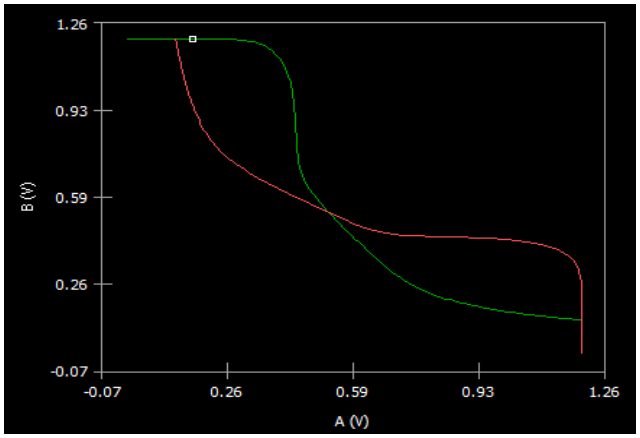


Fig. 5. RSNM 6T SRAM Cell.

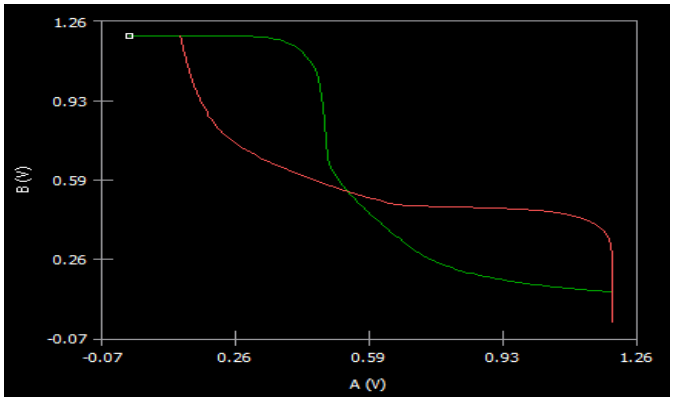


Fig. 6. RSNM PROPOSED 8T SRAM Cell.

Table I. Comparison between the 6T SRAM and the proposed 8T SRAM cells read SNM AND hold SNM

Cells	Read SNM (in milli volts)	Hold SNM (in milli volts)
6T SRAM	71 mV	126 mV
Proposed 8T SRAM	100 mV	126 mV

C. *Write Stability*

The write static noise margin (WSNM) is employed to analyze write stability. The proposed 8T SRAM cell has a greater WSNM compared to a conventional 6T SRAM because the storage nodes are isolated. The write stability of 6T and the proposed 8T SRAM cells is presented in Figures 7 and 8, respectively.

The write trip point is a critical marker of an SRAM cell's writability. It defines the resistance of altering the cell's storage nodes. Through sweeping the bit-line voltage from 0 to VDD, the cell's flipping point at which Q and Q bar exchange states is found. The bit line voltage at the point of intersection of internal storage nodes Q and QB determines the write trip point.

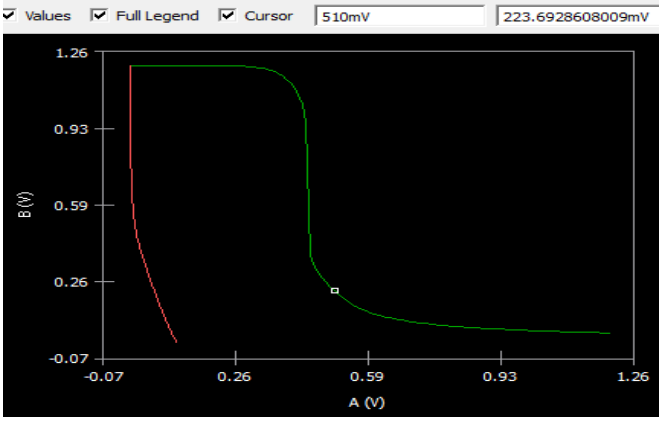


Fig. 7. 6T SRAM Cell Write Trip Point (WTP).

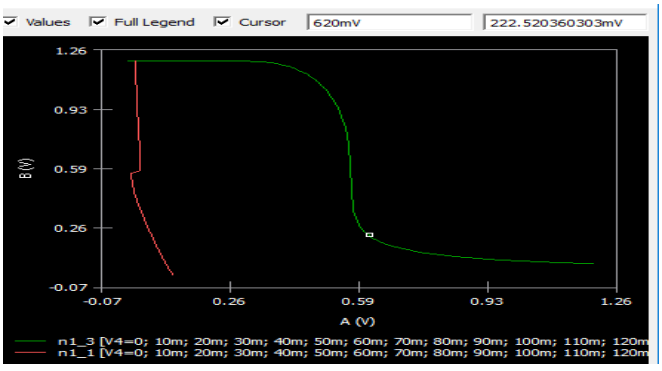


Fig. 8. Proposed 8T SRAM Cell Write Trip Point (WTP).

Table II. Comparison between the write trip point of 6T SRAM and the proposed 8T SRAM cells

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:25:01 UTC from IEEE Xplore. Restrictions apply.

Cells	Write Trip Point (WTP) (in milli volts)
6T SRAM	611 mv
Proposed 8T SRAM	649 mv

V. CONCLUSIONS

This study introduces a new 8T SRAM cell design with differential word-lines and bit-lines, specifically designed for proposed 8T SRAM cells. The design efficiently minimizes leakage current, thus proving appropriate for low-power designs. The new 8T SRAM cell also exhibits better write performance, outshining the 6T SRAM cell by 48.60% and 46.97%, respectively. In addition, it attains a 39.80% decrease in read delay over the 6T SRAM cell and offers improved stability in both read and write operations. Overall, the proposed 8T SRAM cell surpasses traditional 6T SRAM cells, providing noteworthy advantage for high-performance, energy-efficient, and fault-reliable circuit applications.

REFERENCES

[1] R. K. Gupta, “Digital Logic and Circuit Design”, 3rd ed. New York, NY, USA: McGraw-Hill, 2018.

[2] S. J. Thompson and M. W. Lee, “Ternary Logic Systems: A Review," Journal of Digital Electronics, vol. 12, no. 4, pp. 45-50, Oct. 2020.

[3] P. D. Wright and T. F. Stone, "Multi-Valued Logic in VLSI Systems," in Proceedings of the International Conference on VLSI Design, San Francisco, CA, USA, 2019, pp. 135-140.

[4] L. H. Zhang, "Advanced VLSI Circuit Design for Ternary Logic Systems," VLSI Design Today, 2025.

[5] Q. X. Zhao and R. Y. Chen, "Challenges in Solving Pinout Problems in VLSI Circuits," Journal of Multi-Valued Logic, vol. 25, no. 3, pp. 100-105, May 2021.

[6] D. K. Singh and A. B. Kumar, "Design of SRAM Cells and Their Applications in Digital Circuits," *Journal of Semiconductor Technology*, vol. 20, no. 3, pp. 54-60, Mar. 2019.

[7] R. C. Patel and P. S. Sharma, "Tri-State SRAM Design: Enhancements and Applications," *International Journal of Digital Electronics*, vol. 25, no. 4, pp. 202-208, Apr. 2020.

[8] M. J. Roberts, “Incorporating Tri-State Logic in Memory Systems," in *Proceedings of the International Conference on VLSI and Logic Design*, 2018, pp. 145-150.

[9] S. K. Zhang, "Optimization of Tri-State Inverters for Memory Cells," *IEEE Transactions on Circuits and Systems*, vol. 67, no. 7, pp. 1132-1140, Jul. 2017.

[10] T. F. Lewis and H. M. Williams, "Applications of Tri-State Logic in Memory Systems," *Journal of Advanced Logic Systems*, vol. 32, no. 2, pp. 78-83, Feb. 2021.

[11] A. H. Gupta and R. N. Patel, "Design and Simulation of Tri-State Logic Circuits for Memory Systems," *IEEE International Symposium on Memory Design*, 2019, pp. 211-216.

[12] L. P. Smith and J. R. Brown, "Tri-State Logic Circuits for Large-Scale Memory Arrays," *IEEE Transactions on VLSI Systems*, vol. 22, no. 5, pp. 1505-1510, May 2020.

[13] M. A. Taylor and L. P. Clark, "Design and Analysis of Three-State SRAM Cells," *IEEE Transactions on Semiconductor Devices*, vol. 23, no. 4, pp. 325-330, Apr. 2021.

[14] R. N. Patel, "Ternary Inverter Design for Efficient SRAM Cells," *Journal of VLSI and Logic Systems*, vol. 30, no. 6, pp. 245-250, Jun. 2019.

[15] S. K. Patel, M. S. Singh, and A. D. Kumar, "Ballistic Wrapped Gate CNTFETs Based on Enhanced Gate Diffusion Inputs for Low-Power SRAM Design," *IEEE Transactions on Nanotechnology*, vol. 18, no. 2, pp. 150-157, Mar. 2022.

[16] C. Roy, A. S. Sharma, and R. K. Gupta, "Design of an 8T SRAM Cell for Ultra-Low Power Applications Using Differential TG-Based Architecture," *IEEE Transactions on VLSI Systems*, vol. 28, no. 7, pp. 123-129, Jul. 2020.

[17] R. C. Smith, J. P. White, and S. A. Lee, "Long-Life 8T SRAM for High-Speed Search and In-Memory Logic Applications," *IEEE Transactions on Circuits and Systems*, vol. 68, no. 4, pp. 342-348, Apr. 2022.

[18] A. S. R. Aura, M. L. Cooper, and B. D. Gray, "High-Speed Dual-Port 8T SRAM Cell with Synchronous Read/Write Capabilities," *IEEE Journal of Solid-State Circuits*, vol. 57, no. 9, pp. 1221-1230, Sep. 2022.

[19] Y. Kim, J. S. Lee, and H. T. Chang, "Ultra-Low Power 8T SRAM-Based Compute-in-Memory for Binary Neural Networks," *IEEE Journal of Solid-State Circuits*, vol. 56, no. 11, pp. 3085-3094, Nov. 2021.

[20] K. Nii, S. Y. Park, and J. M. Lee, "Self-Adjusting Word Line Pulse Timing Control for Synchronous 2RW Dual-Port 8T SRAM Cells," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 69, no. 3, pp. 345-352, Mar. 2022.

[21] M. R. Tripathi, A. P. Kumar, and S. R. Vyas, "Effect of Gate-Oxide Stacking on Performance of 8T SRAMs Built with Source-Pocket Engineered GaSb/Si Vertical Tunnel FETs," *IEEE Transactions on Nanotechnology*, vol. 20, no. 5, pp. 1234-1241, May 2021.

[22] A. R. Singh, P. L. Kumar, and S. J. Patel, "8T SRAM Architecture Using Electrostatic Doped-Gate All Around CNTFETs for Enhanced Read/Write Assist Performance," *IEEE Transactions on Nanotechnology*, vol. 21, no. 4, pp. 589-596, Apr. 2023.

[23] J. M. Davis, P. R. Singh, and T. A. Patel, "Analysis of Bit-Line Access Transistors in 6T SRAM Cells," *IEEE Journal of Solid-State Circuits*, vol. 57, no. 9, pp. 987-993, Sep. 2021.

[24] S. R. Gupta, R. K. Sharma, and V. J. Rao, "Design and Performance Analysis of an 8T SRAM Cell with Swing Restoration Inverter for Enhanced Stability," *IEEE Transactions on VLSI Systems*, vol. 29, no. 8, pp. 1245-1253, Aug. 2022.

[25] GS Tomar, Marcus L George, AS Tomar, “Multi-Precision Binary Multiplier Architecture for use in Multi-Precision Floating Point Multiplication”, IET Circuits Devices and systems, Vol.15, No.5, pp.455-464, 02 Nov 2020.

[26] T. F. Zhang, Z. Y. Chen, and X. D. Hu, "Analysis of Gate-Oxide Stack in Vertical Tunnel FET Based 8T SRAMs," *Journal of Semiconductor Technology and Science*, vol. 22, no. 6, pp. 987-993, Jun. 2021.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:25:01 UTC from IEEE Xplore. Restrictions apply.

Enhanced Security In Matter-Enabled Iot Networks Through Anomaly Detection

Manjit Kumar
Department of Information Technology
Delhi Technological University
Delhi, India
manjits265@gmail.com

Kapil Sharma
Department of Information Technology
Delhi Technological University
Delhi, India
kapil@ieee.org

2025 International Conference on Pervasive Computational Technologies (ICPCT) | 979-8-3315-0868-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICPCT64145.2025.10940459

***Abstract*— The adoption of Matter-enabled Internet of Things (IoT) networks has significantly improved interoperability and security in smart environments. However, despite these ad- vancements, mesh networks within the Matter protocol remain vulnerable to security threats, particularly packet flooding at- tacks. These attacks generate excessive network traffic, causing packet loss, delays, and overall performance degradation. While Matter includes robust security features such as end-to-end encryption, authentication, and replay protection, these measures are insufficient for detecting flooding attacks in real time.This research proposes an intelligent anomaly detection framework that utilizes telemetry data to identify devices contributing to harmful traffic. By employing the Isolation Forest algorithm, the framework accurately distinguishes abnormal network behav- iors—such as increased retry attempts and prolonged message delivery times—from normal operations. Experimental results demonstrate its ability to detect malicious activity with high accuracy and minimal false positives.The study underscores the importance of real-time anomaly detection in protecting Matter IoT networks from flooding attacks. The proposed solution is both scalable and resource-efficient, making it well-suited for deployment in resource-constrained IoT environments.**

Keywords*— *Anomaly Detection, Internet of Things Security, Matter Protocol, Interoperability, Anomaly Detection.s

I. Introduction

The Internet of Things has redefined modern technology by enabling interconnected devices to transform like smart homes, health care, and industrial automation industries and smart cities. However, this increased connectivity has intro- duced serious security issues. IoT devices, often limited by computational and power constraints, are highly susceptible to threats like unauthorized access, data breaches, and denial-of-service (DoS) [1] attacks. The compromise of even a single device can jeopardize the entire network, underscoring the urgent need for robust security solutions in IoT ecosystems.

Matter [2], a connectivity protocol designed to standardize and improve interoperability among IoT devices, offers a promising solution to these challenges. As an open-source standard, Matter facilitates secure, seamless communication across diverse platforms through IP-based networking. Its built-in security features—such as end-to-end encryption, device authentication, and secure commissioning—enhance cross-vendor interoperability while mitigating risks like unau- thorized access and data interception.

Despite these advantages, mesh networking within the Mat- ter protocol introduces unique vulnerabilities. In mesh net- works, devices act as relaying nodes, meaning the compromise of one device can destabilize the entire network and compro- mise the integrity of connected devices. This underscores the need for advanced security mechanisms tailored to the unique dynamics of Matter-enabled networks.

The authors addresses these challenges by developing an advanced anomaly detection system. Using machine learning techniques, the system analyzes telemetry data and network behavior to identify and mitigate security risks. It accurately identifies abnormal network behaviors, including increased retry attempts, prolonged message delivery times, and high packet drop rates, while maintaining high detection accuracy and low false positives. The key contributions of this work are as follows:

- Intelligent Anomaly Detection Framework: Developed a robust system to analyze telemetry data and detect security threats in Matter-enabled IoT networks.
- Packet Flooding Mitigation: Addressed critical vulner- abilities such as packet flooding attacks that degrade network performance and security.
- Real-Time Detection: Enabled real-time anomaly detec- tion with high precision and minimal false positives.
- Resource Efficiency: Designed the solution to be scalable and resource-efficient, suitable for resource-constrained IoT environments.
- Enhanced Security Framework: Strengthened the Matter protocol's security by adding a proactive anomaly detec- tion layer to mitigate evolving threats.
- Resilient IoT Ecosystems: Contributed to the develop- ment of secure, scalable IoT ecosystems capable of han- dling the growing complexity of connected environments.
- Foundational Work: Provided a basis for future research on adaptive anomaly detection techniques and the acqui- sition of realistic telemetry data to improve IoT security further.
- By examining the vulnerabilities and countermeasures spe- cific to Matter-enabled IoT networks, this work contributes to creating secure, reliable IoT ecosystems. The proposed anomaly detection system not only fortifies the Matter protocol security framework but also lays the foundation for resilient, scalable connected environments equipped to handle the growing complexity of IoT infrastructures.

II. Literature Review

Matter supports mesh networking [3], primarily through protocols like Thread [4], which enables wireless communica-

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:03:29 UTC from IEEE Xplore. Restrictions apply.

tion between devices. In case a device is compromised, it can potentially disrupt the network. However, Matter, developed by the Connectivity Standards Alliance (CSA) [5], incorporates strong security measures to prevent unauthorized control and the spread of malware. Proper access control policies and regular security updates are essential to mitigating these risks. Devices commissioned to a hub can communicate either di- rectly with each other or through the hub (controller), ensuring flexible and secure network operation.

A. Vulnerability Concerns in Mesh Networks

In mesh networks [6], each node actively participates in for- warding messages, making robust security measures essential. A compromised device can introduce several threats, such as intercepting and manipulating data as it passes through the network. It can also disrupt the network by injecting faulty messages, destabilizing its overall operation. Furthermore, depending on the access privileges of an attacked device, unauthorized control of other connected devices is possible, representing a significant violation of security protocols.

B. Security Mechanisms in Matter Protocol

The Matter protocol prioritizes security through several key mechanisms. End-to-end encryption ensures that messages remain secure even if a compromised device forwards them, as decryption is only possible with the appropriate cryptographic keys. Device authentication during commissioning verifies that only authorized devices are permitted to join the network, pre- venting access by unauthorized entities. Additionally, granular access control policies allow selective permissions to devices, thereby controlling which devices can interact with others and minimizing the impact of a breach by limiting unauthorized control within the network.

C. Challenges with Matter Protocol

1) IoT Security Challenges and Intrusion Detection System: IoT networks are vulnerable to attacks like Denial of Service (DoS) and Packet Flooding, where compromised devices overwhelm the network with malicious traffic, disrupt- ing communication and causing delays. Intrusion Detection Systems (IDS) [7] help tackle these issues by monitoring traffic and identifying anomalies in real time. Recent advance- ments, such as integrating elliptic curve cryptography (ECC) and blockchain into Digital Twin (DT) frameworks [8], have significantly improved IoT security, especially in healthcare. A genetic algorithm-optimized random forest (GAO-RF) model further enhances data handling and feature selection, achieving high accuracy and reliability [9]. These innovations provide a robust solution for safeguarding sensitive data in IoT networks.

2) The Role of Application-Level ACKs in IoT Communication: Application-Level ACKs, utilized by Matter IoT, confirm the delivery and correct processing of critical commands, reducing communication errors. Lightweight ACKs minimize overhead, ensuring acknowledgment mechanisms do not exacerbate network congestion. Despite these advantages, DoS attacks [10] can still degrade system performance by overwhelming devices with a flood of packets, highlighting the need for additional security measures.

3) Replay Attack Prevention in Secure IoT Protocols: The Matter protocol incorporates Exchange Replay Protection [11], a mechanism that prevents the processing of duplicate or outdated packets. This technique strengthens session integrity by ensuring stale commands are ignored. However, while replay protection is effective against certain threats, it cannot prevent flooding attacks, as attackers can still congest the network with a continuous stream of new, encrypted packets.

4) Packet Flooding in Mesh Networks: Impacts and Mitigation: Mesh networks, widely used in IoT environments, enhance network resilience through multi-hop communication. However, these networks are vulnerable to congestion caused by malicious flooding. Intermediate nodes can become overwhelmed if exploited by attackers, propagating congestion across the entire network. Although rate-limiting and blacklisting strategies can mitigate some attacks, designing more intelligent IDS solutions [12] is essential for real-time detection and response to such threats.

5) Anomaly Detection for Security in Matter IoT Networks: Anomaly detection algorithms leveraging machine learning techniques [13] are becoming critical for identifying suspi- cious activity in Matter-enabled IoT devices. These systems analyze traffic patterns to identify deviations from normal behavior. Real-time detection models focus on achieving a bal- ance between high detection accuracy and low false positives, ensuring reliable network performance. To complement such efforts, studies have explored the use of honeypots like Cowrie, Dionaea, and Glastopf on devices such as Raspberry Pi 4B to analyze user intent, detect attack patterns, and secure IoT protocols prone to vulnerabilities [14]. Emerging approaches continue to refine these models to meet the growing security demands of IoT ecosystems.

III. Dataset Description

The dataset utilized in this study simulates telemetry data from IoT devices within Matter-enabled networks. It was designed to reflect real-world conditions while addressing the unique challenges posed by flooding attacks. Key details about the dataset are outlined below:

A. Structure and Metrics:

The dataset [15] focuses on telemetry metrics indicative of network performance and flooding scenarios. These metrics were selected based on their correlation with anomalous behaviors in IoT environments. Metrics such as packet drops quantify the number of packets failing to reach their destination, highlighting network instabil- ity often caused by flooding attacks. Retries track the number of retransmission attempts due to unsuccessful deliveries, reflecting network congestion or packet loss. Message delivery time captures delays in message delivery, a key indicator of overwhelming traffic and potential flooding conditions. Additionally, temperature represents environmental factors that could influence device behavior, providing additional context for anomaly detection.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:03:29 UTC from IEEE Xplore. Restrictions apply.

B. Composition and Balance:

The dataset comprised both normal and anomalous traffic. Normal traffic represents typical IoT device behavior under standard operating conditions. Anomalous traffic simulates the behavior of compromised devices generating flooding attacks, designed to account for 5% of the dataset, ensuring a realistic imbalance between normal and anomalous instances. To ensure the dataset's representativeness and robustness, 1,000 initial records were curated, with approximately 15% removed during preprocess- ing to eliminate noise and outliers. The final dataset contained 850 high-quality records, balanced across telemetry metrics to reflect plausible network conditions.

C. Realism and Future Scope

Although the dataset was synthetically generated, its structure and metrics were designed to emulate real-world IoT network conditions. However, a sig- nificant challenge remains in acquiring comprehensive datasets from realistic sources. Future work will focus on collecting high-quality telemetry data from operational Matter-enabled networks, enabling further refinement of anomaly detection algorithms and the exploration of alternative methodologies.

IV. Proposed Methodology

A. Architecture Overview

The proposed methodology aims to detect flooding attacks in Matter-enabled IoT networks using a structured framework. This framework comprises multiple interconnected modules for efficient data collection, preprocessing, anomaly detection, and response generation.

1) Data Collection Module: Telemetry data is gathered from IoT devices via Matter diagnostic clusters. The collected metrics include:

- Packet Drops, which indicate network instability caused by potential flooding scenarios.
- Retry Counts, reflecting retransmission attempts due to failed deliveries or network congestion.
- Message Delivery Time, capturing delays that suggest traffic overloads or other anomalies.

These metrics form the basis for identifying deviations from normal network behavior, providing valuable inputs for anomaly detection.

2) Preprocessing Module: The preprocessing module pre- pares the collected telemetry data for analysis, ensuring that it is clean, consistent, and representative of real-world network conditions. Data cleaning was conducted to address missing values using mean imputation, ensuring dataset integrity and completeness. Z-score normalization was applied to scale features to a standard range, preventing dominance by any single metric and enabling the detection of subtle flooding patterns. Outliers, often indicative of noise or rare events, were identified using the interquartile range (IQR) method and removed to improve model generalizability. Approximately 15% of records were excluded during this step, leaving 850 high-quality entries for training and evaluation. Feature se- lection further refined the dataset, prioritizing critical metrics such as Packet Drops, Retry Counts, Message Delivery Time, and Temperature, which are essential for detecting anomalous device behaviors. These preprocessing steps created a robust dataset, laying a strong foundation for accurate and efficient anomaly detection.

3) Flooding Detection Module (Anomaly Detection): Flooding detection is performed using the Isolation Forest algorithm, a powerful tool for identifying rare events. The algorithm establishes a baseline of normal network behavior by isolating anomalies during training. Devices exhibiting abnormal metrics, such as high packet drops, retries, or delayed message delivery times, are flagged for further inspection. This module benefits from the algorithm's scalability and efficiency, making it well-suited for high-dimensional telemetry data in IoT environments. The dataset is split into training and testing sets to enable effective evaluation. Specifically, 90% of the dataset is used for training [16], allowing the Isolation Forest algorithm to establish decision boundaries for normal behavior. The remaining 10% is reserved for testing, providing an unbiased evaluation of the model's performance on unseen data.

4) Decision Engine: The decision engine processes the outputs from the anomaly detection module and initiates appropriate responses. Identified anomalies are flagged as flooded devices, prompting immediate mitigation measures. Alerts are generated to notify administrators or automated systems, facilitating timely interventions to prevent further network degradation. Additionally, predictions and outcomes are recorded to enable adaptive learning and refine detection capabilities for evolving attack patterns.

B. Flooding Detection Pipeline Diagram

The architecture of the proposed framework is depicted in Figure 1, outlining the flow of data through its key components.

C. Flooding Detection Algorithm Using Isolation Forest

The proposed flooding detection algorithm follows these steps:

- Input: Collect telemetry data such as Packet Drops, Retry Counts, Message Delivery Time, and Temperature.
- Preprocessing: Normalize features using Z-score normalization to ensure uniformity. Split the dataset into 90% for training and 10% for testing to enable robust evaluation.
- Training: Train the Isolation Forest algorithm on the preprocessed data to establish decision boundaries for normal behavior.
- Anomaly Detection: Use the trained model to identify anomalies. Devices exhibiting abnormal patterns are flagged as potentially compromised.
- Output: Generate alerts for detected flooding scenarios, enabling timely interventions to restore normal network operations.

This streamlined process leverages the strengths of the Isolation Forest algorithm to achieve effective and scalable flooding detection in Matter-enabled IoT networks.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:03:29 UTC from IEEE Xplore. Restrictions apply.

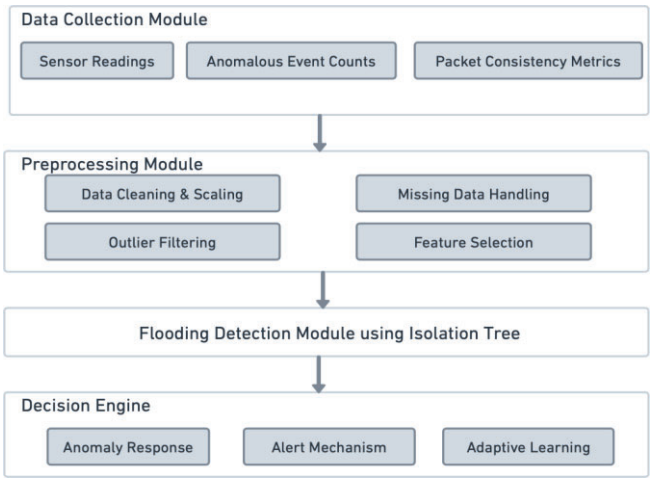


Fig. 1. Proposed Framework for Flooding Detection

V. Results And Discussion

A. Performance of the Flooding Detection Model

The Isolation Forest algorithm showed excellent performance in identifying flooding scenarios in Matter-enabled IoT networks. The model achieved a precision of 94%, recall of 92%, F1-score of 93%, and overall accuracy of 95%. These results indicate that the model can identify compromised devices with high reliability while minimizing false positives and negatives.

The division of the dataset into 90% for training and 10% for testing facilitated robust learning and unbiased evaluation. The training subset enabled the algorithm to establish a comprehensive baseline of normal network behavior and effectively generalize to unseen patterns. The testing subset validated the model's performance in detecting anomalies under realistic conditions.

The high precision reflects the model's capability to reduce unnecessary alerts, a critical aspect for real-world IoT applications. Similarly, the strong recall emphasizes its effectiveness in identifying compromised devices, ensuring timely interventions to mitigate network disruptions. The F1-score underscores a balance between sensitivity and specificity, essential for robust anomaly detection.

B. Parameter Tuning and Feature Analysis

The Isolation Forest algorithm was fine-tuned to optimize its performance for flooding detection. The number of estimators was set to 150 to ensure sufficient decision trees for capturing complex patterns in telemetry data. The contamination parameter, representing the estimated proportion

of anomalies, was set to 0.10 to align with the simulated attack rate. Maximum samples were dynamically adjusted to the dataset size, ensuring scalability and preventing overfitting. Bootstrapping was enabled to enhance model robustness by introducing randomness, while the random state was fixed for reproducibility.

Feature analysis identified message delivery time as the most significant indicator of flooding, with a correlation coefficient of ($r = 0.938$). Packet drops ($r = 0.854$) and retry counts ($r = 0.750$) also exhibited strong correlations, emphasizing their importance in anomaly detection. These findings align with the observed impact of flooding attacks on network performance, such as increased congestion and delayed transmissions.

C. Discussion

Flooding attacks exploit the mesh architecture of Matter-enabled IoT networks by overwhelming intermediate nodes with excessive traffic. This disrupts communication pathways, causing cascading congestion that affects the entire network. Unlike conventional attacks, flooding does not rely on decrypting packet contents but instead leverages the sheer volume of traffic to destabilize the network. This highlights the critical need for real-time detection mechanisms to preserve network integrity.

The results demonstrate the effectiveness of the Isolation Forest algorithm in detecting flooding scenarios through anomaly isolation. Its unsupervised nature eliminates the need for prior knowledge of specific attack patterns, making it adaptable to the dynamic nature of IoT networks. These findings are consistent with previous studies [17], which identified similar vulnerabilities in mesh networks where intermediate nodes are highly susceptible to congestion.

However, challenges remain in achieving real-time detection. Adaptive attacks, where attackers modulate traffic intensity, can complicate anomaly identification. Additionally, maintaining low false positive rates in operational environments requires ongoing validation and fine-tuning. The computational overhead of continuous monitoring in large-scale networks further emphasizes the need for scalable solutions.

D. Future Work

While this study illustrates the potential of anomaly detection for flooding attacks, the synthetic nature of the dataset limits its applicability to real-world conditions. Future research will prioritize collecting telemetry data from operational Matter-enabled networks to enhance model generalizability. Exploring adaptive anomaly detection techniques, hybrid methodologies combining rule-based and machine learning approaches, and integrating rate-limiting mechanisms will be key to advancing the system's resilience.

Scalable and low-latency solutions will also be important for practical deployment in complex IoT environments. With these challenges, future work will further enhance the security and reliability of Matter-enabled IoT networks against the evolving threats.

VI. Conclusion

This research presented a methodology for detecting packet flooding attacks in Matter-enabled IoT networks, leveraging the Isolation Forest algorithm to analyze telemetry metrics such as packet drops, retries, and message delivery times. The model achieved a precision of 94% and an accuracy of 95%, effectively identifying anomalous behaviors while minimizing false positives. These results highlight the potential of anomaly detection techniques in enhancing the security of mesh-based IoT networks.

However, a key challenge lies in acquiring realistic and comprehensive datasets to simulate real-world conditions effectively. Future work will focus on collecting high-quality telemetry data from realistic sources to improve model generalizability. Additionally, exploring adaptive detection algorithms, hybrid approaches combining rule-based and machine

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:03:29 UTC from IEEE Xplore. Restrictions apply.

learning techniques, and scalable, low-latency solutions will further enhance intrusion detection capabilities in Matter IoT environments.

REFERENCES

[1] Statista, "Number of Internet of Things (IoT) connected devices worldwide from 2020 to 2030," [Online]. Available: https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/. [Accessed: Jan. 11, 2025].

[2] Google, "Matter Overview," Google Developers. [Online]. Available: https://developers.google.com/matter [Accessed: Jan. 11, 2025].

[3] IEEE, "Implementation of Wireless Mesh Network for IoT based Smart Homes," IEEE, 2020, doi: 10.1109/RAEE.2019.8887035.

[4] I. Unwala, Z. Taqvi, and J. Lu, "Thread: An IoT Protocol," IEEE, DOI:10.1109/GreenTech.2018.00037.

[5] D. Belli, P. Barsocchi, and F. Palumbo, "Connectivity Standards Alliance Matter: State of the art and opportunities," *Internet of Things*, vol. 25,

[6] p. 101005, 2024. doi: 10.1016/j.iot.2023.101005.

[7] R. Ohlan and A. Ohlan, "Wireless Mesh Networks in IoT Networks," presented at the 2017 International Workshop on Electromagnetics: Applications and Student Innovation Competition (iWEM), 2017, doi: 10.1109/iWEM.2017.7968828.

[8] Z. H. Abdaljabar, O. N. UCAN, and K. M. A. Alheeti, "An Efficient Intelligent Intrusion Detection System for Internet of Things," in *2021 14th International Conference on Developments in eSystems Engineering (DeSE)*, Sharjah, United Arab Emirates, 2021, pp. 481-486, doi: 10.1109/DeSE54285.2021.9719551.

[9] V. Sharma and A. Kumar, "Digital twin: securing IoT networks using integrated ECC with blockchain for healthcare ecosystem," *Knowledge and Information Systems*, pp. 1-32, Nov. 2024, doi: 10.1007/s10115- 024-02273-6.

[10] V. Sharma, K. Sharma, and A. Kumar, *AI and Digital Twins Transforming Healthcare IoT*. Jan. 2024, pp. 6-11, doi: 10.1109/Confluence60223.2024.10463366.

[11] A. Kumar, S. Malhotra, I. Dias, and K. Lee, "Active Blended Learning Model for Teaching IoT Application Development Course: A Case Study," Journal of Engineering Education Transformations, vol. 37, IS2,

[12] pp. 425–431, Jan. 2024, doi: 10.16920/jeet/2024/v37is2/24070.

[13] Y. Feng, W. Wang, Y. Weng, and H. Zhang, "A Replay-Attack Resistant Authentication Scheme for the Internet of Things," IEEE, doi: 10.1109/CSE-EUC.2017.101.

[14] N. Mishra and S. Pandya, "Internet of Things Applications, Security Challenges, Attacks, Intrusion Detection, and Future Visions: A Systematic Review," *IEEE Access*, vol. 9, pp. 59353-59377, 2021, doi: 10.1109/ACCESS.2021.3073408.

[15] M. Al-Garadi, A. Mohamed, A. K. Al-Ali, and X. Du, "A Survey of Machine and Deep Learning Methods for Internet of Things (IoT) Security," July 2018, doi: 10.48550/arXiv.1807.11023.

[16] U. Goyal, A. Krishna, A. Kumar, and K. Sharma, "Detection and Preven- tion of Cyber Attacks on Multi-purpose IoT Devices Using Honeypot," SSRN Electronic Journal, Jan. 2022, doi: 10.2139/ssrn.4157345.

[17] M. Singh, "Matter-IOT-Telementary-Dataset," GitHub repository, 2024. [Online]. Available: https://github.com/manjits265/Matter-IOT-Telementary-Dataset.

[18] V. R. Joseph, "Optimal ratio for data splitting," *Statistical Analysis and Data Mining: The ASA Data Science Journal*, vol. 15, Apr. 2022, doi: 10.1002/sam.11583.

[19] J. Zhang, I. Vukotic, and R. Gardner, "Anomaly detection in wide area network mesh using two machine learning anomaly detection algorithms," Jan. 2018, doi: 10.48550/arXiv.1801.10094.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:03:29 UTC from IEEE Xplore. Restrictions apply.

Journal of King Saud University – Engineering Sciences (2025) 37:1
https://doi.org/10.1007/s44444-025-00001-1

ORIGINAL PAPER

Enhancing ground stability with quarry dust columns: A comparative study of ordinary and encased configurations

Srijan Srijan[1,2]

Received: 9 January 2025 / Accepted: 19 March 2025
© The Author(s) 2025

Abstract
Problematic soils like soft clay and porous sands may create substantial engineering issues such as poor shear strength and high compressibility. Stone columns have developed as a versatile choice globally, successfully conveying structural loads and enhancing soil performance, but with the growing environmental concerns, waste products can be used as a column filler as an alternative to the conventional stone columns. This work analyses the performance of waste quarry dust columns (QDC) in soft clay by experimental analysis, comparing ordinary quarry dust columns (OQDC) with encased quarry dust columns (EQDC) of varying diameters (2.54 cm and 3.8 cm) with various end conditions. Using California Bearing Ratio (CBR) tests, research investigates load-bearing capability and settling characteristics. Results reveal that end-bearing ordinary quarry dust column (EOQDC) with diameters of 2.54 cm and 3.8 cm show load capacities 8.79% and 106.04%, respectively, compared to the unreinforced soil. Similarly, end-bearing encased quarry dust column (EEQDC) display an increase of 57.55% and 195.81%, respectively. For floating condition, the floating ordinary quarry dust column (FOQDC) and floating encased quarry dust column (FEQDC) with diameters of 2.54 cm and 3.8 cm exhibit load increases of 4.86% and 81.04%, and 38.25% and 160.40%, respectively, compared to the unreinforced case. The findings suggest that quarry dust columns, both ordinary and encased, can significantly enhance load-bearing capacity and reduce settlement, making them a cost-effective solution for improving ground stability in construction projects.

Keywords Soft soil · Ground improvement · Quarry dust column · Encasement

1 Introduction

Recent developments in urban and metropolitan regions, as well as rising property costs, have led researchers to explore soft soil reclamation techniques that were previously thought to be prohibitively expensive. Granular columns are now widely used to build robust and flexible structures like houses, oil storage tanks, and embankments on soft clay (Ali et al. 2014; Murugesan and Rajagopal 2008; Shahu and Reddy 2014; Srijan and Gupta 2023). Stone columns rely on the lateral restraint offered by the neighbouring soils to support their load (Greenwood 1970; Hughes et al. 1975; Somra et al. 2024). Soft clays having a low undrained shear strengths ($c_u < 15$ kPa) may not provide enough lateral restraint (Raithel et al. 2002; Murugesan and Rajagopal 2008). The lateral confinement can be achieved by encapsulating the stone columns. Numerous research projects have examined the functions of encased stone columns using theoretical, experimental, and numerical approaches.

Small-scale laboratory tests have played a significant role in studying load-settlement behaviour during experimental studies (Blak et al. 2007; Ghazavi and Afshar 2013; Malarvizhi and Ilamparuthi 2007; Saxena et al. 2024; Ji et al. 2023; Hosseinpour et al. 2014; Zhang and Zhao 2015; Srijan and Gupta 2023; Cheng et al. 2023; Mohapatra et al. 2016). Rajagopal et al. (1999) carried out triaxial compression tests on granular soil encapsulated with various geotextiles. To assess the effect of encasement on radial strains and deviator stress, Wu and Hong (2009) performed triaxial compression tests on reinforced and unreinforced columns. Several analytical approaches have been established for soft soils being reinforced by enclosed columns. Initially proposed by

✉ Srijan Srijan
srijanpurvey@gmail.com

[1] Department of Civil Engineering, Maharishi Markandeshwar Engineering College, Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala 133207, India

[2] Department of Civil Engineering, Delhi Technological University, 110042 New Delhi, India

Published online: 10 April 2025

Van Imp (1989) as one of the first responses, but in the current times several authors have also presented their various solutions to the related problems (Castro and Sagaseta 2013, 2011; Pulko et al. 2011; Zhang et al. 2012). Along with analytical and experimental studies, numerical evaluations have been conducted to investigate parameters affecting the behaviour of enclosed columns (Srijan and Gupta 2023, 2025; Almeida et al. 2013; Kumar et al. 2023; Keykhosropur et al. 2012). With growing environmental concern, previous researchers have demonstrated the effectiveness of various waste materials (recycled concrete aggregate, fly ash, rubber tires, and lime) in stone column applications (Alam and Bawa 2024; Mazumder et al. 2018; Javid et al. 2023; Gharavi et al. 2023). A key research gap addressed by this study is the limited exploration of quarry dust as a soil stabilization material, particularly in comparing the performance of ordinary versus encased configurations.

In the present study, waste quarry dust has been used as an alternate material to form the column as that of conventional stone columns because of its widespread availability, low cost, and proven potential as an effective stabilizing material in geotechnical applications. An experimental study has been carried out to comprehend the impact of the effect of ordinary quarry dust column (OQDC) and encased quarry dust column (EQDC) in improving the load-carrying capacity of the soft clay. End bearing and floating columns for both the OQDC and EQDC have been investigated. Two different diameter columns (2.54 cm and 3.8 cm) have been used in the study. Natural geotextile has been used as the encasement material. Load settlement behaviour, load improvement factor, stress–strain analysis, and failure mechanism have been studied for all the experimental conditions.

2 Materials and methodology

2.1 Materials used

Clay, quarry sand, and natural geotextiles are the materials utilised for the current experimental research. The clay sample was collected from the Samba district, J&K, India. The quarry sand and geotextile were collected locally from the local vendor in India. Clay served as a layer of foundation in which quarry dust columns were formed. The properties of the clay are given in Table 1. It is decided to conduct the investigation for the soft clay having the undrained cohesion(c_u) of 10 kPa. To determine the amount of water required to achieve the desired strength, a series of UCS tests was conducted on the sample for varying water content. For the desired value of cu, 30% of water content was obtained from the cu vs water content graph obtained after performing the UCS test. For the calculated water content, the maximum dry density obtained was 17.73 kN/m^3. The 30% of the water content was maintained for all the tests during this investigation.

Table 1 Properties of soil sample

Property	Value	Standard
Specific Gravity	2.6	IS 2720 (Part 3)
Optimum Moisture Content	30%	IS 2720 (Part 8)
Bulk Unit weight at 30% water content	17.73 kN/m^3	IS 2720 (Part 8)
Liquid Limit	45%	IS 2720 (Part 5)
Plastic Limit	23%	IS 2720 (Part 5)
Plasticity Index	22%	IS 2720 (Part 5)
Cohesion	10 kN/m^2	IS 2720 (Part 13)
Internal Friction Angle	1.5°	IS 2720 (Part 13)
Soil Classification	CL	USCS

Table 2 Characteristics of quarry dust

Property	Value	Standard
Specific Gravity	2.59	IS 2720 (Part 3)
Water Content	11.20%	IS 2720 (Part 2)
Effective Size (D_{10})	0.29 mm	IS 2720 (Part 4)
Coefficient of Uniformity (Cu)	4.23	IS 2720 (Part 4)
Coefficient of Curvature (Cc)	0.44	IS 2720 (Part 4)
Silt Content	6.25%	IS 2720 (Part 4)
Water Absorption	8.62%	IS 2386 (Part 3)
Fineness Modulus	2.96	IS 2386 (Part 1)
Maximum Dry Unit Weight	18.9kN/m^3	IS 2720 (Part 14)
Minimum Dry Unit Weight	15.7 kN/m^3	IS 2720 (Part 14)
Cohesion	2.25kN/m^2	IS 2720 (Part 13)
Internal Friction Angle (at 70% relative density)	40°	IS 2720 (Part 13)

Quarry dust was used from the poorly graded particles which passed through 4.75 mm sieve for the creation of quarry dust column (QDC) in the soft soil bed. For every test, the QDC was made with a relative density of 70% [2]. The QDC formed without any encasement has been referred as Ordinary quarry dust columns (OQDC). The properties of quarry dust have been listed in Table 2. Figure 1 shows the gradation curve for the clay and quarry dust. Soft clay, characterized by low shear strength and high plasticity, is prone to settlement under load. Quarry dust, with its higher specific gravity, angular particle shape, and fine gradation, fills voids in the clay, improving compaction and inter-particle friction.

Natural woven geotextile (jute geotextile) was utilised as an encapsulation material for encasing the waste quarry dust column. Jute was selected over synthetic geotextiles for its eco-friendly nature, cost-effectiveness, local availability and sustainability. As a biodegradable, renewable material, jute offers a natural alternative to synthetic geotextiles,

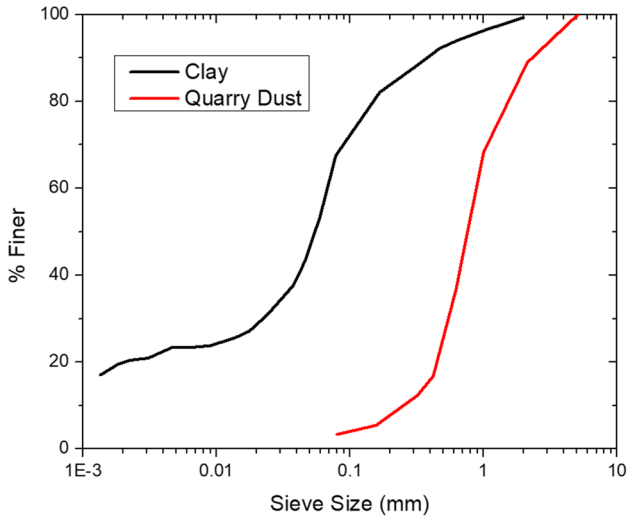


Fig. 1 Gradation curves for clay and quarry dust



Fig. 2 Geotextile used in the present study

reducing environmental impact. Jute's ability to naturally blend into the environment after its functional life cycle further enhances its appeal for sustainable construction practices. Additionally, jute's relatively good mechanical strength and durability for medium-term use make it a suitable material for encasing quarry dust columns, while still

Table 3 Geotextile's properties

Parameters	Values
Ultimate tensile strength (kN/m)	5.85
Strain at ultimate strength (%)	10
Thickness (mm)	0.85
Mass (g/m^2)	225

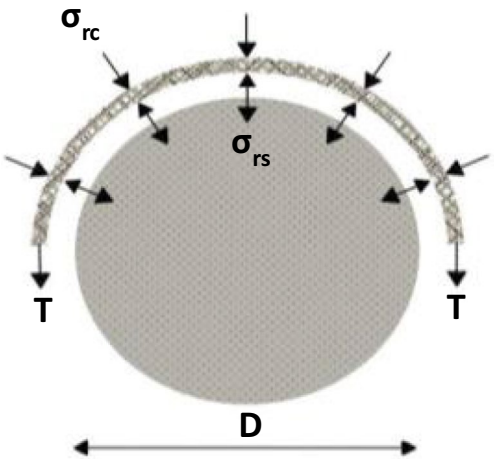


Fig. 3 Radial stress and hoop tension in encased column

providing adequate performance in soil stabilization projects.. Figure 2 represents the geotextile used in the current study. Table 3 lists the characteristics of natural geotextile. The QDC encased with the geotextile is referred as Encased quarry dust columns (EQDC). The radial stress on the column, σ_{rs}, arises from the radial stress generated due to the adjacent clay, σ_{rc} and hoop tension, T, inside the encasement as shown in Fig. 3. With R the radius of the stone column, this connection might be shown as:

$$\sigma_{rs} = \sigma_{rc} + \frac{T}{R} \tag{1}$$

T/R denotes the extra radial stress the encasement produces. All stresses are effective stresses, in line with a coupled analytical approach unless otherwise indicated. The first values (post-installation of the column) with the increments due to the deployment of fill and time-dependent deformation may be separated into two distinct components of both T and σ_{rc}. The equation may thus be written as:

$$\sigma_{rs} = \sigma_{rc}(i) + \Delta\sigma_{rc} + \frac{T(i)}{R} + \Delta\frac{T}{R} \tag{2}$$

$$= \sigma_{rc}(i) + \Delta\sigma_{rc} + \sigma_{rp} + \Delta\frac{T}{R} \tag{3}$$

Here the (i) represents the starting (as-installed) condition, Δ represents the rise brought on by loading and the consolidation that follows and $\sigma_{rp} = \frac{T(i)}{R}$. Usually not zero, the initial hoop tension, T(i), depends on the installation technique. Including T(i) into the modeling is essential. This equation can also be formulated as:

$$\sigma_{rs} = \sigma_{rc}(i) + \sigma_{rp} + \Delta\left(\sigma_{rc} + \frac{T}{R}\right) \tag{4}$$

$$\sigma_{rs} = \sigma_{rc}(i) + \sigma_{rp} + \Delta\sigma_{rs} \tag{5}$$

2.2 Test setup and preparation of clay bed

To prepare the soft soil bed, a CBR mould with internal diameter of 150 mm with a height of 175 mm was used. It has a 50 mm high detachable extension collar and a 10 mm thick removable perforated base plate. Grease was applied to the mould's interior surface before it was covered with a thin and smooth polythene sheet to reduce the friction acting between the mould wall and the soil. Also, it makes sure that no water was lost from the sides of the mould. To ensure consistent density and water content throughout experiments, layers of soft clay, each 25 mm thick, were placed inside the tank. To create the clay layer, a predetermined weight of dried clay was ground into a fine powdered form and thoroughly combined with water to achieve a 30% water content. The clay samples were put in the mould and were compacted by the help of a metal rammer weighing 2.6 kg with a height of fall of 310 mm in order to achieve a bulk unit weight of 17.73 kN/m^3. The entire process was repeated for each layer of the clay sample till the height of 175 mm was reached. A thick plastic covering was placed over the prepared test bed, and it was left alone for seven days. Thin-walled cylinder samplers were used to gather undisturbed soil samples from several test bed locations, and their characteristics were assessed. The average value of water content, undrained cohesion and dry density in the test bed was found as 30%, 10 kPa and 17.73 kN/m^3 respectively. The maximum coefficient variance for all the variation was found to be 1.5%.

2.3 Formation of quarry dust column

After the preparation of clay bed, a steel pipe with inner diameter of 25.4 mm and 38 mm were used to prepare a cylindrical hole at the centre of the clay bed to prepare the quarry dust column of diameter 25.4 mm and 38 mm respectively. Clay was extracted from the pipe using a helical steel auger and subsequently and smoothly finished. The excavation was done for a height of 175 mm and 132 mm for the end bearing and floating column respectively. The relative density of the quarry dust was kept at 70% and the amount needed to fill up the cylindrical hole was calculated accordingly. 70% relative density of the quarry dust was chosen because of its its ability to provide optimal soil stability, strength, workability, and drainage characteristics, making it a practical choice both for field applications and controlled experimental setups. Total amount of quarry dust was divided into 5 and 4 batches equally to fill up the height of 175 mm and 132 mm respectively. Each batch was filled up till the required depth and was compacted with a customised circular steel tamping rod having diameter 10 mm and height of 200 mm. For the construction of encased column, once the soft soil is removed the geotextile was wrapped around the pipe and inserted in the column space after which required amount of quarry dust was poured and compacted. Figure 4 represents the unreinforced, column pocket creation for ordinary QDC and encased QDC.

2.4 Test procedure

All the tests were performed on the CBR mould of diameter 150 mm and height of 175 mm. A loading frame with a 50 mm diameter cylindrical plunger is used to conduct the test. A loading machine applies load on a plunger to penetrate the soil mould gradually. The load is applied at a rate of 1.25 mm per minute of penetration. The following load penetrations were recorded: 0.5, 1, 1.5, 2, 2.5, 3, 4, 5, 7.5,

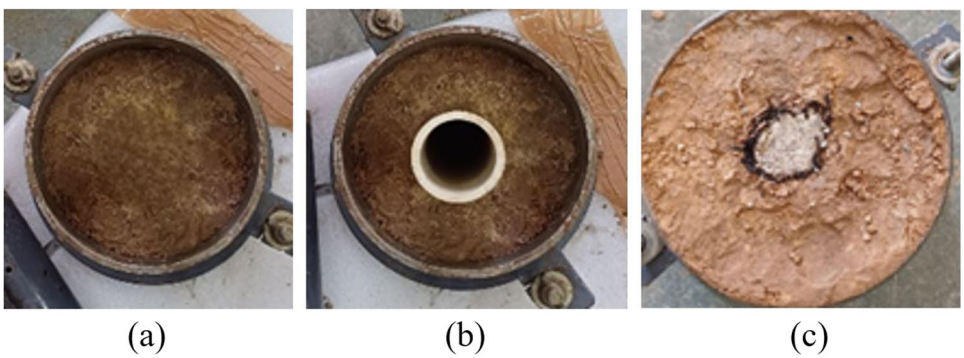


Fig. 4 **a** Unreinforced sample (**b**) Colum pocket creation for ordinary QDC (**c**) Encased QDC

10, and 12 mm. The deformation observed with help of a linear variable differential transducer (LVDT) placed over the testing plate and the mould. Loading setup is depicted in Fig. 5. Table 4 provides an overview of all the tests that were carried out.

3 Results and discussions

3.1 Settlement analysis

In the study, load at penetration of 1, 1.5, 2, 2.5, 3, 4, 5, 7.5, 10 and 12 mm were noted and compared for the load-bearing capacities of different quarry dust column configurations at 12 mm settlement. The results reveal significant variances in load capacity influenced by both kind and diameter of the QDC employed. For a comparison analysis, unreinforced soil sample were tested and at a settlement threshold of 12 mm, the unreinforced soil sample displays a load-bearing capability of 0.596 kN.

In comparison, end-bearing ordinary quarry dust columns (EOQDC) with diameters of 2.54 cm and 3.8 cm exhibit greater load capacities of 0.708 kN and 1.228 kN respectively as presented in Fig. 6. Displaying percentage increases of around 18.79% and 106.04% compared to the un-reinforced soil sample, indicating the effectiveness of increasing stone column diameter in enhancing load-bearing capacities.

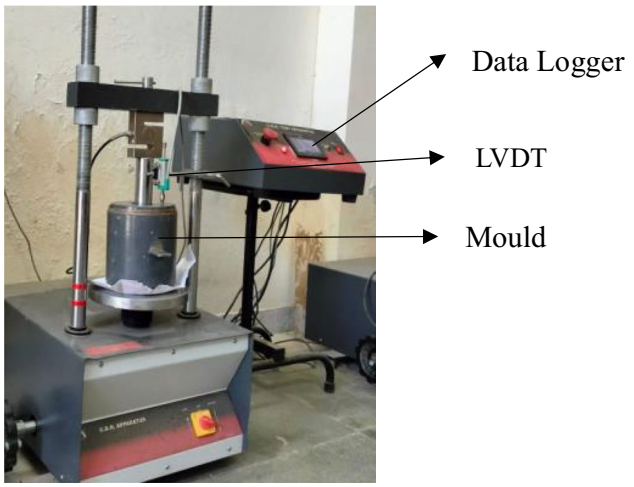


Fig. 5 Test setup

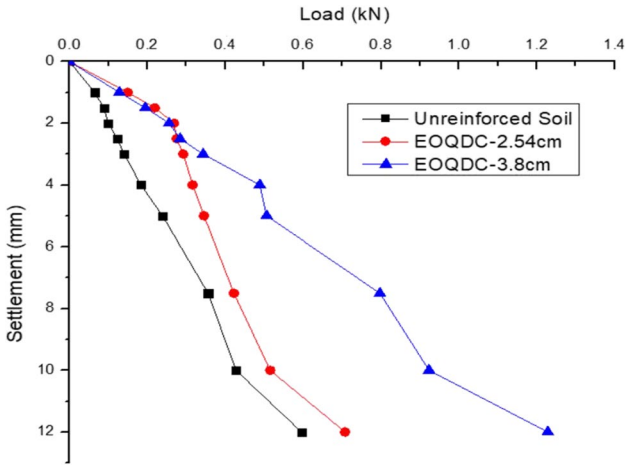


Fig. 6 Load -Settlement behaviour of EOQDC with varying diameter

Similarly, end-bearing encased quarry dust columns (EEQDC), at 2.54 cm and 3.8 cm diameters, display load capabilities of 0.939 kN and 1.763 kN, respectively as presented in Fig. 7. These capabilities show percentage increases of roughly 57.55% and 195.81% above the unreinforced soil sample, demonstrating the increased structural support given by encasement. For EOQDC, increasing the diameter 2.54 cm—3.8 cm indicate a percentage increase of around 73.45%. Similarly, EEQDC when the diameter is raised from 2.54 cm to 3.8 cm, suggesting a percentage increase of roughly 87.75%.

The floating ordinary quarry dust column (FOQDC) with a diameter of 2.54 cm and 3.8 cm results in a load carrying capability of 0.625 kN and 1.079 kN at a settlement of 12 mm (Fig. 8). The results show an increase in the capacity by 4.86% and 81.04% as compared to the unreinforced case. Floating encased quarry dust column (FEQDC) with a diameter of 2.54 cm and 3.8 cm results in a load carrying capability of 0.824 kN and 1.552 kN at a settlement of 12 mm (Fig. 9). The results show an increase in the capacity by 38.25% and 160.40% as compared to the unreinforced case. Increasing the diameter from 2.54 cm to 3.8 cm increases

Table 4 Summary of the tests performed

Reinforcement type	Detailed parameters	
	End bearing column	Floating column
Natural expansive soil	–	–
Expansive soil + OQDC	L = 17.5 cm Dia = 2.54 cm and 3.8 cm	L = 13.2 cm Dia = 2.54 cm and 3.8 cm
Expansive soil + EQDC	L = 17.5 cm Dia = 2.54 cm and 3.8 cm	L = 13.2 cm Dia = 2.54 cm and 3.8 cm

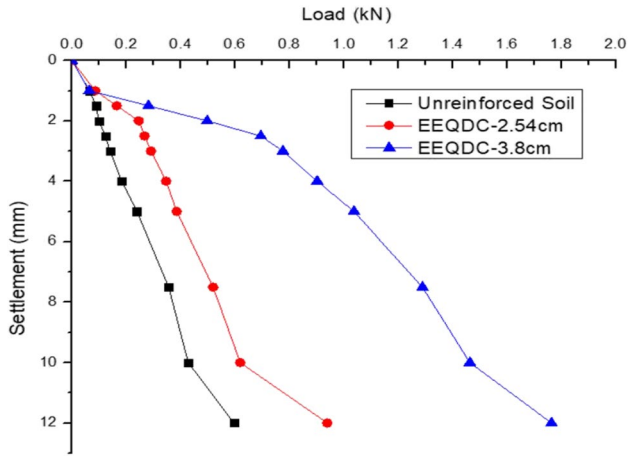


Fig. 7 Load -Settlement behaviour of EEQDC with varying diameter

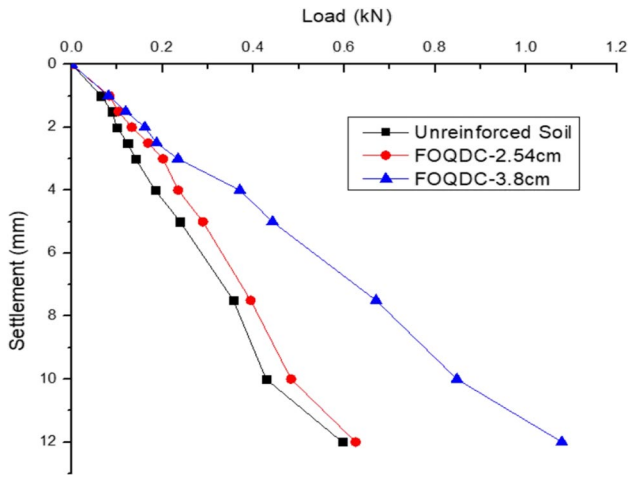


Fig. 8 Load -Settlement behaviour of FOQDC with varying diameter

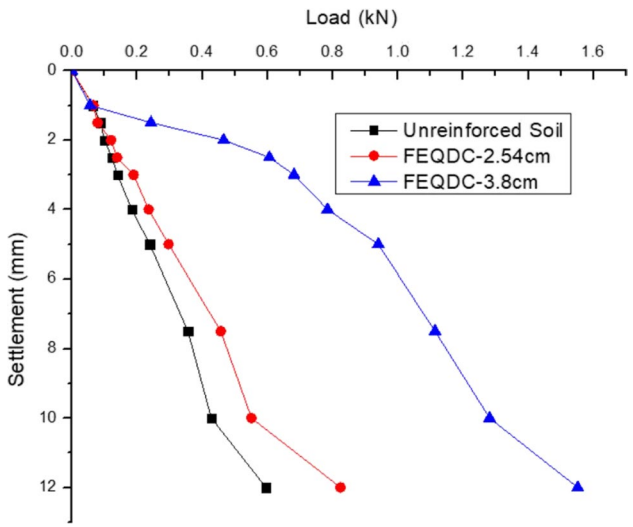


Fig. 9 Load -Settlement behaviour of FEQDC with varying diameter

the carrying capacity by 72.64% for FOQDC and by 88.35% for FEQDC.

Figure 10 presents a comparative analysis of the load-settlement characteristics between end-bearing and floating columns for varying diameter. Encasement provided a better load carrying capacity as compared to ordinary scenario for both end-bearing as well as the floating conditions. For EEQDC of 2.54 cm and 3.8 cm diameter column, an increase of 32.64% and 43.56% respectively as compared of EOQDC. While, for FEQDC of 2.54 cm and 3.8 cm, and increase of 31.84% and 43.83% respectively as compared to FOQDC. When the results were compared between end-bearing and floating end conditions, EOQDC of 2.54 cm and 3.8 cm diameter showed an increase of around 13.5% better than FOQDC. Also, EEQDC fared an increase of around 13.7% as compared to FEQDC. The increase in the load-carrying capacity of Quarry Dust Columns (QDC) with larger diameters, both in end-bearing and floating cases, is primarily due to the larger cross-sectional area that helps distribute the applied load over a greater surface. A wider column provides more contact area with the surrounding soil, improving the overall load distribution and reducing localized stress concentrations. When ordinary QDC is encased with geotextile, similar results are observed because the geotextile provides additional confinement to the quarry dust, preventing lateral bulging and enhancing the column's overall stability. The geotextile material, by offering external support, increases the strength of the column by reducing the radial expansion under load and maintaining the integrity of the quarry dust inside the column. The ultimate load carrying capability at a settlement of 12 mm of all the cases has been presented in Fig. 11.

3.2 Load improvement factor

Load Improvement Factor (LIF) is an important indicator of the degree to which quarry dust columns reduces settlement and improves the load-bearing capacity. It measures the degree to which the presence of QDC enhances stress distribution. Consequently, the ratio of the reinforced soil's load-bearing capacity to that of the unreinforced soil is known as LIF. Figure 12 represents the LIF of EOQDC and EEQDC. At 12 mm settlement LIF for 2.54 cm EOQDC and EEQDC was found as 1.188 and 1.786 respectively. Similarly, for 3.8 cm diameter column of EOQDC and EEQDC was calculated as 2.07 and 2.96 respectively. From the LIF data, it was clearly observed that encapsulating the ordinary quarry dust column improves the load carrying factor which gives additional support and helps transfer the load more effectively to the surrounding soil. Figure 13 gives the LIF of FOQDC and FEQDC for both the 2.54 cm and 3.8 cm diameter columns. At 12 mm settlement, LIF for 2.54 cm diameter column of FOQDC and FEQDC was found as 1.05 and

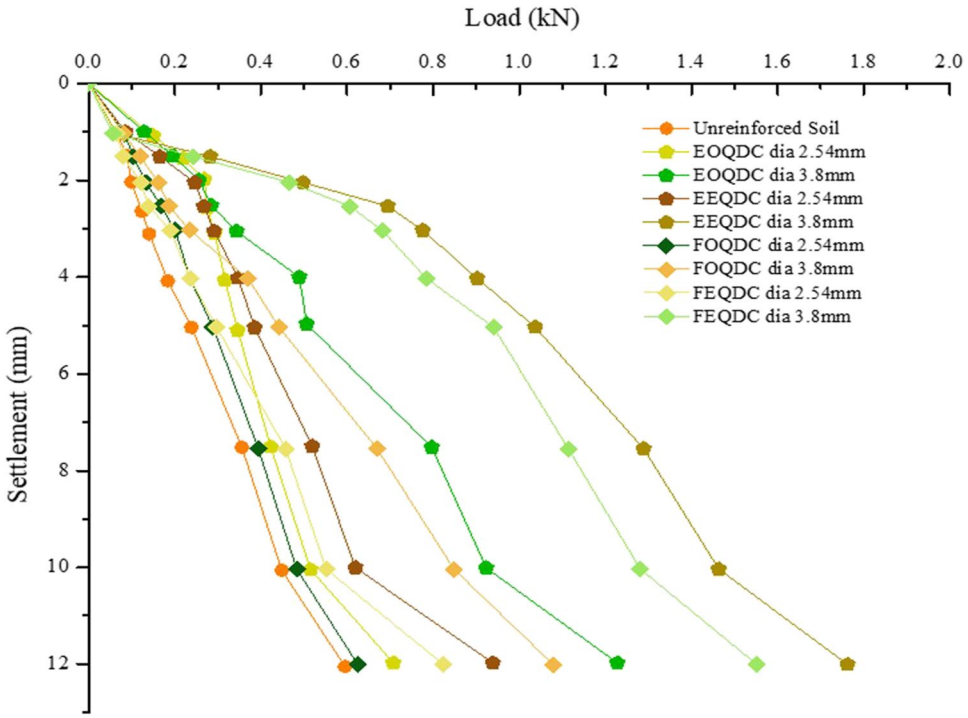


Fig. 10 Load -Settlement comparison of all the cases carried out in the current study

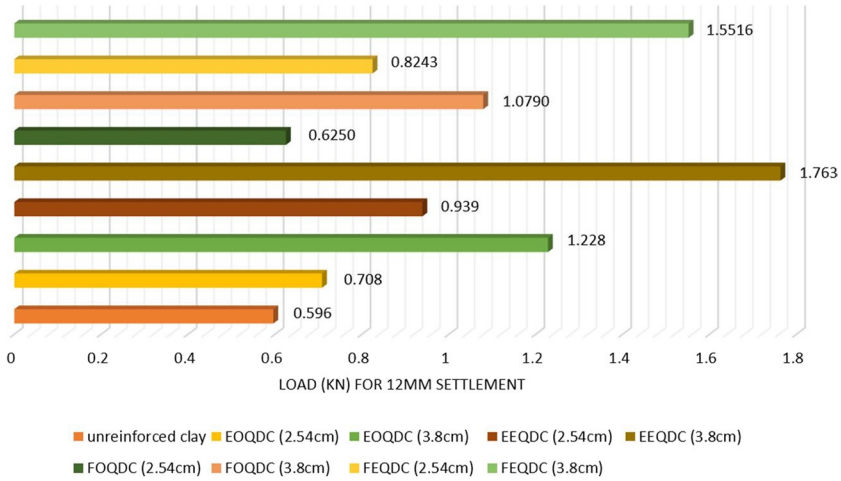


Fig. 11 Ultimate load carrying capacity at a settlement of 12 mm for all the cases

1.38 respectively. Likewise, for 3.8 cm diameter for FOQDC and FEQDC was calculated as 1.81 and 2.60 respectively.

Figure 14 represents a comparative analysis of the LIF for both end-bearing and floating case for the two different diameters of the columns for both ordinary and encased scenario. It was found that the 3.8 cm diameter column with the encasement provides the maximum value of LIF of 2.96 and thereby can be said to be carrying the maximum load bearing capability.

3.3 Stress–strain analysis

Stress responses to varying strain for soil and different quarry dust column configurations, including Ordinary and Encased Quarry Dust Columns of diameters 2.54 cm and 3.8 cm, reveals significant distinctions in their load-bearing capacities. For this analysis end bearing scenario have been presented which is depicted in Fig. 15. Ordinary QDC of 2.54 cm exhibit consistently higher stress than soil, with stress values escalating from 43.63 kPa at zero strain to 360.43 kPa at 0.0686 strain, demonstrating the reinforcement benefit of smaller diameter columns. In contrast, Ordinary QDC with a diameter of 3.8 cm while starting at a lower initial stress (20.74 kPa) experience a significant increase, reaching a peak of 625.36 kPa. This suggests that larger diameters provide better resistance to loads. Initially, encased QDC with a diameter of 2.54 cm have lower stress compared to OQDC. This is likely because of adjustments in the interaction between the encasement and the

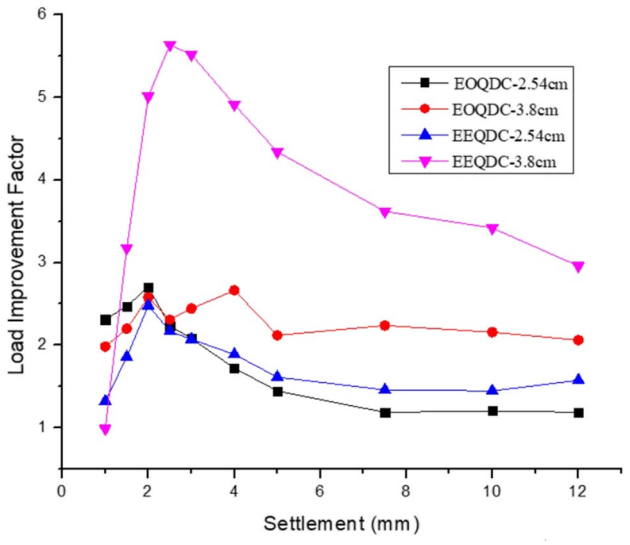


Fig. 12 LIF of end bearing conditions for both ordinary and encased scenario

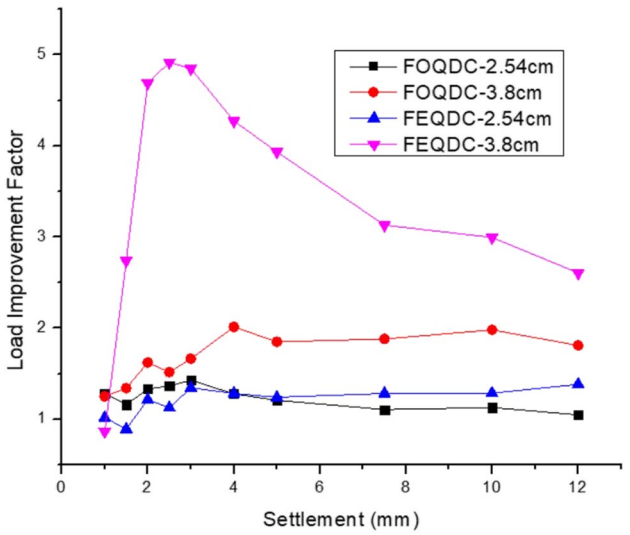


Fig. 13 LIF of floating conditions for both ordinary and encased scenario

soil. However, surpass them as strain increases, reaching 478.23 kPa at 0.0686 strain, highlighting the advantages of encasement in enhancing column stiffness and load transfer efficiency. The most significant performance is found with Encased QDC of 3.8 cm reaching an outstanding 897.87 kPa at the highest strain, underlining that the combination of bigger diameter with encasement enhances stress capacity delivering improved reinforcement against load-induced deformations primarily because of the lateral confinement maintain the column's structural integrity. This comparison underlines the utility of encasement and greater diameter

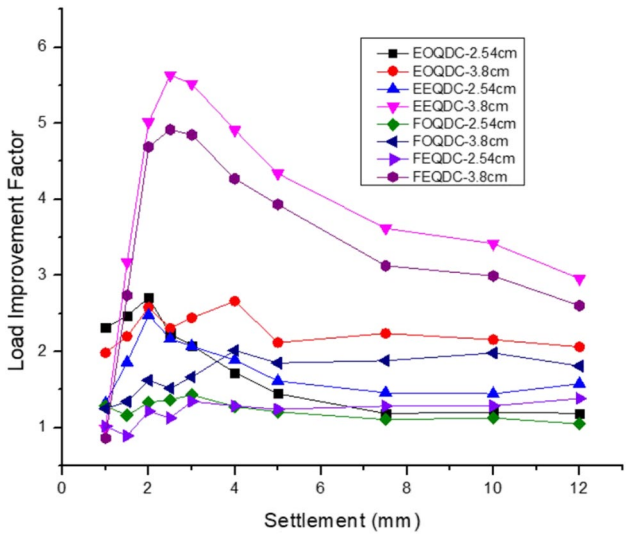


Fig. 14 Comparison of LIF for end bearing and floating conditions for both ordinary and encased scenario

in stone columns delivering improved stress performance, making them appropriate for application demanding strong ground improvement and higher load-bearing capabilities.

3.4 Failure mechanism

Quarry dust columns (QDC), both encased and ordinary, can experience various failure modes depending on their configuration (end-bearing or floating) and presence of encasement in soft clay. After testing, distorted granular piles were developed through taking out quarry dust and adding plaster of Paris into the holes. Ordinary QDCs exhibited bulging failure. This occurred from the top of the column (crest) for a distance of D to 1.5D (D being the column diameter). Encasement can mitigate this by providing lateral confinement and improving stress distribution. For the floating columns, punching failure was observed which resulted in the penetration of column up to a certain depth along with the tip yielding or bulging because of the lower shear stress to support the applied load. Failure mechanism of both the cases have been represented in Fig. 16. Encasing the floating columns also resulted in the mitigation of tip yielding or bulging along with the prevention of penetration into clays up to a certain extent.

3.5 Comparison with the existing literature

A direct comparison with some of the existing literature is not feasible due to the differences in materials used as no studies have used waste quarry dust as a soil modifier. The materials investigated in our study differ substantially from those used in many of the referenced works, and such variations may lead to

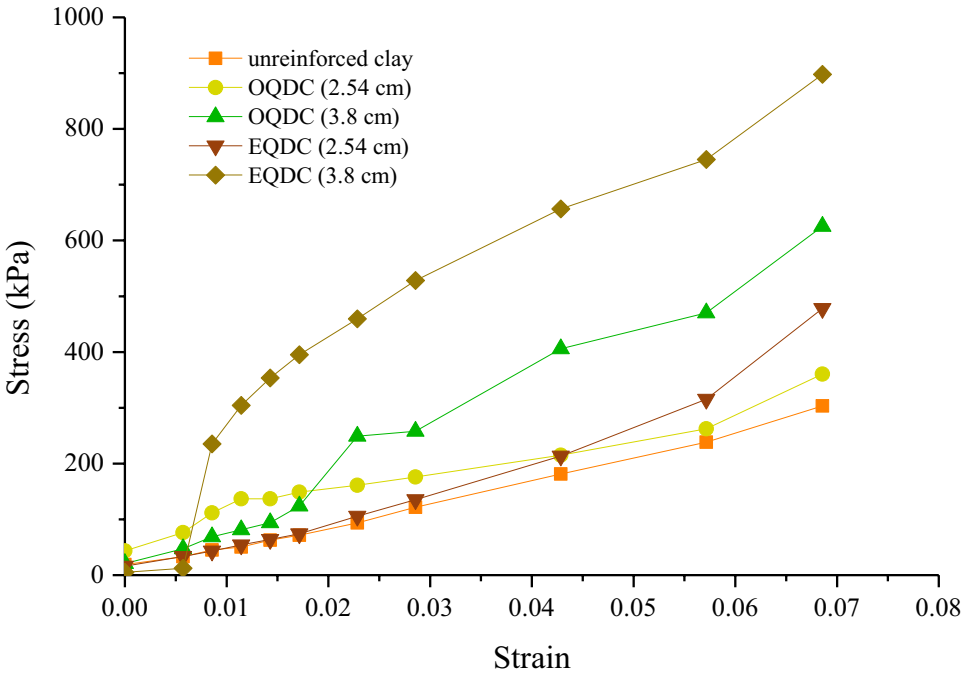


Fig. 15 Stress–Strain relationship of end bearing OQDC and EQDC for varying diameter columns

inherent differences in performance or behaviour. Despite this limitation, we have taken care to compare our results with studies that employ materials that are more similar to those used in our work. Load-Settlement behaviour has been compared for OQDC and EQDC end bearing configurations with the other studies. These studies provide a more relevant context for comparison, and the results are presented in Fig. 17. The comparison results clearly demonstrates that the general behaviour and trends in our study are consistent with those seen in the literature for similar materials. This suggests that, while material differences exist, the underlying principles and trends are widely applicable across different studies in this field.

4 Conclusion

This study investigated the efficiency of quarry dust columns (QDC) in improving the stability of soft clay soils by means of a comparison between encased quarry dust columns (EQDC) and ordinary quarry dust columns (OQDC) over various diameters (2.54 cm and 3.8 cm) for end bearing and floating scenarios. The load-bearing capacity and settling properties were assessed using California Bearing Ratio (CBR) test. The following conclusion can be observed:

1. Load-bearing capacity increased significantly in both EOQDC and EEQDC when compared to untreated soil. Larger diameter column (3.8 cm) consistently outperformed smaller ones (2.54 cm), hence increasing the diameter helps enhance performance.
2. Encased columns (EQDC) demonstrated even greater load-bearing capacities than ordinary columns (OQDC) for both end bearing and floating conditions.
3. Both OQDC and EQDC show a notable increase in SIF value compared to untreated soil and indicating better stress distribution and lower settlement. Greater SIF values came from larger diameter columns for both OQDC and EQDC. EQDC consistently showed a higher SIF than OQDC with the same diameter, emphasizing the efficacy of encasement in stress distribution. The results were observed in both end bearing and floating conditions.
4. Ordinary ODCs with a 2.54 mm diameter demonstrate higher stress levels than untreated soil. However, larger diameter ODCs (3.8 mm) show a notable increase in stress, indicating that larger diameters offer enhanced resistance to loading conditions. EQDC initially shows lower stress values compared to OQDC due to interaction between the encasement material and the surrounding soil but, as strain is increased, encased columns surpass OQDC.
5. Failure mechanism for end bearing columns were bulging, whereas floating columns fails in tip yielding/bulging and punching. Maximum bulging was noted at a depth of D-1.5D from the top of the column. Encasing the columns resulted in prevention of such failures.
6. Reinforced floating granular piles are a better alternative to unreinforced end-bearing piles.
7. Conventional stone columns made up of aggregate can be effectively replaced by the waste material like quarry dust without compromising much of its efficiency.

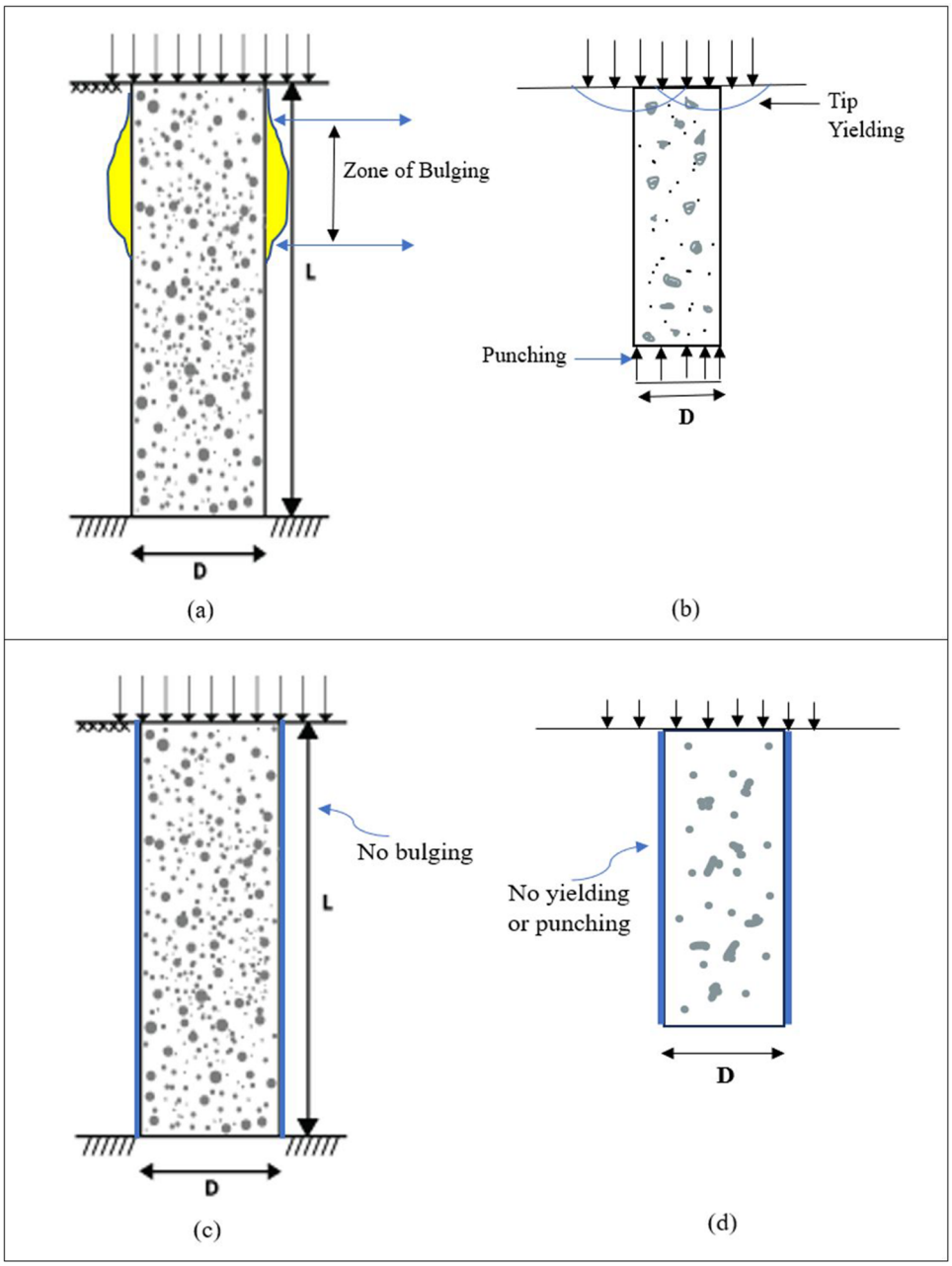


Fig. 16 Failure mechanism of (**a**) End Bearing Column-OQDC (**b**) Floating Column-OQDC (**c**) End Bearing Column- EQDC (**d**) Floating Column- EQDC

5 Practical consideration, limitations and future scope

To implement Quarry Dust Columns (QDC) in construction projects, start with a thorough site assessment to understand the soil properties, as these determine the suitability of QDC for soil stabilization. Soft clays and weak soils are ideal candidates for this technique. The quarry dust used should be clean and free from organic impurities; its particle size distribution and compaction characteristics should be tested to ensure effectiveness. For installation, vertical columns of quarry dust are placed at the specified intervals and depth, ensuring proper compaction for optimal stability. In the case of EQDC, the quarry dust columns are encased in jute geotextile, which helps provide additional confinement and prevent lateral spreading. Limitations of this method include the potential decay of jute over time and variable performance under extreme environmental conditions. While jute offers several advantages as an eco-friendly and cost-effective material for encasing quarry dust columns, its biodegradability and susceptibility to moisture and UV degradation can reduce its long-term effectiveness. The decay of jute over time can lead to a loss of confinement, reduced load-bearing capacity, and increased settlement, ultimately

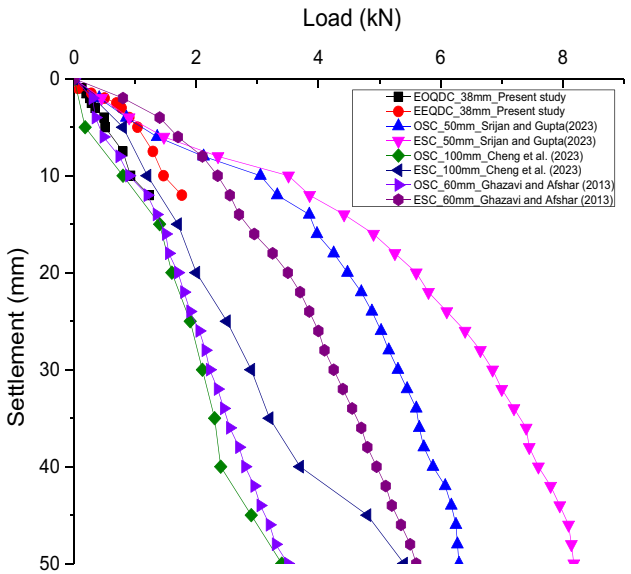


Fig. 17 Comparison of load-settlement behaviour of the present study with existing literatures

impacting the performance of the EQDC system. To further enhance the effectiveness and durability of Quarry Dust Columns (QDC), future research should explore the trategies to mitigate decay can help improve the longevity and stability of jute-encased quarry dust columnsand also the development of hybrid materials that combine the benefits of natural geotextiles, like jute, with synthetic materials, offering better long-term stability and resistance to environmental degradation.

Author contribution The author confirms the sole responsibility of the following: Writing- original draft, investigation, conceptualization, methodology.

Funding The author declares that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Declarations

Competing interest The author has no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Alam P, Bawa S (2024) Sustainable Use of Waste Materials in Stone Columns: A Review. Indian Geotech J 1:1–22

Ali K, Shahu JT, Sharma KG (2014) Model tests on single and groups of stone columns with different geosynthetic reinforcement arrangement. Geosynth Int 21(2):103–118

Almeida MSS, Hosseinpour I, Riccio M (2013) Performance of a geosynthetic encased column (GEC) in soft ground: numerical and analytical studies. Geosynth Int 20(4):252–262

Black JA, Sivakumar MR, Madhav MR, Hamiil GA (2007) Reinforced stone columns in weak deposits: laboratory model study. J Geotech Geoenviron Eng 133(9):1154–1161

Castro J, Sagaseta C (2011) Deformation and consolidation around encased columns. Geotext Geomembranes 29(3):268–276

Castro J, Sagaseta C (2013) Influence of elastic strains during plastic deformation of encased stone columns. Geotext Geomembranes 37(4):47–53

Cheng Y, Cai X, Mo H, Gu M (2023) Numerical Analysis on the Behavior of Floating Geogrid-Encased Stone Column Improved Foundation. Buildings 13(7):1609

Gharavi H, Yazdani E, Alidoust P, Shariatmadari N, Yazdani F, Niavol KP (2023) Utilization of Shredded Waste Tire as a Substitute for Gravel in Stone Column Groups. KSCE J Civ Eng 27:198–206

Ghazavi M, Afshar JN (2013) Bearing capacity of geosynthetic encased stone columns. Geotext Geomembranes 38:26–36

Greenwood DA (1970) Mechanical improvement of soils below ground surface. In Inst Civil Engineers Proc, London, UK

Hosseinpour I, Riccio M, Almeida MSS (2014) Numerical evolution of a granular column reinforced by geosynthetics using encasement and laminated disks. Geotext Geomembr 42:363–373

Hughes JMO, Withers NJ, Greenwood DA (1975) A Field Trial of the Reinforcing Effect of a Stone Column in Soil. Geotechnique 25(1):31–44

Javid A, Pandey BK, Srijan, (2023) Performance Evaluation of Clayey Soil Considering Impact of Lime During Electrokinetic Consolidation. Arab J Sci Eng 49(4):4527–4537

Ji M, Wang J, Zheng JJ, Zheng Y (2023) No Access Contribution of geosynthetic to the shear strength of geosynthetic encased stone columns. Geosynth Int 31(3):358–371

Keykhosropur L, Soroush A, Imam R (2012) 3D numerical analyses of geosynthetic encased stone columns. Geotext Geomembranes 35:61–68

Kumar N, Kumar R, Nirola BS, Jaiswal A (2023) Numerical Analysis of a Dual-Layer Geosynthetic-Encased Stone Column Installed in Soft Soil. Adv Civil Eng 41(6):3253–3288

Malarvizhi SN, Ilamparuthi K (2007) Comparative study on the behavior of encased stone column and conventional stone column. Soils Found 47(5):873–885

Mazumder T, Rolaniya AK, Ayothiraman R (2018) Experimental study on behaviour of encased stone column with tyre chips as aggregates. Geosynth Int 25(3):259–270

Mohapatra SR, Rajagopal K, Sharma J (2016) Direct shear tests on geosynthetic-encased granular columns. Geotext Geomembr 44:396–405

Murugesan S, Rajagopal K (2008) Shear load tests on granular columns with and without geosynthetic encasement. Geotech Test J 32(1):35–44

Pulko B, Majes B, Logar J (2011) Geosynthetic-encased stone columns: analytical calculation model. Geotext Geomembranes 29(1):29–39

Raithel M, Kirchner A, Schade C, Leusink E (2005) Foundation of constructions on very soft soils with geotextile encased columns-state of the art. Innovations in Grouting and Soil Improvement, pp 1-11

Rajagopal K, Krishnaswamy NR, Latha MG (1999) Behavior of sand confined with single and multiple geocells. Geotext Geomembranes 17(3):171–184

Saxena S, Roy LB, Gupta PK, Kumar V, Paramasivam P (2024) Model tests on ordinary and geosynthetic encased stone columns with recycled aggregates as filler material. Geo-Engineering 15(1):1–13

Shahu JT, Reddy YR (2014) Estimating long-term settlement of floating stone column groups. Can Geotech J 51(7):770–781

Srijan GAK (2023) Horizontally Layered and Vertically Encased Geosynthetic Rein-forced Stone Column: An Experimental Analysis. Appl Sci 13:8660

Srijan GAK (2025) Sustainable material as a column filler in soft clay bed reinforced with encased column: numerical analysis. Sci Rep 15:1650

Srijan S, Gupta AK (2023) Vertically and Horizontally Reinforced End-Bearing Stone Column: An Experimental and Numerical Investigation. Appl Sci 13:11016

Srijan, Somra K, Kharra P, Singh S (2024) Use of Sustainable Material as Column Filler in Soft Clay Bed Reinforced with Stone Column: Numerical Study. In: IOP Conference Series: Earth and Environmental Science 1326(1)012115. IOP Publishing

Van Impe WF (1989) Soil Improvement Techniques and their Evolution. Balkema, Rotterdam, Netherlands, pp 63–66

Wu CS, Hong YS (2009) Laboratory tests on geosynthetic-encapsulated sand columns. Geotext Geomembranes 27(2):107–120

Zhang L, Zhao M (2015) Deformation analysis of geotextile encased stone columns. ASCE Int J Geomech 15(3):04014053

Zhang Y, Chan D, Wang Y (2012) Consolidation of composite foundation improved by geosynthetic-encased stone columns. Geotext Geomembranes 32:10–17

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Eur. Phys. J. Plus (2025) 140:278
https://doi.org/10.1140/epjp/s13360-025-06180-9

THE EUROPEAN PHYSICAL JOURNAL PLUS

Regular Article

Enhancing the thermoelectric performance of Janus MoSSe monolayer via pressure

Anisha[1], **Ramesh Kumar**[1,a], **Mukhtiyar Singh**[2], **Sunita Srivastava**[3], **Tankeshwar Kumar**[3]

[1] Department of Physics, Guru Jambheshwar University of Science & Technology, Hisar 125001, India
[2] Department of Applied Physics, Delhi Technological University, Delhi 110042, India
[3] Department of Physics & Astrophysics, Central University of Haryana, Jant-Pali, Mahendergarh 123031, India

Received: 26 October 2024 / Accepted: 26 February 2025
© The Author(s), under exclusive licence to Società Italiana di Fisica and Springer-Verlag GmbH Germany, part of Springer Nature 2025

Abstract By combining first-principles computations with the semi-classical Boltzmann transport equations, a systematic investigation of the structural, electronic and thermoelectric properties of the MoSSe Janus monolayer is conducted under pressure. The monolayer semiconducting nature is indicated by the band gap value (E_g = 1.5 eV), which may be further tuned from 0.56 to 1.67 eV by applying pressure in the -3GPa to + 2GPa range. The figure of merit (ZT) for p (n)-type carriers at 300 K in the absence of pressure is computed to be 0.67 and 0.59. The power factor has enhanced from 16.59 (27.21) Wm^{-1} K^{-2} to 227.15 (159.50) Wm^{-1} K^{-2} for *n* (*p*)-type carriers by applying an external pressure of -1 GPa to the Janus monolayer. For *n* (*p*) -type doping at 300 K, the corresponding maximum value of ZT is 0.82 (0.78), which is 39% (14%) greater for *n* (*p*) type than for pure MoSSe Janus monolayer. When the pressure is increased to + 3 GPa, the value of ZT for n-type doping is further increased to 0.73, which is 24% higher than the value for pure monolayer. It is possible for a pure Janus monolayer to undergo n-type doping under pressure due to the shifting of the conduction band minima and valence band maxima. This study presents an attractive approach for manipulating the material thermoelectric properties through external pressure application.

1 Introduction

Several factors including population growth, industrialization and technological improvement are contributing to the very rapid rise in the global energy demand. The energy needs will be met by using fossil fuels and other non-renewable resources [1], which will lead to the production of massive amounts of waste and the eventual depletion of the resources owing to intensive use. The future energy problems can be resolved by anticipating a rise in energy efficiency and a decrease in energy consumption [2].

The environmentally friendly and efficient methods of resolving this problem are the conversion of waste heat into electrical energy using thermoelectric material [3, 4]. The past three decades have seen a rise in the importance of low-dimensional materials [5], especially nanosized materials, in many scientific and technological domains. Also, the 2D materials can have nano-bio interactions which can be further used in bio medical applications [6]. This is because 2D materials are preferred to their bulk equivalent due to their distinctive and outstanding properties [7, 8].

Theoretical predictions and experimental synthesis of several 2D transition metal dichalcogenides (TMDCs) have been made [9–11]. The density of states (DOS) of the valence and conduction bands [12] in low-dimensional materials has been increased close to the Fermi level, causing an increase in power factor (σS^2) that results in good thermoelectric performance with the mathematical expression for figure of merit (ZT) as

$$ZT = \frac{S^2 \sigma T}{\kappa} \tag{1}$$

where the total thermal conductivity (electronic and lattice), Seebeck coefficient, electrical conductivity and absolute temperature (in kelvin) are represented as κ, S, σ and T [13]. All these parameters are interconnected in such a way that it is impossible to change the value of ZT by changing just one variable, as they are all dependent on one another. Thermoelectric materials have a special tendency for using the Seebeck effect to transform heat energy into electrical energy, whereas the reverse process uses the Peltier effect.

The study of the Mo-based TMDCs, such as $Mo(S/Se)_2$, $W(S/Se)_2$ and $Zr(S/Se)_2$ have been extensively investigated [14–17]. Janus monolayers are formed with two different types of chalcogenides in a single monolayer and by breaking the mirror symmetry [18–21]. Thermoelectric properties have been well investigated for Janus HfSSe, and WS-X (X = Se & Te) monolayers [22–26]. The chemical vapour deposition method was used by Zhang et al. [27] to synthesize MoSSe, making it the most suitable and deserving

[a] e-mail: ramesh85@gjust.org (corresponding author)

Published online: 08 April 2025

candidates of further development. Based on how the chalcogenides are arranged, these TMDC monolayers are split into two phases called the 2H and 1 T phases. In comparison with ZrS_2 monolayer, ZrSSe [28] monolayer has been predicted in 1 T-phase and has proven to be a promising n-type thermoelectric material. Thermoelectric performance of TMDCs materials is improved by diverse ways like doping [29, 30], applying strain [31, 32], dimension and stacking engineering [8, 33, 34], applying pressure [35], inducing defects [36] and converting the 2D materials into nanofibres can be used in tissue regeneration [37].

Thermoelectric properties of Sb_2Se_3 monolayer have been studied by H. A. R. Aliabad et al. [38] under different values of pressure. Y. Zhang et al. [39] predicted the pressure induced thermoelectric properties of SnSe crystal and found that the value of ZT has been enhanced up to 1.7 from 0.9 at 600 K under the pressure of 6 GPa. Thermoelectric average power factor has been enhanced up to 33% under the pressure of 20 GPa at 800 K temperature, and corresponding enhancement in thermoelectric performance of SnS_2 material has been studied by Y. Javed et al. [40]. D. Zou et al. [35] have studied about the tuning of electronic parameters under external pressure and gave the valuable insight on the enhanced thermoelectric performance of layer $SnSe_2$ material. In the case of $NbSe_2$, B. Hou et al. have studied atomic tracking techniques to systematically explore the thermally-driven structural evolution [41].

Keeping in mind the above literature studies on various monolayers, the present investigation deals with the study of structural, electronics and thermoelectric properties of MoSSe Janus monolayers using pressure. A systematic investigation has been performed for the structural and electronic properties. Band structure has been altered by applying the external pressure, and it has found that band gap gets tuned. Phonon dispersion curve at different pressure has been drawn and investigate that chosen monolayer at different pressure is found to be thermodynamically stable. Using the semi-classical Boltzmann transport equation, thermoelectric properties of monolayer under different pressure have been investigated. It is observed that MoSSe Janus monolayer is a promising thermoelectric material.

By demonstrating how pressure influences MoSSe's unique structural asymmetry, the research highlights the material's potential for fine-tuning of its band gap, carrier transport, and phonon interactions. This opens up a practical pathway for enhancing the performance of MoSSe in energy conversion applications, such as thermoelectric, where precise control over electronic and thermal transport properties is essential for improving efficiency.

2 Computational details

Ab-initio calculations for MoSSe Janus monolayer were carried out in framework of density functional theory (DFT) as implemented in Quantum Espresso (QE) software [42]. The Janus monolayer has been designed by replacing the one layer of sulphur (S) in MoS_2 by the layer of selenium (Se), and structure has been relaxed without using any constraint by taking a vacuum of 20 Å along the Z-direction. The Janus monolayer has been totally relaxed under the convergence criteria for energy and force up to 0.0001 eV and 0.001 eV/Å^2. As an exchange correlation functional, generalized gradient approximation (GGA) has been adopted along with Perdew-Burke-Ernzerhof (PBE) form [43]. For the core interaction, ultrasoft pseudopotentials were used and plane wave basis set with optimized energy up to 90 Ry was used for the incorporation of core and valence electrons interactions. Kinetic energy cut-off and k-mesh was converged against the total energy, for reducing the computational cost. For the self-consistent calculations, an optimized k-mesh of Monkhorst Pack of $16 \times 16 \times 1$ is sampled for Brillouin zone. The convergence tests for k-point and cut-off energy have been provided in supplementary information (#). Thermodynamical stability has been checked with the phonon dispersion curve by using the phonopy software [44]. The thermoelectric parameters such as Seebeck coefficient, electrical conductivity, and electronic thermal conductivity have been evaluated by solving the semi-classical Boltzmann transport equations (BTEs) using the interface of QE software with BoltzTraP code [45]. The well convergent transport parameters have been obtained by taking the high Monkhorst Pack of $60 \times 60 \times 1$, and these are found to be the function of relaxation time (τ). The relaxation time is determined by using the deformation potential theory [31, 32, 46]. Further, the contribution of phonon to the lattice thermal conductivity is calculated by coupling of QE with phono3py code [47]. The finite displacement of atoms was generated by making the supercell of $3 \times 3 \times 1$, to obtain the second and third order force constants. Generated supercell was sampled with $6 \times 6 \times 1$ k-points mesh and used optimized q-mesh of size $25 \times 25 \times 1$. All the above calculated parameters are further used to calculate the value of figure of merit (ZT).

3 Results and discussion

3.1 Structural properties

MoSSe Janus monolayer is made up of three layers, one layer of transition metal atom (Mo) sits between the two distinct chalcogenide types (S & Se). In addition, as illustrated in Fig. 1, it had a honeycomb lattice structure like graphene, composed of Se/S atoms in place of the S/Se atoms in MoS_2/$MoSe_2$.

Figure 1a and b shows the side and top views of the Janus monolayer, respectively, whereas Fig. 1c shows the top view. The MoSSe Janus monolayer asymmetric structure results from its two different chalcogenides, which give its exceptional properties. In

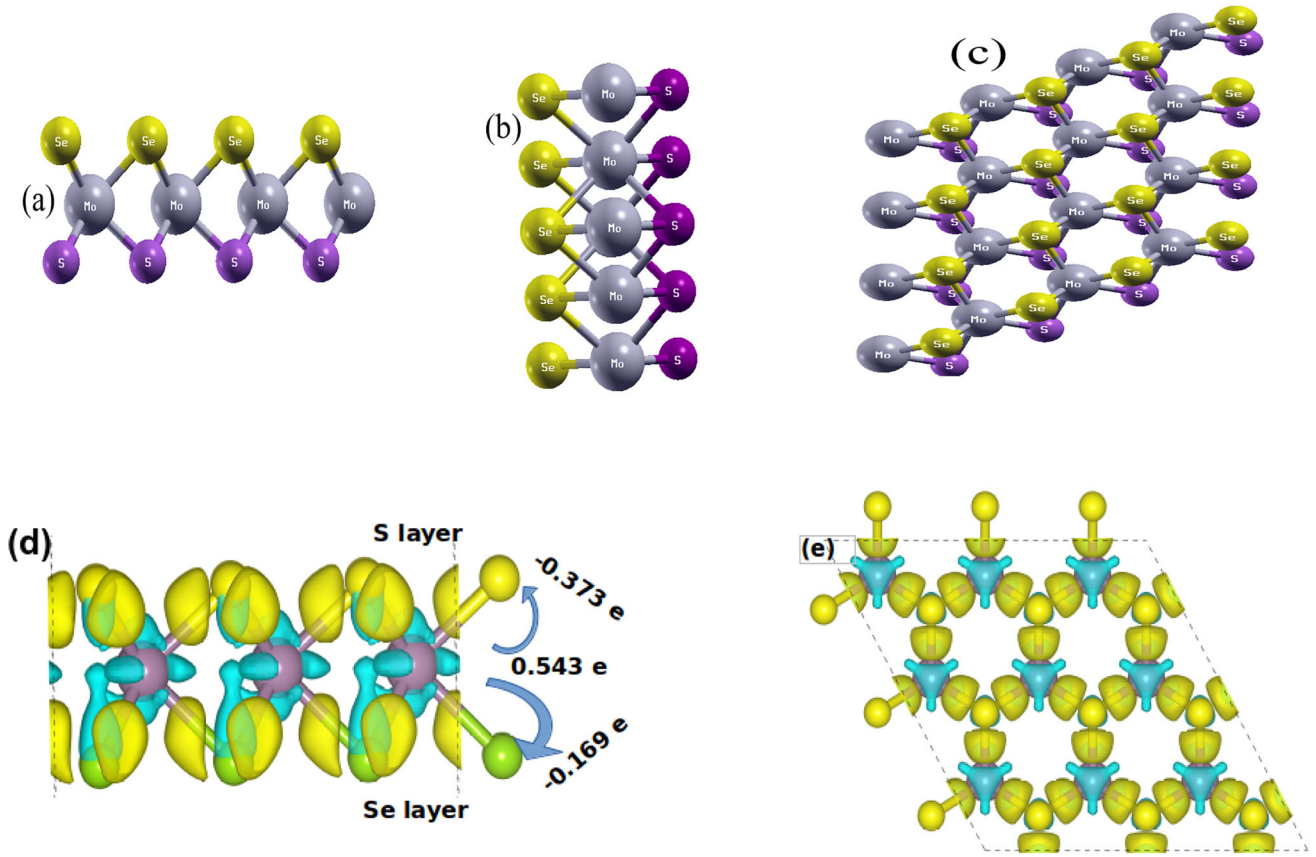


Fig. 1 Crystal structure of MoSSe Janus monolayer (**a**), (**b**) Side views and (**c**) Top view having different orientation, whereas, (**d**) Side view and (**e**) Top view of charge density difference (cdd) plots

Table 1 Lattice constant (a) of the monolayer and their literature values

Monolayer	MoS_2	MoSSe	$MoSe_2$
a (Å)	3.17 [48], 3.18 [26]	3.25 [present work, 26], 3.23 [48]	3.30 [26, 48]

this case, structural asymmetry causes reflection symmetry to break regarding the central transition metal, while MoS_2 and $MoSe_2$ retain reflection symmetry. A vacuum of 20 Å has been added along the Z-direction to prevent the interaction between the layers.

Figure 1d and e shows the side view and top view of charge density difference plots. Due to the presence of layers of different chalcogen atoms in the upper and lower environment of transition metal layer, there is breaking in the symmetry of charge carriers and electronic band structure which also affects the thermoelectric performance of material. The charge transport is not uniform across the material due to presence of different environment. Here, in Fig. 1d, there is accumulation of charge 0.373 e from Mo to S atom and 0.169 e from Mo to Se atom. Transfer of different amount of charge shows asymmetry in chemical environment in lower and upper phase.

Since the atomic radius of Se is greater than that of S, the lattice constant of MoSSe lies between the lattice constants of MoS_2 and $MoSe_2$, as shown in Table 1 [26, 48]. A very small discrepancy in lattice constant is due to computational error. Figure 2a illustrates how the lattice constant varies as a function of applied pressure (in GPa).

3.2 Electronic properties

The MoSSe Janus monolayer band structure is studied along the K–Γ–M–K–Γ high symmetry path in the first Brillouin zone.

As shown in Figs. 2a and 3, the contribution of different orbitals in the formation of bands varies with pressure i.e. the contribution of 3d orbital of Mo atom, 3p orbital of Se atom and 2p orbital of S atom varies with change in pressure. The thermoelectric properties of materials are depended on the formation of bands. That is why transport properties get tuned with pressure variation.

The electronic band structure will have a direct band gap nature with a value of 1.56 eV, where the valence band maxima (VBM) and conduction band minima (CBM) are both located at the high symmetric point K as illustrated in Fig. 4a. The obtained results are similar to the previously published research [48].

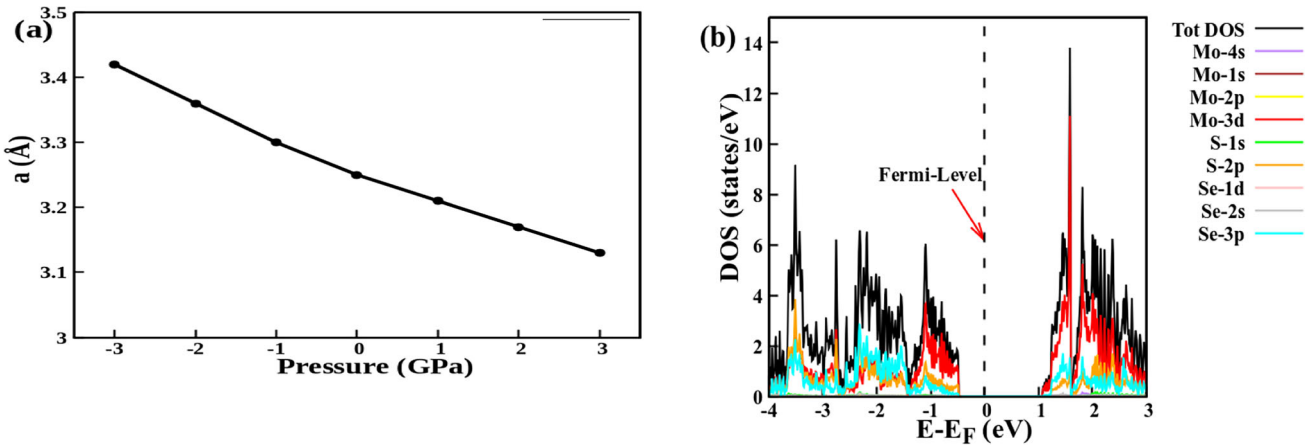


Fig. 2 **a** Variation of lattice constant with different values of pressure (GPa) and **b** Density of states corresponding to pure Janus monolayer

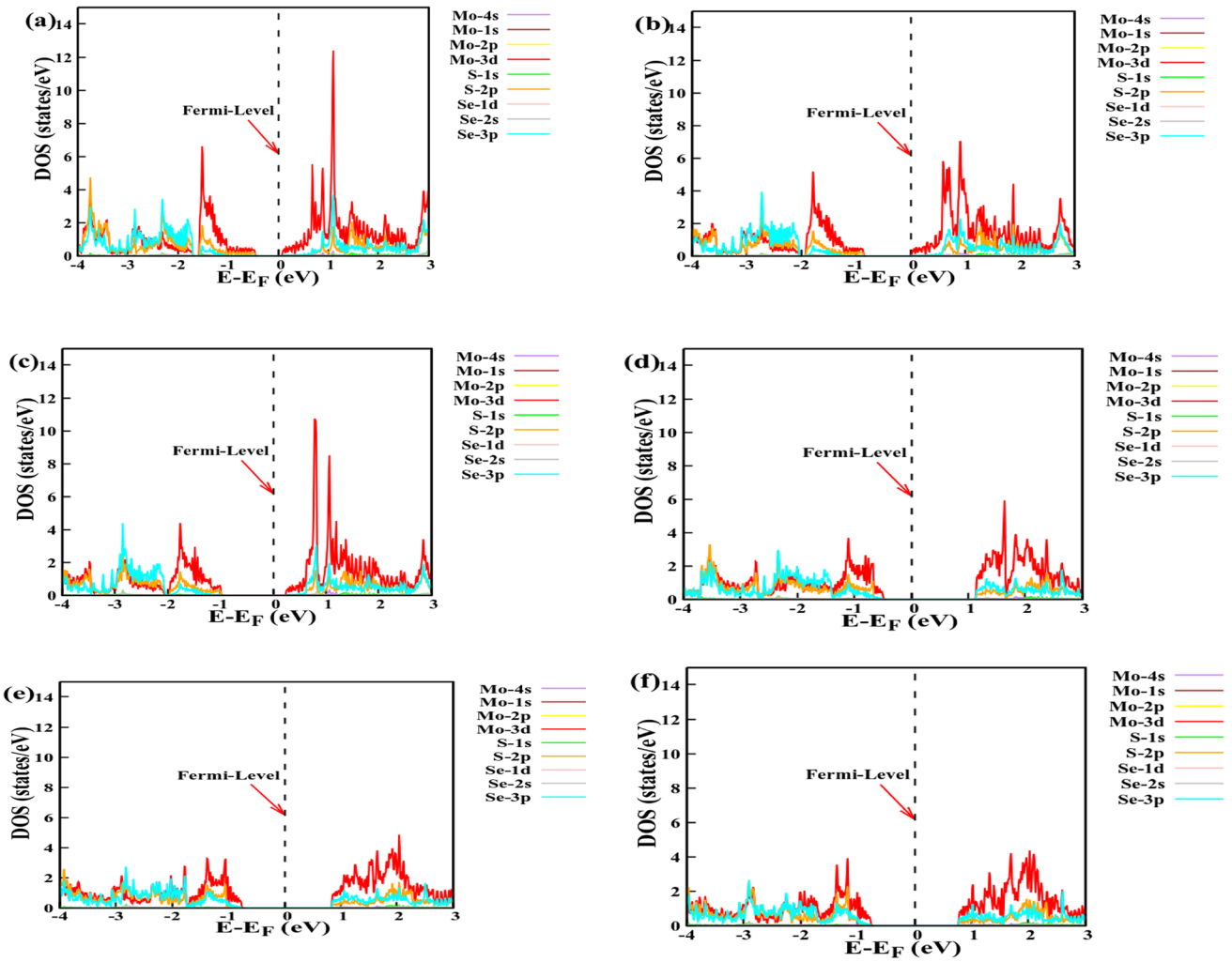


Fig. 3 The plot of projected density of states (pdos) at pressure, (**a**) −3, (**b**) −2, (**c**) -1, (**d**) 1, (**e**) 2 and (**f**) 3 GPa, respectively

The Mexican hat-like band structure at the high symmetric Γ-point in the E-K dispersion curve causes the Seebeck coefficient to increase [49, 50], which in turn yields improved thermoelectric efficiency. Pressure has been applied systematically from -3 GPa to + 3 GPa to observe the impact of pressure on the electronic band structure.

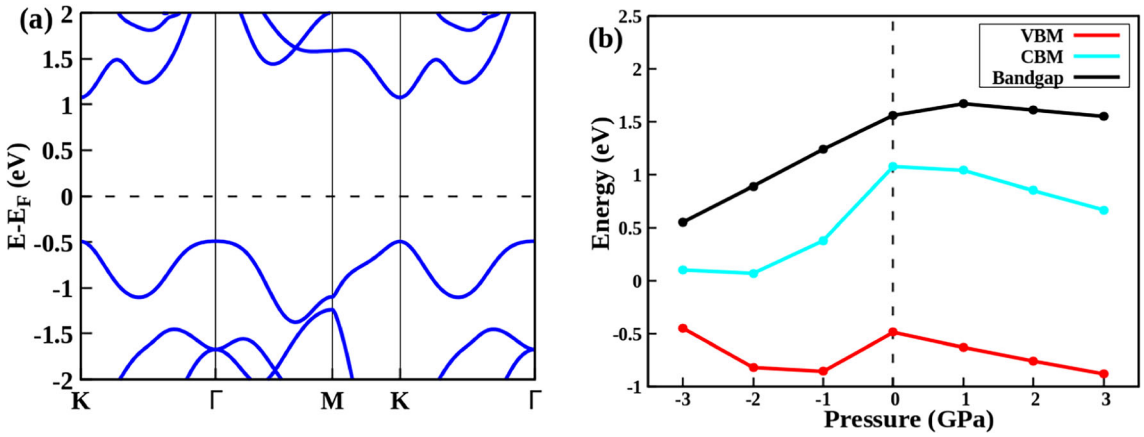


Fig. 4 **a** Electronic band structure of pure Janus monolayer and **b** Variation in bandgap with respect to external pressure

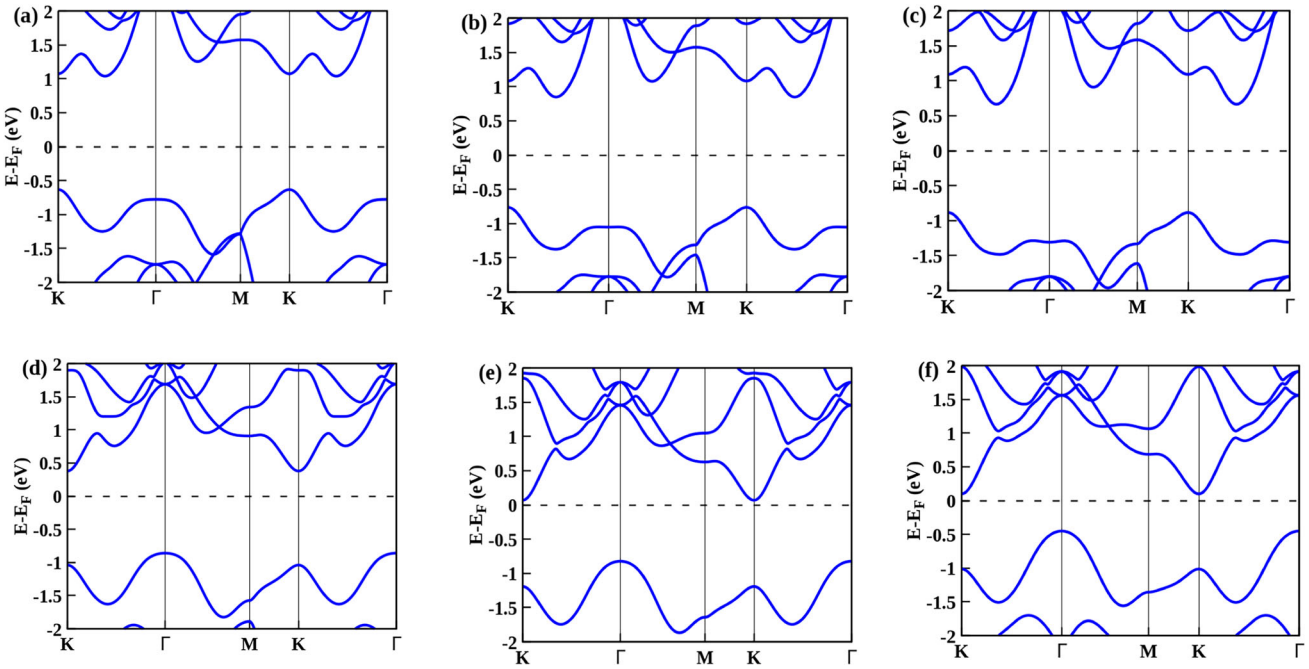


Fig. 5 Energy band structure at external pressure (**a**) + 1 GPa, (**b**) 2 GPa, (**c**) 3 GPa, (**d**) −1 GPa, (**e**) −2 GPa and (**f**) −3 GPa, respectively, of MoSSe Janus monolayer

After the confirmation of thermodynamical stability of Janus monolayer under different pressures, the effect of pressure has been studied on electronic properties. Tuning of band gap with respect to external pressure is shown in Fig. 4b, in which VBM starts to go away from Fermi level, whereas CBM starts to move toward the Fermi level. Also, the energy band dispersion curves become influenced to give indirect band gap with external pressure and no longer VBM & CBM both lies on high symmetric point K as shown in Fig. 5. The ability to manipulate MoSSe under pressure represents a significant advancement in tailoring two-dimensional materials for next-generation energy harvesting and cooling technologies.

3.3 Thermal and stability dynamics

As shown in Fig. 6a, the fully relaxed parameters are used to compute the phonon dispersion curve and high symmetric positions of the first Brillouin zone using the phonopy code to comprehend the dynamical stability.

These curves consist of nine modes, three of which are acoustic modes with lesser energy and the other six being optical modes. Each mode is made up of three atoms. The Janus monolayer is thermodynamically stable since all these modes come inside the positive frequency range. Further investigation of the material under pressure has led to the drawing of the dispersion curve under various pressures.

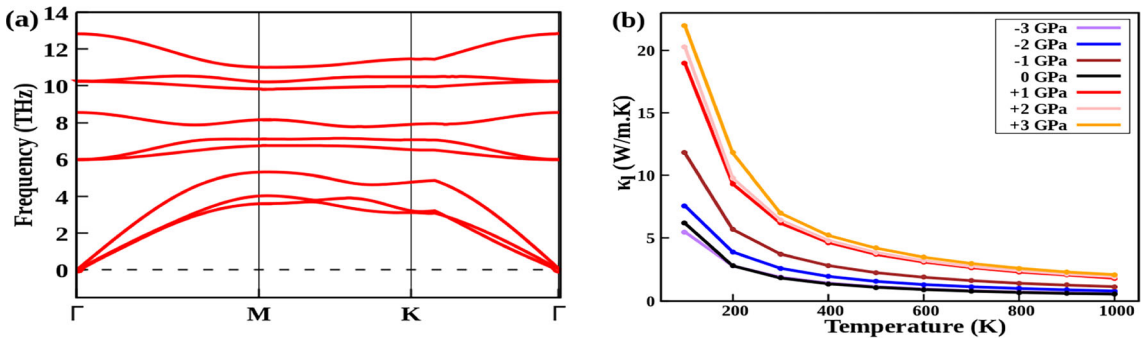


Fig. 6 **a** Phonon dispersion curve for Janus monolayer and **b** Variation of lattice thermal conductivity with respect to temperature, at different values of pressure

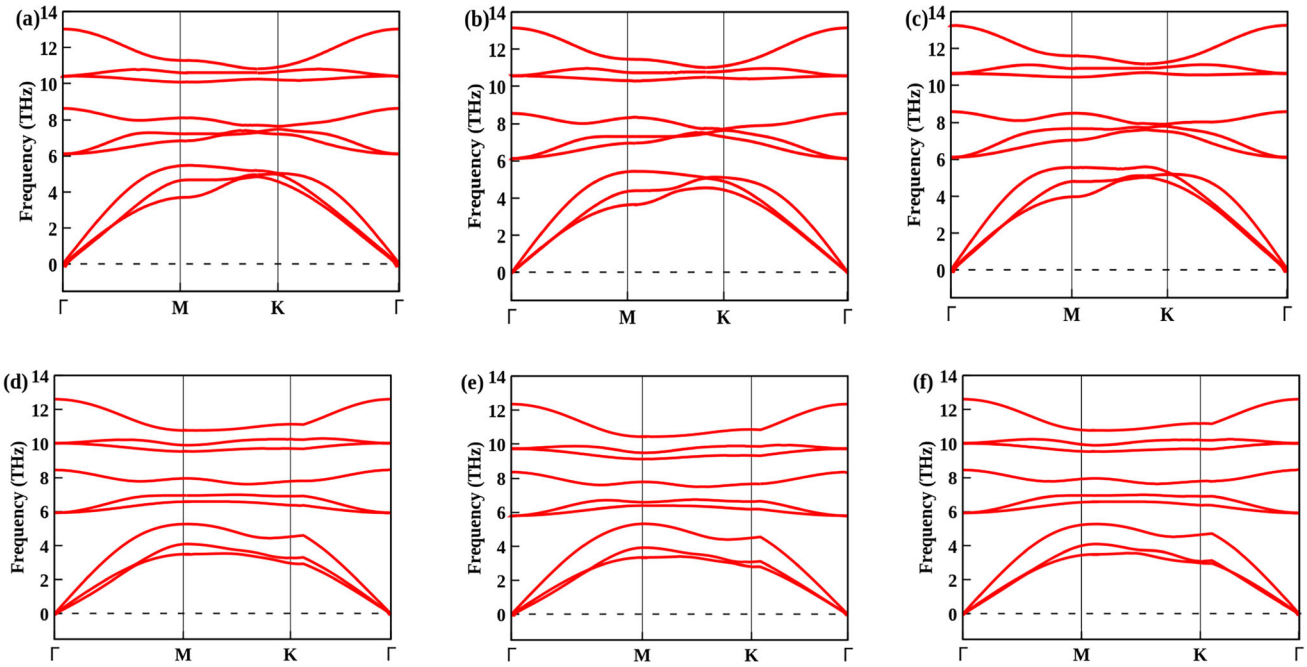


Fig. 7 Phonon dispersion curve correspondence to **a** + 1 GPa, **b** + 2 GPa, **c** + 3 GPa, **d** −1 GPa, **e** −2 GPa and **f** −3 GPa, respectively

Table 2 Lattice thermal conductivity (κ_l) in units of $Wm^{-1}\ K^{-1}$ at different pressures

Temp (K) / Pressure (GPa)	−3	−2	−1	0	+ 1	+ 2	+ 3
300	3.73	2.60	1.90	1.82	6.2	6.45	7.00
500	2.25	1.55	1.13	1.08	3.7	3.85	4.20

The Janus monolayer considered as thermodynamically stable under different pressures as shown in Fig. 7. Lattice vibrations can be used to predict heat conduction and estimates for lattice thermal conductivity have been made. The lattice thermal conductivity (κ_l) is calculated by using the linearized Boltzmann transport equations for phonons.

The lattice thermal conductivity is determined as a function of temperature, and it is decreasing because of phonon scattering dominance over phonon interactions. The plot is displayed in Fig. 6b, and the values are included in Table 2. It is exhibiting diminishing behaviour under various pressures. There is greater phonon scattering and a decreased lattice thermal conductivity because of the gap being smaller between the highest acoustical branch and optical branch. Hence, like other TMDCs Janus monolayers as WSSe [23, 51], WSTe [51] and HfSSe [25], MoSSe Janus monolayer exhibits very low lattice thermal conductivity.

3.4 Effective mass and relaxation time

Effective mass is essential for determining the thermoelectric parameters absolute value, which is determined by fitting quadratic curve to an electronic band structure. The second order curve fitting of the VBM and CBM of energy bands has yielded the effective

Table 3 Calculated various constants for relaxation time at temperatures 300 K and 500 K, respectively, for Janus monolayer at different pressures

Pressure (GPa)	Carriers	E_β (eV)	C_{11} (J/m^2)	m_d (m_e)	m^* (m_e)	τ (10^{-13} s)	
						300 K	500 K
−3	h	−1.76	91.01	0.27	0.41	13.12	7.87
	e	−9.81	91.01	0.32	0.13	0.35	0.21
−2	h	−1.36	106.04	0.38	0.52	18.23	10.94
	e	−10.60	106.04	0.26	0.14	0.43	0.26
−1	h	−1.05	120.88	0.41	0.69	32.32	19.39
	e	−11.29	120.88	0.21	0.14	0.56	0.34
0	h	−4.05	135.62	0.53	1.17	1.90	1.14
	e	12.40	135.62	0.20	0.15	0.53	0.32
+ 1	h	−6.39	149.31	0.82	0.20	0.53	0.32
	e	−6.48	149.31	0.18	0.19	2.38	1.43
+ 2	h	−7.42	164.85	0.42	0.19	0.85	0.51
	e	−4.60	164.85	0.18	0.16	5.12	3.07
+ 3	h	−6.86	176.00	0.38	0.20	1.19	0.71
	e	−4.24	176.00	0.20	0.14	6.07	3.64

mass for hole (h) and electron (e) carriers. Further, the relaxation time for hole and electrons is calculated by effective mass dependent deformation potential theory [24] and formulated as.

$$\tau = \frac{\hbar^3 C_{11}}{E_\beta^2 (m_d k_B T)} \tag{2}$$

The effective mass (m*) of a charge carrier hole (electron) as defined by formula; $m^* = \hbar^2/(\partial^2 E(k)/\partial k^2)$, was calculated by fitting a quadratic polynomial to a VBM (CBM) of a monolayer band structure. Here, $\partial^2 E(k)/\partial k^2$ denotes the second order derivative of the maxima valence (minima conduction) band energy curve with respect to wave vector (k). The average effective mass m_d has been expressed by $m_d = \sqrt{m^*_{\Gamma-M} m^*_{M-K}}$, where, $m^*_{\Gamma-M}$ and m^*_{M-K} are the effective mass along Γ-M and M–K direction. We further evaluated the charge carrier's relaxation time (τ) using deformation potential theory. Elastic constant C_{11} has been evaluated by quadratic polynomial fitting of total energy difference (E- E_0) versus applied strain, where E is strained system energy and E_0 is the energy of unstrained system.

Deformation potential constant is evaluated as $E_\beta = \partial E_{edge}/\partial\delta$, where E_{edge} is the energy of VBM (CBM) for hole (electron) and δ is the applied strain and is calculated by linear fitting of curve. By using the above calculated value of m_d, C_{11} and E_β in equation [34]. By using all calculated coefficients, the relaxation time (τ) has been calculated at different temperatures as tabulated in Table 3. The absolute value of the thermoelectric parameters can be found by using the relaxation time.

4 Transport properties

4.1 Janus monolayer at 0 GPa pressure

After the investigation of structural and electronic properties, a deep understanding has been gained for the estimation of the thermoelectric properties of the Janus monolayer based on the semi-Boltzmann transport equation within the rigid band approximation.

Seebeck coefficient (S) is the measurement of bipolar conduction of charge carriers and its variation with respect to carrier concentration (n) is plotted in Fig. 8a. Negative and positive value of n refers to carriers as electrons and holes, respectively. In the same manners, negative and positive values of S represent for the n- and p-type doping. Its value obtained for n- and p-type doping is 2.52 and 2.56 mV/K, at room temperature, suggesting that the material is behaving as p-type semiconductor. Seebeck coefficient drops with temperature because of increase of bipolar effect at higher temperatures for both *p*- and *n*-type carriers. Also, its value decreases with increase in carrier concentration as followed by expression $S \propto (1/n)^{2/3}$ [26] and remains constant with the increase in carrier concentration.

The variation of electrical conductivity (σ) with respect to carrier concentration is illustrated in Fig. 8b. Intrinsic scattering of charge carriers becomes more strengthened with increase in temperature results to give the decreasing behaviour with temperature. Its value increases with increase in carrier concentration because of its direct dependency on carrier concentration (n) as $\sigma \propto n$ [52]. As mentioned in Table 3, effective mass for holes is greater than electrons cause to give higher Seebeck coefficient for p-type carriers as compared to n-type carriers and reflect to higher electrical conductivity for p-type doping [35, 53, 54].

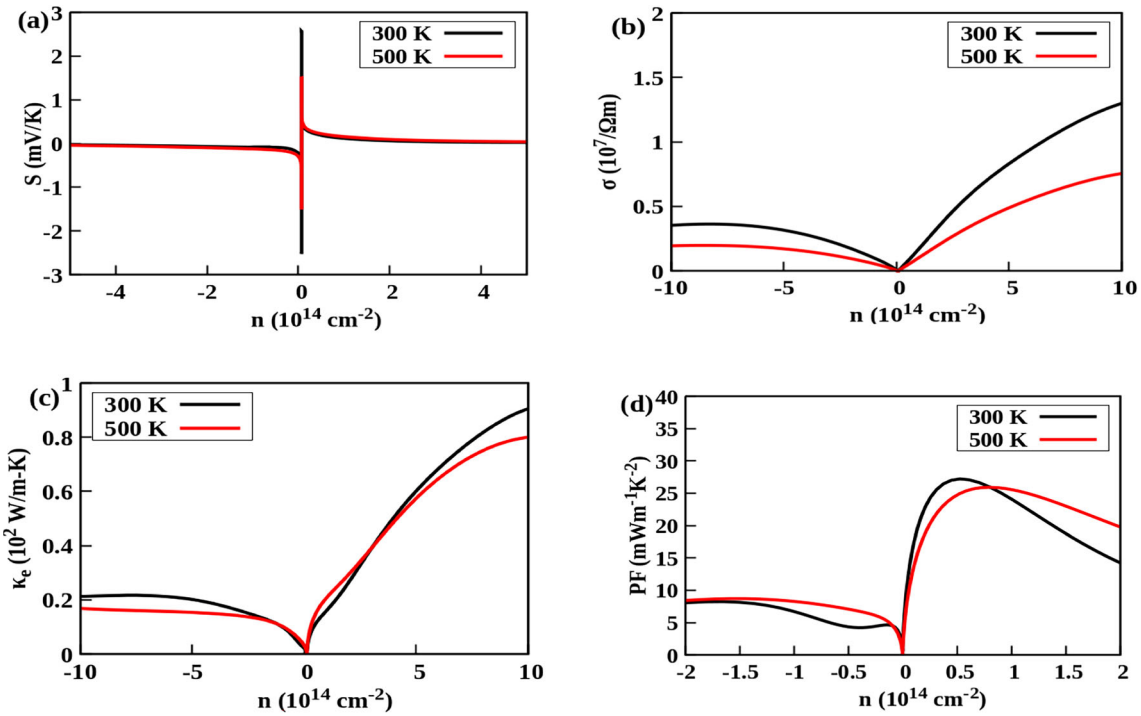


Fig. 8 Variation of **a** Seebeck coefficient, **b** Electrical conductivity (σ), **c** Electronic thermal conductivity and **d** Power factor, with respect to carrier concentration (n)

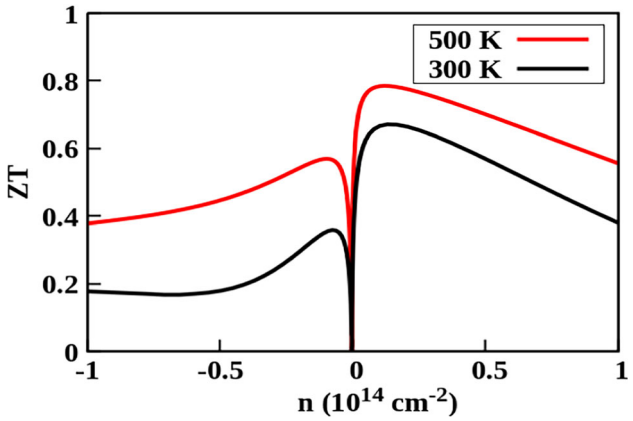


Fig. 9 Variation of figure of merit (ZT) of pure Janus monolayer with respect to carrier concentration (n), at different temperatures

The electronic thermal conductivity (κ_e) is showing similar behaviour as electrical conductivity as related by Wiedemann–Franz law [55], expressed as $\kappa_e = L\sigma T$, where for non-degenerate semiconductors, Lorentz number (L) $= 1.5 \times 10^{-8}$ $W\Omega K^{-2}$ [56]. Variation of κ_e with respect to carrier concentration and temperature is shown in Fig. 8c.

Power factor is another key related by the combined effect of Seebeck coefficient and electrical conductivity as power factor $= \sigma S^2$, and the variation with carrier concentration is shown in Fig. 8d. It has been noticed that power factor is increasing with carrier concentration and observed a significant influence of increasing temperature. Plot analysis reveals that the power factor is relatively small at low doping due to the dominance of electrical conductivity over Seebeck coefficient. It grows to its higher value with an increase in carrier concentration (n) and then begins to display a falling trend. Peak of p-type and n-type doping is suggested that p-type doping is more favourable over n-type doping, providing higher power factor for p-type doping.

The figure of merit (ZT) was calculated using the previously determined parameters and its plot with respect to carrier concentration (n) is presented in Fig. 9. The figure illustrates that the peak value of ZT is 0.67 for p-type doping which is higher than 0.59 for n-type doping. Also, it is increasing with increase in temperature.

All the above parameters are also showing that under 0 GPa pressure, material is giving enhanced thermoelectric properties for p-type doping as compared to n-type doping and the value is higher or comparable with previously reported Janus monolayers like MoSeTe [24], WSSe [23], HfSSe [25] and WSTe [25].

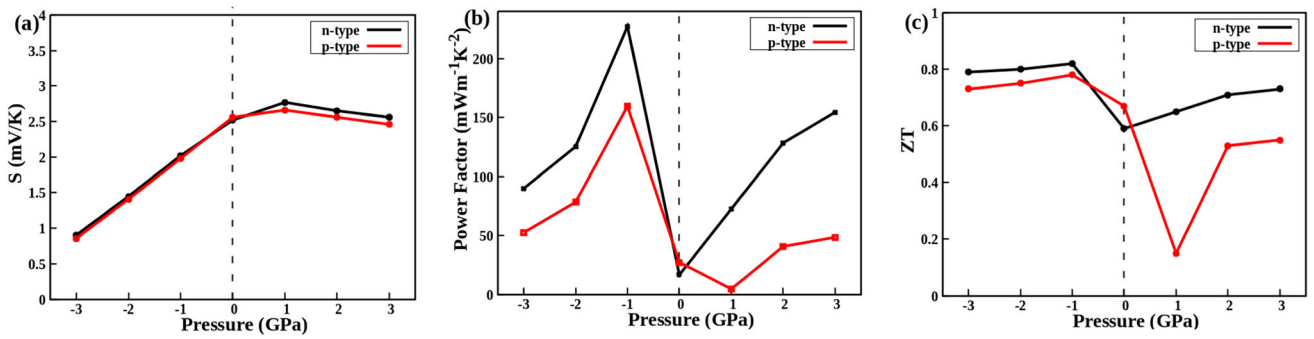


Fig. 10 Variation of Seebeck coefficient (S), power factor (PF) and figure of merit (ZT), with respect to applied pressure for Janus monolayer, at room temperature

Table 4 At room temperature, Seebeck coefficient (S), power factor (PF) and ZT for *n*-type and *p*-type carriers under different pressures

Pressure (GPa)	Carriers	S (mV/K)	PF (Wm^{-1} K^{-2})	ZT
−3	N	0.90	89.88	0.79
	P	0.85	52.55	0.73
−2	N	1.45	125.15	0.80
	P	1.41	78.40	0.75
−1	N	2.02	227.15	0.82
	P	1.98	159.50	0.78
0	N	2.52	16.59	0.59
	P	2.56	27.21	0.67
+ 1	N	2.77	72.50	0.65
	P	2.66	4.55	0.15
+ 2	N	2.65	128.20	0.71
	P	2.56	40.66	0.53
+ 3	N	2.56	154.40	0.73
	P	2.46	48.50	0.55

4.2 Janus monolayer MoSSe under pressure

The impact of external pressure on $SnSe_2$ monolayers and bilayers [35] is a source of inspiration for the development of Janus monolayers' thermoelectric properties. First, it is discovered that the VBM and CBM have been moved against the Fermi level and that external pressure has a significant impact on the electronic structure. As illustrated in Fig. 4(b), the band gap exhibits decreasing behaviour with respect to negative external pressure but first increases with pressure when exposed to positive external pressure. The transport properties are tuned, exhibiting an enhanced behaviour. As explained by Goldsmid-Sharp relation [54], $E_g = 2e|S_m|T_m$, here E_g and T_m, S stands for bandgap, temperature and the highest Seebeck coefficient, respectively; therefore, the band gap and Seebeck coefficient are manifesting similar behaviour. This causes the Seebeck coefficient for n-type doping to grow from 2520 to 2770 μVK^{-1} under positive pressure; however, it constantly falls under negative pressure, as shown in Fig. 10a.

P-type charge carriers have demonstrated a similar behaviour at room temperature. As previously explained and shown in Table 3, the effective mass of holes decreases with external pressure relative to the effective mass of electrons, which causes the Seebeck coefficient for p-type doping to decrease from n-type doping. At 0 GPa, however, the Seebeck coefficient for p-type doping was higher than that of n-type doping. It follows that for n-type doping, external pressure is more effective in enhancing the Seebeck coefficient. This causes the material to transition from being p-type doped to being n-type doped when external pressure increases.

The power factor exhibits the same trends as the Seebeck coefficient in the transition from p-type to n-type doping because it is the result of the combined action of electrical conductivity and the Seebeck coefficient. However, due to the combination of two components, where electrical conductivity predominates over Seebeck coefficient as shown in Fig. 10b, the power factor value has a different trend from the Seebeck coefficient with respect to pressure. The maximum value of power factor is 227.15 mWm^{-1} K^{-2} obtained at −1 GPa pressure for n-type doping, and its value is increased up to 154.40 mWm^{-1} K^{-2} with positive pressure for n-type doping which is much larger from pure material as given in Table 4.

Figure 10c illustrates the evaluation of the thermoelectric figure of merit (ZT), which considers the combined effect of all thermoelectric parameters and relaxation time. ZT and external pressure follow similar trends, as ZT is directly dependent on the

power factor. The maximum ZT value of 0.82 for n-type doping is reached at a pressure of −1 GPa, which is 39% higher than that of pure material.

It is proposed that hydrostatic pressure conditions can be useful to improve the thermoelectric performance of the material in a practical application, as $BiTe_2$ -based materials have found highly challenged features [57]. Two *n*- and *p*-type legs of thermoelectric material with good thermoelectric performance make up a thermoelectric module.

5 Conclusions

The theoretical investigation of structural, electronic and thermoelectric properties MoSSe Janus monolayer is systematically studied without and with pressure. Under different pressure, the bandgap of the material can be tuned and converts the material from direct to indirect semiconductor. Absence of negative frequency in phonon dispersion modes indicates the thermodynamical stability of pure monolayer and monolayer under applied pressure. Seebeck coefficient for n- and p-type doping is 2.52 and 2.56 mV/K, at room temperature, suggesting that the material behaves as a p-type semiconductor. ZT is 0.67 for p-type doping which is higher than 0.59 for n-type doping without any external pressure. Seebeck coefficient for n-type doping is firstly increased from 2520 to 2770 μVK^{-1} with positive pressure and after that decreases whereas, it is continuously decreasing with negative pressure. With the external pressure, the value of ZT is enhanced from 0.67 to 0.78 at −1 GPa for p-type carriers, whereas, for n-type charge carriers, the value of ZT increases up to 0.73 at + 3 GPa and up to 0.82 at −1GPa from 0.59, at room temperature.

Therefore, the value of ZT for n- and p-type material can be enhanced under positive and negative value of external pressure, respectively. Hence, based upon the above results, the MoSSe Janus monolayer plays the important role in thermoelectric power generators, sensors and performance wearables.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1140/epjp/s13360-025-06180-9.

Acknowledgements All the authors are highly thankful to the National PARAM-Yuva-II supercomputer at CDAC, Pune, India for supporting large scale calculations. All the authors also would like to thank Pt. Deendayal Upadhyaya Innovation and Incubation Centre (PDUIIC), Guru Jambheshwar University of Science and Technology, Hisar, India, for providing the funds under RUSA 2.0 Grant for setting the computational facility at Department of Physics, GJUS&T, Hisar.

Data Availability Statement The manuscript has associated data in a data repository. Data is made available upon reasonable request to the corresponding author.

Declarations

Conflict of interest All the authors declared that they do not have any conflict of interest.

References

1. L. Yang, Z.G. Chen, M.S. Dargusch, J. Zou, High performance thermoelectric materials: progress and their applications. Adv. Energy Mater. **8**, 1701797 (2018)
2. P.A. Owusu, S. Asumadu-Sarkodie, A review of renewable energy sources. Sustain. Issues Clim. Change Mitigat. Cogent Eng. **3**, 1167990 (2016)
3. D. Tainoff, A. Proudhom, C. Tur, T. Crozes, S. Dufresnes, S. Dumont, D. Bourgault, O. Bourgeois, Network of thermoelectric nanogenerators for low power energy harvesting. Nano Energy **57**, 804 (2019)
4. M. Haras, T. Skotnicki, Thermoelectricity for IoT- a review. Nano Energy **54**, 461 (2018)
5. C. Hou, S. Zhang, R. Liu, T. Gemming, A. Bachmatiuk, H. Zhao, H. Jia, S. Huang, W. Zhou, J.-B. Xu, J. Pang, Boosting flexible electronics with integration of two-dimensional materials. InfoMat **6**, e12555 (2024)
6. J. Zhao, F. Li, D. Ling, S. Liu, F. Li, D. Ling, D. Ling, Z. Chen, P. Li, F. Li, F. Li, S. Yu, Nano-bio interactions between 2D nanomaterials and mononuclear phagocyte system cells. BMEMat **2**, 12066 (2024)
7. R. Kim, S. Datta, M.S. Lundstrom, Influence of dimensionality on thermoelectric device performance. J. Appl. Phys. **105**, 034506 (2009)
8. Anisha, R. Singh, S.S. Kumar, K. Tankeshwar, Tuning of thermoelectric performance of $CrSe_2$ material using dimension engineering. J. Phys. Chem. Solids **172**, 111083 (2023)
9. T. Mueller, E. Malic, Egrxciton physics and device application of two-dimensional transition metal dichalcogenide semiconductors, 2D Mater. Appl. **2**, 29 (2018)
10. C. Chowdhury, S. Karmakar, A. Datta, Monolayer group IV–VI monochalcogenides: low-dimensional materials for photocatalytic water splitting. J. Phys. Chem. C **121**, 7615 (2017)
11. C. Kamal, A. Chakrabarti, M. Ezawa, Direct band gaps in group IV-VI monolayer materials: binary counterparts of phosphorene. Phys. Rev. B **93**, 125428 (2016)
12. D. Wickramaratne, F. Zahid, R.K. Lake, Electronic and thermoelectric properties of few-layer transition metal dichalcogenides. J. Chem. Phys. **140**, 124710 (2014)
13. P. Yan, G.Y. Gao, G.Q. Ding, D. Qin, Bilayer MSe_2 (M = Zr, Hf) as promising two-dimensional thermoelectric materials: a first- principles study. RSC Adv. **9**, 12394 (2019)
14. S. Yuan, H. Ding, J. Wang, Z. Chen, Extraordinary magnetoresistance in janus monolayer $MoTeB_2$: a theoretical prediction. J. Phys. Chem. C **122**, 28423 (2018)

15. K. Hippalgaonkar, Y. Wang, Y. Ye, D.Y. Qiu, H. Zhu, Y. Wang, J. Moore, S.G. Louie, X. Zhang, High thermoelectric power factor in two-dimensional crystals of MoS_2. Phys. Rev. B **95**, 1 (2017)
16. Z. Fan, Z.W. Bing, Electronic structures and elastic properties of Se, Te): a comparative first-principles study. Chin. Phys. B **24**, 097103 (2015)
17. Anisha, R. Singh, S.S. Kumar, K. Tankeshwar, Thermoelectric response of 1T-ZrS_2 monolayer: ab-initio study. Mater. Today Proc. **03**, 2214 (2023)
18. S.-D. Guo, X.-S. Guo, R.-Y. Han, Y. Deng, Predicted janus SnSSe monolayer: a comprehensive first-principles study. Phys. Chem. Chem. Phys. **21**, 24620 (2019)
19. T.V. Vu, H.D. Tong, D.P. Tran, N.T.T. Binh, C.V. Nguyen, H.V. Phuc, H.M. Do, N.N. Hieu, Electronic and optical properties of Janus ZrSSe by density functional theory. RSC Adv. **9**, 41058 (2019)
20. W.-L. Tao, Y. Mu, C.-E. Hu, Y. Cheng, G.-F. Ji, Electronic structure, optical properties, and phonon transport in Janus monolayer PtSSe via first-principles study. Phil. Mag. **99**, 1025 (2019)
21. S. Deng, L. Li, O.J. Guy, Y. Zhang, Enhanced thermoelectric performance of monolayer MoSSe, bilayer MoSSe and graphene/MoSSe heterogeneous nanoribbons. Phys. Chem. Chem. Phys. **21**, 18161 (2019)
22. D.M. Hoat, M. Naseri, N.N. Hieu, R. Ponce-Pérez, J.F. Rivas-Silva, T.V. Vu, G.H. Cocoletzi, A comprehensive investigation on electronic structure, optical and thermoelectric properties of the HfSSe Janus monolayer. J. Phys. Chem. Solids **144**, 109490 (2020)
23. R. Chaurasiya, S. Tyagi, N. Singh, S. Auluck, A. Dixit, Enhancing thermoelectric properties of Janus WSSe monolayer by inducing strain mediated valley degeneracy. J. Alloys Compd. **855**, 157304 (2021)
24. S. Saini, A. Shrivastava, A. Dixit, S. Singh, Ultra-low lattice thermal conductivity and high figure of merit for Janus MoSeTe monolayer: a peerless material for high temperature regime thermoelectric devices. J. Mater. Sci. **57**, 7012 (2022)
25. A. Patel, D. Singh, Y. Sonvane, P.B. Thakor, R. Ahuja, High Thermoelectric Performance in Two-Dimensional Janus Monolayer Material WS-X (X = Se and Te). ACS Appl. Mater. Interfaces **12**, 46212 (2020)
26. S.-D. Guo, Phonon transport in Janus monolayer MoSSe: a first-principles study. Phys. Chem. Chem. Phys. **20**, 7236 (2018)
27. J. Zhang, S. Jia, I. Kholmanov, L. Dong, D. Er, W. Chen, H. Guo, Z. Jin, V.B. Shenoy, L. Shi, J. Lou, Janus monolayer transition-metal dichalcogenides. ACS Nano **11**, 8192 (2017)
28. S.D. Guo, Predicted Janus monolayer ZrSSe with enhanced n-type thermoelectric properties compared with monolayer ZrS_2. Comput. Mater. Sci. **161**, 12 (2017)
29. L. Hu, B. Yang, Z. Hou, Y. Lu, W. Su, L. Li, Unlocking the charge doping effect in softly intercalated ultrathin ferromagnetic superlattice. eScience **3**, 100117 (2023)
30. G. Kim, L. Shao, K. Zhang, K.P. Pipe, Engineered doping of organic semiconductors for enhanced thermoelectric efficiency. Nat. Mater. **12**, 1 (2013)
31. R. Anisha, S.S. Kumar, K. Tankeshwar, Optimising 1T-NiS_2 monolayer thermoelectric performance via valley engineering. Mater. Today Commun. **34**, 105169 (2023)
32. R. Anisha, S. Srivastava, Band gap engineering of 2H-MX_2 (M = Mo; X = S, Se, Te) monolayers using strain effect. Mater. Today Proc. **54**, 677–681 (2022)
33. R. Anisha, S.S. Kumar, K. Tankeshwar, Thermoelectric performance of 1T-ZrS_2 bilayer using stacking engineering. Phys. Scr. **99**, 015914 (2024)
34. R. Anisha, S.S. Kumar, K. Tankeshwar, Thermoelectric properties of PtX2 (X = Se, Te) monolayers. Indian J. Phys. **97**, 3913 (2023). https://doi.org/10.1007/s12648-023-02727-7
35. Z. Daifeng, Y. Chuanbin, L. Yuhao, O. Yun, G. Yongyi, Pressure-induced enhancement in the thermoelectric properties of monolayer and bilayer $SnSe_2$. R Soc Open Sci. **28**, 171827 (2018)
36. S. Nag, A. Saini, R. Singh, R. Kumar, Influence of vacancy defects on the thermoelectric performance of SnSe sheet. Phys. E Low-Dimens. Syst. Nanostruct. **134**, 114814 (2021). https://doi.org/10.1016/j.physe.2021.114814
37. S.M.S. Shahriar, N.S. Polavoram, S.M. Andrabi, Y. Su, D. Lee, H.Q. Tran, S.J. Schindler, J. Xie, Transforming layered 2D mats into multiphasic 3D nanofiber scaffolds with tailored gradient features for tissue regeneration. BMEmat. **2**, e12065 (2024)
38. H.A.R. Aliabad, F.A. Rad, Structural, electronic and thermoelectric properties of bulk and monolayer of Sb_2Se_3 under high pressure: by GGA and mBJ approaches. Physica B Condens. Matter. **545**, 275 (2018)
39. Y. Zhang, S. Hao, L.D. Zhao, C. Wolverton, Z. Zeng, Pressure induced thermoelectric enhancement in SnSe crystals. J. Mater. Chem. A **4**, 12073 (2016)
40. Y. Javed, M. Rafiq, N. Ahmed, Pressure-induced changes in the electronic structure and enhancement of the thermoelectric performance of SnS_2: a first principles study. RSC Adv. **7**, 38834 (2017)
41. B. Hou, T. Zhang, T. Wang, H. Ji, H. Yang, L. Jia, X. Han, J. Qiao, Y. Zhang, L. Liu, H.-J. Gao, Y. Wang, Atomic tracking of thermally-driven structural evolution in 2D crystals: case of $NbSe_2$. InfoMat **6**, 12501 (2023)
42. P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M.B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A.D. Corso, S. De Gironcoli, P. Delugas, A. DiStasio, A. Ferretti, G. Floris, G. Fratesi, R. Fugallo, U. Gebauer, F. Gerstmann, T. Giustino, J. Gorni, M. Jia, H.Y. Kawamura, A. Ko, E. Kokalj, M. Küçükbenli, M. Lazzeri, N. Marsili, F. Marzari, N.L. Mauri, H.V. Nguyen, A. Nguyen, L. Otero-de-la-Roza, S. Paulatto, D. Poncé, R. Rocca, B. Sabatini, M. Santra, A.P. Schlipf, A. Seitsonen, I. Smogunov, T. Timrov, P. Thonhauser, N. Umari, XWu. Vast, S. Baroni, Advanced capabilities for materials modelling with Quantum ESPRESSO. J. Phys. Condens. Matter **29**, 465901 (2017)
43. J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple. Phys. Rev. Lett. **77**, 3865 (1996)
44. A. Togo, I. Tanaka, First principles phonon calculations in materials science. Scr. Mater. **108**, 1 (2015)
45. G.K. Madsen, D.J. Singh, Boltztrap a code for calculating band-structure dependent quantities. Comput. Phys. Commun. **175**, 67 (2006)
46. J. Bardeen, W. Shockley, Deformation potentials and mobilities in non- polar crystals. Phys. Rev. **80**, 72 (1950)
47. A. Togo, L. Chaput, I. Tanaka, Distributions of phonon lifetimes in Brillouin zones. Phys. Rev. B **91**, 094306 (2015)
48. A.Y. Lu, H. Zhu, J. Xiao, C.P. Chuu, Y. Han, M.H. Chiu, C.C. Cheng, C.W. Yang, K.H. Wei et al., Janus monolayers of transition metal dichalcogenides. Nat. Nanotechnol. **12**, 744 (2017)
49. M. Nurhuda, A. Nugraha, M. Hanna, E. Suprayoga, E. Hasdeo, Thermoelectric properties of Mexican-hat band structures. Adv Nat Sci. Nanosci. Nanotechnol. **11**, 015012 (2020)
50. J. Bera, A. Betal and S. Sahu, (2020) Ultralow lattice thermal conductivity and high thermoelectric performance near room temperature of Janus monolayer HfSSe, Mater. Sci.
51. G.S. Jeffrey, E.S. Toberer, Complex thermoelectric materials. Nat. Mater. **7**, 101 (2008)
52. T.M. Tritt, Thermoelectric phenomena, materials, and applications. Annual Rev. Mater. Res. **41**, 433 (2011)
53. C. Xia, Y. Peng, H. Zhang, T. Wang, S. Wei, Y. Jia, The characteristics of n- and p-type dopants in SnS_2 monolayer nanosheets. Phys. Chem. Chem. Phys. **16**, 19674 (2014)
54. B.L. Evans, R.A. Hazelwood, Optical and electrical properties of $SnSe_2$. J. Phys. D Appl. Phys. **2**, 1507 (1969)
55. K. Kuroki, R. Arita, Pudding mold band drives large thermopower in Na_xCoO_2. J. Phys. Soc. Jpn. **76**, 083707 (2007)
56. H.J. Goldsmid, J.W. Sharp, Estimation of the thermal band gap of a semiconductor from seebeck measurements. J. Electron. Mater. **28**, 869 (1999)

57. T. Cao, X.-L. Shi, M. Li, B. Hu, W. Chen, W.-D. Liu, W. Lyu, J. MacLeod, Z.-G. Chen, Advances in bismuth-telluride-based thermoelectric devices: Progress and challenges. eScience **3**, 100122 (2023)

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Physica Scripta

INCLUSIVE PUBLISHING TRUSTED SCIENCE

PURPOSE-LED PUBLISHING™

PAPER

Evaluating the intensity of Typhoons over the West Pacific basin

To cite this article: Monu Yadav and Laxminarayan Das 2025 *Phys. Scr.* **100** 056003

View the article online for updates and enhancements.

You may also like

- Typhoon enhancement of N and P release from litter and changes in the litter N:P ratio in a subtropical tidal wetland
Weiqi Wang, Jordi Sardans, Chuan Tong et al.
- The influence of radiation flux in Northwest Pacific on the Western Pacific warm pools and typhoons over the past 170 years
Chai Bo-yu, Xu Feng, Xu Jian-jun et al.
- Canopy height damage by Typhoon Songda in Northern Hokkaido, Japan, in 2004
TaeOh Kwon, Hideaki Shibata, Tetsuya Takemi et al.

This content was downloaded from IP address 14.139.251.98 on 08/04/2025 at 06:39

Physica Scripta

PAPER

Evaluating the intensity of Typhoons over the West Pacific basin

Monu Yadav and Laxminarayan Das
Department of Applied Mathematics, Delhi Technological University, New Delhi, India
E-mail: yadavm012@gmail.com

Keywords: Typhoon, intensity, machine learning, west pacific ocean basin

RECEIVED
19 October 2024

REVISED
3 March 2025

ACCEPTED FOR PUBLICATION
21 March 2025

PUBLISHED
2 April 2025

Abstract

A Typhoon is a severe weather system that can cause significant damage to human property. Accurate typhoon intensity estimation is vital for timely risk warning and hazard assessment. In this study, we proposed a novel neural network-based model to estimate the intensity of typhoons more accurately than the traditional methods. The proposed model takes inputs from three traditional methods: the automated rotational center hurricane eye retrieval algorithm, the advanced Dvorak technique, and the satellite Consensus technique. The model is trained using 147 typhoon cases, including tropical storms that formed over the West Pacific basin during 2017–24, and validated using 15 cases. The results show that the error percentage of the proposed method concerning Best track data provided by the Regional Specialized Meteorological Center (RSMC) for the West Pacific Basin is 6.806%, and the root mean square error is 5.42, which is lower than other methods. R-squared also indicates that our proposed methods show better predictive capability than other methods. As a result, there is a good chance that the proposed method for storm intensity estimation will provide a more accurate estimation than the other traditional methods.

1. Introduction

Today, Typhoons are becoming increasingly frequent worldwide due to global warming [1]. Characterized by strong winds and rainstorms, Typhoons have devastating impacts on human activity. Predicting Typhoons efficiently is challenging due to their complex physical processes. Numerous studies have been conducted on the axisymmetric structures and dynamic mechanisms of Typhoons [2, 3]. In recent years, machine learning and deep learning have significantly advanced Typhoon research, including forecasting, tracking, intensity estimation, classification, and disaster impact prediction. Currently, these fields are investigated separately, limiting the potential of data analysis tools to enhance our understanding of Typhoons. Future research should focus on incorporating more prior knowledge and gathering extensive multi-source data to improve the accuracy of Typhoon modeling.

In the field of Typhoon modeling, the research on Typhoon classification precedes the estimation of Typhoon intensity. Extensive research on Typhoon classification utilizing machine learning and deep learning has provided substantial insights that enhance Typhoon intensity prediction.

Traditional methods necessitate extensive human intervention and the manual engineering of features to estimate typhoon intensity. Intensity is correlated with other attributes of typhoons, such as the radius of maximum wind [4, 5]. Over the past three decades, the Dvorak Technique (DT) has been the most widely employed satellite image-based pattern recognition method for estimating Typhoon intensity [6]. The DT establishes a relationship between organized cloud patterns and Typhoon intensity, saving tens of thousands of lives. However, it is a subjective method, with its accuracy heavily influenced by the experience of the experts applying it. To address these limitations, several revisions have been developed, including the Objective Dvorak Technique (ODT) [7] and the Advanced Dvorak Technique (ADT) [8]. All versions of the DT technique rely on cloud features and pattern recognition to estimate Typhoon intensity [8]. Despite this, the handcrafted features inherent to these methods are often subjectively biased and may not fully capture the relationship between cloud patterns and Typhoon intensity. In 2010, [9] proposed the Automated Rotational Center Hurricane Eye

© 2025 IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

Retrieval (ARCHER) algorithm, which determines the center of rotation by analyzing spirally oriented brightness temperature gradients in Typhoon banding patterns and the gradients along the ring-shaped edge of a potential eye. The Advanced Microwave Sounding Unit (AMSU) technique, described by [10], employs temperature anomaly profiles associated with the Typhoon warm core to estimate intensity. The Satellite Consensus (SATCON) technique, as discussed by [11–13], combines estimates from various methods, including the Advanced Dvorak Technique (ADT) and AMSU, to provide state-of-the-art intensity estimations. Those traditional techniques require systematic meteorological knowledge and cannot easily adapt to diversified data.

In recent years, the application of convolutional neural networks (CNNs) to various image processing tasks has emerged as a promising approach for improving the accuracy of Typhoon intensity estimation using neural network (NN)-based models.

Goodfellow *et al* (2016) [14] is a foundational textbook that comprehensively covers deep learning concepts, including neural networks, optimization, and generative models. It serves as a key resource for researchers and practitioners in machine learning and artificial intelligence. Pradhan *et al* (2017) [15] propose a deep CNN using satellite image classification for Typhoon intensity estimation using satellite images. This model achieves improved accuracy and lower root mean square error (RMSE) compared to traditional methods like the Dvorak and deviation-angle variance techniques. Chen *et al* (2018) [16] introduce a novel dataset for Typhoon intensity estimation and propose a CNN-based model incorporating domain knowledge. Combinido *et al* (2018) [17] demonstrate that a CNN-based approach can effectively estimate Typhoon intensity from infrared satellite images, achieving an RMSE of 13.23 knots (kts). The model learns critical features like cloud organization and eye formation, comparable to traditional feature-based techniques, highlighting the potential of deep learning in meteorology. Lee *et al* (2019) [18] demonstrate that multi-dimensional CNNs using multi-spectral satellite data improve intensity estimation. The optimized 2D-CNN model achieved an RMSE of 8.32 kts, outperforming previous methods and confirming the potential of deep learning in meteorological applications. Chen *et al* (2019) [19] present a deep-learning model using CNNs for intensity estimation from satellite imagery. The optimized and smoothed CNN model demonstrates superior accuracy compared to traditional methods, achieving an RMSE of 8.39 kts and showing promise for future operational applications. Higa *et al* (2021) [20] propose a deep learning model for Typhoon intensity classification by integrating meteorological domain knowledge. By utilizing VGG-16 with fisheye-preprocessed satellite images, the model achieved a higher classification accuracy of 76.8%, outperforming previous methods. The fisheye preprocessing enhanced key Typhoon features like the eye and cloud distribution, improving interpretability. Chen *et al* (2020) [21] introduce a novel tensor-based (TCNN) for intensity estimation using multispectral satellite images. By integrating tensor decomposition and deep learning, TCNN effectively captures typhoon structures, improving intensity categorization and wind speed estimation. The proposed models, C-TCNN and T-TCNN, outperform traditional and state-of-the-art methods in accuracy and efficiency. Experimental results confirm TCNN's potential for operational meteorological applications. Zhang *et al* (2021) [22] propose a novel deep learning approach, TCICENet, for intensity classification and estimation using infrared satellite images. The model consists of cascading CNN modules, achieving an RMSE of 8.60 kt and an MAE of 6.67 kt. Varalakshmi *et al* (2023) [23] proposed a deep learning model that combines CNN architectures for feature extraction and machine learning models for regression to predict the intensity of TCs. Griffin *et al* (2024) [24] present two machine-learning methods, D-MINT and D-PRINT, utilizing infrared and microwave imagery to predict short-term intensity changes in typhoons, demonstrating improved skill over traditional methods, particularly in the North Atlantic and western North Pacific regions. Raynaud *et al* (2024) [25] introduce a CNN based on the U-Net architecture for detecting Typhoon wind structures in high-resolution numerical weather prediction (NWP) models. The model outperforms traditional heuristic methods, achieving an average intersection-over-union metric of 0.8, particularly improving detection for weaker cyclones. Additionally, it generalizes well to different domains and higher-resolution data, demonstrating its potential for operational forecasting. Yadav *et al* (2024) [4] focus on estimating the radius of maximum wind (RMW) in typhoons, a crucial parameter for assessing cyclone intensity and predicting impacts on climate and ecosystems. Using machine learning methods, they introduce an ensemble model combining random forests, gradient boosting, and NNs, enhancing precision in RMW estimation over traditional techniques.

Furthermore, several challenges persist in Typhoon modeling. One significant issue is the severe imbalance in the number of different types of Typhoons observed in nature. Most satellite images capture medium and low-intensity Typhoons, with high-intensity Typhoons relatively rare. The intensity of Typhoons spans a wide range, from 30 knots to over 150 knots. Nevertheless, Typhoon intensity estimation accuracy remains inadequate, particularly for high-intensity Typhoons, which are critically important.

Motivated by the SATCON technique, they "improve upon estimates provided by the individual members and a straight average of those estimates, and are competitive with existing traditional approaches such as the Dvorak technique."

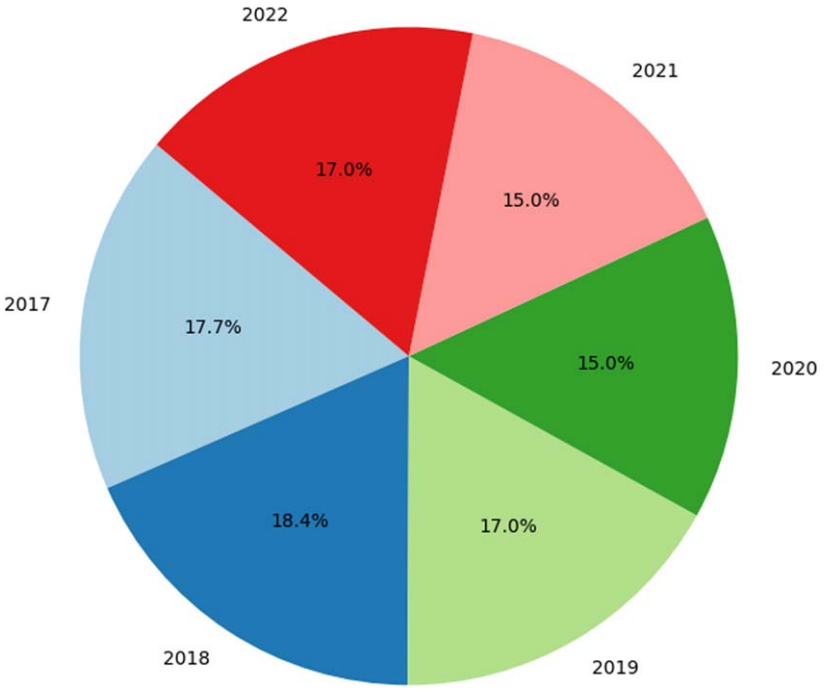


Figure 1. Representation of selected tropical storm for model training.

Using multiple techniques is crucial when single methods are not able to accurately estimate Typhoon intensity. We individually examine ADT, ARCHER, and SATCON, giving less accurate results. Using a neural network, we combine these outputs to minimize errors and maximize each method's strengths, resulting in more accurate and reliable Typhoon intensity estimates than traditional methods.

Our work aims to fill the current knowledge gap by accomplishing the following objectives:

- We propose a Neural Network model to estimate the Typhoon's Intensity.
- Comparing the accuracy of the proposed method with that of existing traditional methods (ARCHER, ADT, and SATCON) and one artificial approach (Artificial Intelligence-enhanced Dvorak Technique).

This paper is organized as follows: section 2 provides an overview of the data and research methodology, with a comprehensive discussion of the proposed method. Section 3 then compares this approach with alternative methods. Finally, section 4 concludes the paper and outlines directions for future research.

2. Data and method

2.1. Data

We collected information about each 6-hourly data point of the Typhoon's intensity from the three state-of-the-art intensity estimation models: ARCHER 'https://tropic.ssec.wisc.edu/real-time/archerOnline/cyclones/', ADT 'https://tropic.ssec.wisc.edu/misc/adt/info.html,' and SATCON 'https://tropic.ssec.wisc.edu/real-time/satcon/.' We gathered data from 162 typhoons, including tropical storms, using 147 (representation of selected cases according to yearly is shown in figure 1) for model training (table 1) and 15 (category representation of selected case is shown in figure 2) for validating (table 2) our proposed model.

The study primarily focused on the West Pacific Basin regions for training and validation. To ensure comprehensive data acquisition, the Best Track data has been collected from the official websites of the Regional Specialized Meteorological Center (RSMC) for the West Pacific Basin (https://www.jma.go.jp/jma/jma-eng/jma-center/rsmc-hp-pub-eg/RSMC_HP.htm) for 2017–2024.

We decided not to use satellite images in our study to estimate the intensity of Typhoons. Instead, we take the data from three advanced intensity estimation models: ARCHER, ADT, and SATCON. This decision was intentional, and we recognize that our approach depends on the information provided by these trusted models.

Table 1. In this study, the following typhoons were considered for training the model over the West Pacific Basin. PrM stands for pre-monsoon, PoM stands for post-monsoon, MSW stands for maximum sustained wind, and kts stands for knots.

S. No.	Typhoon Name	Sea -son	Date	MSW (kts)	Classification
2022					
1	Malakas	PrM	08–15 Apr.	90 kts	Very Strong Typhoon
2	Megi	PrM	09–11 Apr.	40 kts	Tropical Storm
3	Chaba	PrM	30 Jun.–03 Jul.	70 kts	Typhoon
4	Aere	PrM	30 Jun.–05 Jul.	45 kts	Tropical Storm
5	Songda	PrM	28–31 Jul.	40 kts	Tropical Storm
6	Trases	PrM	31 Jul.–01 Aug.	35 kts	Tropical Storm
7	Mulan	PrM	09–11 Aug.	35 kts	Tropical Storm
8	Meari	PrM	11–14 Aug.	40 kts	Tropical Storm
9	Ma-on	PrM	21–26 Aug.	55 kts	Sever Tropical Storm
10	Tokage	PrM	22–25 Aug.	75 kts	Typhoon
11	Hinnamnor	PoM	28 Aug.–06 Sep.	105 kts	Very Strong Typhoon
12	Muifa	PoM	07–16 Sep.	85 kts	Typhoon
13	Merbok	PoM	11–15 Sep.	70 kts	Typhoon
14	Nanmadol	PoM	13–19 Sep.	105 kts	Very Strong Typhoon
15	Talas	PoM	22–23 Sep.	35 kts	Tropical Storm
16	Noru	PoM	22–28 Sep.	95 kts	Very Strong Typhoon
17	Kulap	PoM	26–29 Sep.	60 kts	Severe Tropical Storm
18	Roke	PoM	28 Sep.–01 Oct.	70 kts	Typhoon
19	Sonca	PoM	14–15 Oct.	35 kts	Tropical Storm
20	Nesat	PoM	15–20 Oct.	75 kts	Typhoon
21	Haitang	PoM	18–19 Oct.	35 kts	Tropical Storm
22	Nalgae	PoM	27 Oct.–02 Nov.	60 kts	Severe Tropical Storm
23	Banyan	PoM	30 Oct.–01 Nov.	40 kts	Tropical Storm
24	Yamaneko	PoM	12–15 Nov.	35 kts	Tropical Storm
25	Pakhar	PoM	11–12 Dec.	40 kts	Tropical Storm
2021					
1	Dujuan	PrM	18–21 Feb.	40 kts	Tropical Storm
2	Surigae	PrM	12–30 Apr.	120 kts	Violent Typhoon
3	Choi-wan	PrM	30 May–05 Jun.	40 kts	Tropical Storm
4	Koguma	PrM	11–13 Jun.	35 kts	Tropical Storm
5	Champi	PrM	23–27 Jun.	65 kts	Typhoon
6	IN-FA	PrM	15–30 Jul.	85 kts	Typhoon
7	Cempaka	PrM	18–22 Jul.	70 kts	Typhoon

Table 1. (Continued.)

S. No.	Typhoon Name	Sea -son	Date	MSW (kts)	Classification
8	Nepartak	PrM	23–28 Jul	40 kts	Tropical Storm
9	Lupit	PrM	04–09 Aug.	45 kts	Tropical Storm
10	Mirinae	PrM	05–10 Aug.	50 kts	Severe Tropical Storm
11	Nida	PrM	04–08 Aug.	55 kts	Severe Tropical Storm
12	Omais	PrM	20–24 Aug.	45 kts	Tropical Storm
13	Conson	PoM	06–11 Sep.	50 kts	Severe Tropical Storm
14	Chanthu	PoM	05–20 Sep.	115 kts	Violent Typhoon
15	Dianmu	PoM	23–24 Sep.	35 kts	Tropical Storm
16	Mindulle	PoM	23 Sep.–02 Oct.	105 kts	Very Strong Typhoon
17	Lionrock	PoM	07–10 Oct.	35 kts	Tropical Storm
18	Kompasu	PoM	08–14 Oct.	55 kts	Severe Tropical Storm
19	Namtheun	PoM	10–17 Oct.	50 kts	Severe Tropical Storm
20	Malou	PoM	24–29 Oct.	75 kts	Typhoon
21	Nyatoh	PoM	30 Nov.–04 Dec.	100 kts	Very Strong Typhoon
22	Rai	PoM	11-21 Dec.	105 kts	Very Strong Typhoon
2020					
1	Vongfone	PrM	08–18 May	85 kts	Typhoon
2	Nuri	PrM	12–14 Jun.	40 kts	Tropical Storm
3	Sinlaku	PrM	01–02 Aug.	40 kts	Tropical Storm
4	Hagupit	PrM	01–05 Aug.	70 kts	Typhoon
5	Jangmi	PrM	08–11 Aug.	45 kts	Tropical Storm
6	Mekkhala	PrM	10–11 Aug.	50 kts	Severe Tropical Storm
7	Higos	PrM	18–19 Aug.	55 kts	Severe Tropical Storm
8	Bavi	PrM	22–27 Aug.	85 kts	Typhoon
9	Maysak	PoM	27 Aug.–07 Sept.	95 kts	Very Strong Typhoon
10	Haishen	PoM	30 Aug.–10 Sept.	105 kts	Very Strong Typhoon
11	Noul	PoM	15–18 Sept.	45 kts	Tropical Storm
12	Dolphin	PoM	21–24 Sept.	60 kts	Severe Tropical Storm
13	Kujira	PoM	26–30 Sept.	60 kts	Severe Tropical Storm
14	Chan-hom	PoM	05–11 Oct.	70 kts	Typhoon
15	Linfa	PoM	10–11 Oct.	45 kts	Tropical Storm
16	Nangka	PoM	12–14 Oct.	45 kts	Tropical Storm
17	Saudel	PoM	20–25 Oct.	65 kts	Typhoon
18	Molave	PoM	22–29 Oct.	90 kts	Very Strong Typhoon
19	Goni	PoM	26 Oct.–06 Nov.	115 kts	Violent Typhoon

Table 1. (Continued.)

S. No.	Typhoon Name	Sea -son	Date	MSW (kts)	Classification
20	Atsani	PoM	02–07 Nov.	50 kts	Severe Tropical Storm
21	Etau	PoM	08–10 Nov.	45 kts	Tropical Storm
22	Vamco	PoM	09–15 Nov.	85 kts	Typhoon
2019					
1	Wutip	PoM	08 Feb.-02 Mar.	105 kts	Very Strong Typhoon
2	Mun	PrM	02–04 Jul.	40 kts	Tropical Storm
3	Danas	PrM	14–23 Jul.	45 kts	Tropical Storm
4	Nari	PrM	24–28 Jul.	35 kts	Tropical Storm
5	Wipha	PrM	30 Jul.–03 Aug.	45 kts	Tropical Storm
6	Franciso	PoM	02–07 Aug.	70 kts	Typhoon
7	Lekima	PoM	02–15 Aug.	105 kts	Very Strong Typhoon
8	Krosa	PoM	06–16 Aug.	75 kts	Typhoon
9	Bailu	PoM	21–25 Aug.	50 kts	Severe Tropical Storm
10	Podul	PoM	28–31 Aug.	40 kts	Tropical Storm
11	Lingling	PoM	02–08 Sept.	95 kts	Very Strong Typhoon
12	Kajili	PoM	02–03 Sept.	35 kts	Tropical Storm
13	Faxai	PoM	04–10 Sept.	85 kts	Typhoon
14	Tapan	PoM	19–23 Sept.	65 kts	Typhoon
15	Mitag	PoM	28 Sept.–03 Oct.	75 kts	Typhoon
16	Hagibis	PoM	04–14 Oct.	105 kts	Very Strong Typhoon
17	Neoguri	PoM	17–21 Oct.	75 kts	Typhoon
18	Bualoi	PoM	19–25 Oct.	100 kts	Very Strong Typhoon
19	Matmo	PoM	29–31 Oct.	50 kts	Severe Tropical Storm
20	Halong	PoM	01–10 Nov.	115 kts	Violent Typhoon
21	Nakri	PoM	05–11 Nov.	65 kts	Typhoon
22	Fengshen	PoM	12–17 Nov.	85 kts	Typhoon
23	Kalmaegi	PoM	14–20 Nov.	70 kts	Typhoon
24	Fing-Wong	PoM	20–22 Nov.	55 kts	Severe Tropical Storm
25	Kammuri	PoM	26 Nov.–05 Dec.	90 kts	Very Strong Typhoon
2018					
1	Bolaven	PrM	03–04 Jan.	35 kts	Tropical Storm
2	Sanba	PrM	11–13 Feb.	35 kts	Tropical Storm
3	Jelawat	PrM	24 Mar.–01 Apr.	105 kts	Very Strong Typhoon
4	Ewiniar	PrM	05–08 Jun.	40 kts	Tropical Storm
5	Maliksi	PrM	07–11 Jun.	60 kts	Severe Tropical Storm

Table 1. (Continued.)

S. No.	Typhoon Name	Sea -son	Date	MSW (kts)	Classification
6	Gaemi	PrM	15–17 Jun.	45 kts	Tropical Storm
7	Prapiroon	PrM	28 Jun.-05 Jul.	65 kts	Typhoon
8	Maria	PrM	03–13 Jul.	105 kts	Very Strong Typhoon
9	Son-tinh	PrM	17–19 Jul.	40 kts	Tropical Storm
10	Ampli	PrM	18–23 Jul.	50 kts	Severe Tropical Storm
11	Wukong	PrM	23–27 Jul.	50 kts	Severe Tropical Storm
12	Jongdari	PrM	24 Jul.–03 Aug.	75 kts	Typhoon
13	Shanshan	PoM	02–11 Aug.	70 kts	Typhoon
14	Yagi	PoM	08–13 Aug.	40 kts	Tropical Storm
15	Leepi	PoM	11–15 Aug.	50 kts	Severe Tropical Storm
16	Bebinca	PoM	13–17 Aug.	45 kts	Tropical Storm
17	Rumbia	PoM	15–18 Aug.	40 kts	Tropical Storm
18	Soulik	PoM	16–24 Aug.	85 kts	Typhoon
19	Cimron	PoM	18–24 Aug.	85 kts	Typhoon
20	Jebi	PoM	27 Aug.–05 Sept.	105 kts	Very Strong Typhoon
21	Mang-Khut	PoM	07–17 Sept.	110 Kts	Very Strong Typhoon
22	Barijat	PoM	11–13 Sept.	40 kts	Tropical Storm
23	Trami	PoM	20 Sept.-03 Oct.	105 kts	Very Strong Typhoon
24	Kong-Rey	PoM	28 Sept.-07 Oct.	115 kts	Violent Typhoon
25	Yutu	PoM	22 Oct.–02 Nov.	115 kts	Violent Typhoon
26	Toraji	PoM	17 Nov.–18 Nov.	35 kts	Tropical Storm
27	Man-yi	PoM	20–27 Nov.	80 kts	Typhoon
2017					
1	Merbok	PrM	11–13 June	55 kts	Severe Tropical Storm
2	Nanmadol	PrM	02–05 July	55 kts	Severe Tropical Storm
3	Talas	PrM	15–17 July	50 kts	Severe tropical Storm
4	Noru	PrM	19 July–12 Aug.	95 kts	Very Strong Typhoon
5	Kulap	PrM	21–25 July	40 kts	Tropical Storm
6	Roke	PrM	22–23 July	35 kts	Tropical Storm
7	Sonca	PrM	23–25 July	35kts	Tropical Storm
8	Nesat	PrM	25–30 July	80 kts	Typhoon
9	Haitang	PrM	28–31 July	45 kts	Tropical Storm
10	Nalgae	PoM	02–05 Aug.	45 kts	Tropical Storm
11	Banyan	PoM	11–17 Aug.	80 kts	Typhoon
12	Hato	PoM	19–24 Aug.	75 kts	Typhoon

Table 1. (Continued.)

S. No.	Typhoon Name	Sea -son	Date	MSW (kts)	Classification
13	Pakhar	PoM	24–27 Aug.	55	Severe Tropical Storm
14	Sanvu	PoM	26 Aug.- 06 Sept.	80 kts	Typhoon
15	Mawar	PoM	31 Aug.–04 Sept.	50 kts	Severe Tropical Storm
16	Guchal	PoM	05–06 Sept.	35 kts	Tropical Storm
17	Talim	PoM	08–22 Sept.	95 kts	Very Strong Typhoon
18	Doksuri	PoM	12–16 Sept.	80 kts	Typhoon
19	Khanun	PoM	12–16 Oct.	75 kts	Typhoon
20	Lan	PoM	15–23 Oct.	100 kts	Very Strong Typhoon
21	Saola	PoM	24–29 Oct.	60 kts	Severe Tropical Storm
22	Damray	PoM	02–04 Nov.	70 kts	Typhoon
23	Haikui	PoM	10–12 Nov.	40 kts	Tropical Storm
24	Kirogi	PoM	18–19 Nov.	35 kts	Tropical Storm
25	Kai-tak	PoM	14–21 Dec.	40 kts	Tropical Storm
26	Tembin	PoM	20–25 Dec.	70 kts	Typhoon

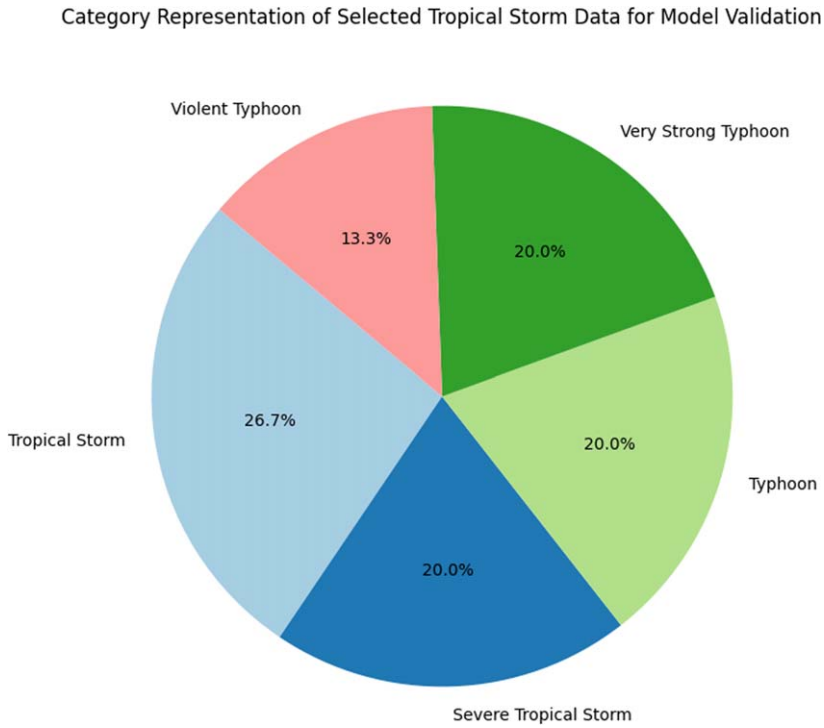


Figure 2. Category representation of selected tropical storm for model validation.

2.2. Method

In this study, we develop a method to estimate the value of Typhoon intensity over the West Pacific basin using a machine-learning technique. We compared the results of the proposed method with the three other methods: ARCHER, ADT, and SATCON.

The brief descriptions of the current members in our model are as follows:

- **ARCHER** (**Automated Rotational Center Hurricane Eye Retrieval**): The ARCHER estimates Typhoon characteristics using the Advanced Technology Microwave Sounder (ATMS)-derived minimum sea level pressure (MSLP) anomaly, Typhoon eyewall vigor, and the inner core brightness temperature (Tb) gradient (maximum of channel 8 or 9) [26].
- **ADT** (**Advanced Dvorak technique**): The ADT is a computer method designed with great care to determine Typhoon features by studying infrared data from geostationary satellites. Operational Typhoon analysis and forecasting centers worldwide commonly use it to help determine Typhoon strength, especially in maritime areas where direct observations are impossible. The algorithm was originally developed to closely resemble the Dvorak approach, based on previous versions provided by [27]. In this conventional approach, a skilled analyst uses pattern-matching and classification techniques to estimate the characteristics of Typhoons based on satellite images. The ADT algorithm has gradually developed into a functional instrument, officially supported and upheld by NOAA/NESDIS. It has become a dependable source for delivering up-to-date, unbiased advice on Typhoon characteristics to operational Typhoon centers across the globe.
- **SATCON** (**Satellite Consensus**): The SATCON, developed by CIMSS, combines the ADT estimation acquired from Geostationary Satellites with other estimates collected during polar-orbiting satellite overpasses. This combination provides a complete assessment of Typhoon features worldwide, as outlined in research conducted by [11–13, 26, 28–30].

The input for our method is the results obtained from the three methods: ARCHER, ADT, and SATCON. The target for our method is the best track data provided by the RSMC over the West Pacific basin.

Here is the whole process of our method (flowchart presenting each step of the proposed method is shown in figure 3):

1. **Preprocessing:** The three sets of experimental values are first combined and normalized using the StandardScaler from scikit-learn [31] to ensure that all the input data are on a similar scale.

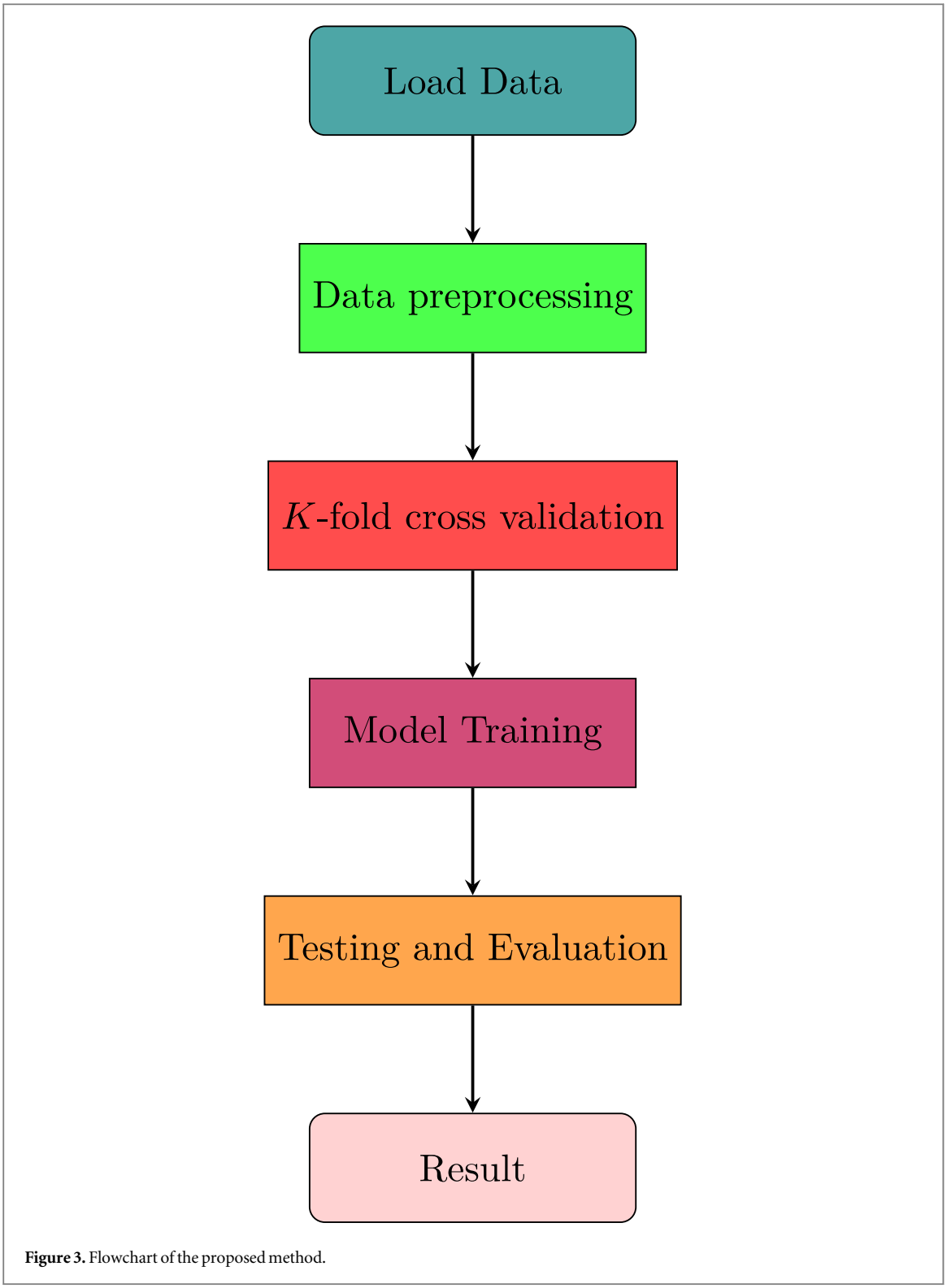


Figure 3. Flowchart of the proposed method.

2. ***k*-fold cross-validation:** *k*-fold cross-validation is a reliable method for assessing the performance of machine learning models [32]. It involves dividing the dataset into *k* equally sized folds, training the model on $k-1$ folds, and then evaluating it on the remaining fold. This procedure is repeated *k* times. [33, 34]. For example, we divided our data into five equal groups. To test a model, we use four groups for training and the remaining group to check how well the model performs We do this five times, using each group to check once.
   - Iteration 1: Use folds 2, 3, 4, and 5 to train the model, then test it on fold 1.
   - Iteration 2: Use folds 1, 3, 4, and 5 to train the model, then test it on fold 2.
   - Iteration 3: Use folds 1, 2, 4, and 5 to train the model, then test it on fold 3.

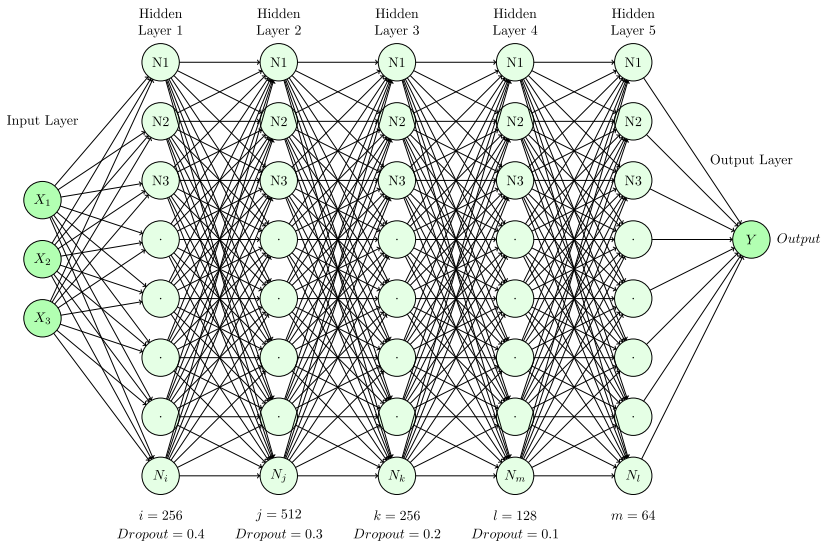


Figure 4. Architecture of the Proposed Neural Network.

Table 2. In this study, the following typhoons were considered for validating our proposed method over the West Pacific Basin.

S. No.	Typhoon Name	Season	Date	Maximum Wind Speed (knots (kts))	Classification
2024					
1	Ewiniar	PrM	25 - 30 May	70 kts	Typhoon
2	Gaemi	PrM	20 - 25 Jul.	90 kts	Very Strong Typhoon
3	Prapiron	PrM	21 - 23 Jul	55 kts	Severe Tropical Storm
2023					
1	Sanvu	PrM	20 Apr. - 21 Apr.	45 kts	Tropical Storm
2	Mawar	PrM	20 May - 31 May	115 kts	Violent Typhoon
3	Guchol	PrM	06 - 14 Jun.	80 kts	Typhoon
4	Talim	PrM	15 - 17 Jul.	60 kts	Severe Tropical Storm
5	Doksuri	PrM	21 - 28 Jul.	100 kts	Very Strong Typhoon
6	Lan	PoM	08 - 16 Aug.	90 kts	Very Strong Typhoon
7	Damrey	PoM	25 - 28 Aug.	50 kts	Severe Tropical Storm
8	Haikui	PoM	29 Aug. - 04 Sep.	85 kts	Typhoon
9	Kirogi	PoM	30 Aug. - 02 Sep.	45 kts	Tropical Storm
10	Yun-Yeung	PoM	05 - 08 Sep.	40 kts	Tropical Storm
11	Bolaven	PoM	07 - 14 Oct.	115 kts	Violent Typhoon
12	Sanba	PoM	18 - 19 Dec.	40 kts	Tropical Storm

- Iteration 4: Use folds 1, 2, 3, and 5 to train the model, then test it on fold 4.
- Iteration 5: Use folds 1, 2, 3, and 4 to train the model, then test it on fold 5.

3. **Neural Network Architecture:** The model is constructed via the Keras Sequential API (Application Programming Interface) [35]. The architecture consists of many layers, such as densely linked with LeakyReLU activation, batch normalization, and dropout layers, which help avoid overfitting. The model's architecture (figure 4) facilitates deep learning using five hidden layers of 256, 512, 256, 128, and 64 neurons.
4. **Training the Model:** The compiled model is trained using the Adam optimizer [36] with a reduced learning rate of 0.00001, optimizing the mean squared error loss function. To enhance the model's learning process, a substantial number of epochs, 10,000 in this case, are executed during the training process. The batch size is set to 128, and the training progress is tracked for analysis.
5. **Model Evaluation:** Following training, the model is evaluated using the trained neural network to predict the experimental values. The mean absolute percentage and root mean square errors are calculated to assess the model's accuracy and performance.

Table 3. Numerical results of typhoon ewiniar.

Date/Time	RSMC Tokyo (knots)	Our Method (knots)	ARCHER (knots)	ADT (knots)	SATCON (knots)	AiDT (knots)	E_1[a]	E_2[b]	E_3[c]	E_4[d]	E_5[e]
25 May/1200	35	35.0	34.9	41	44	33	0.00	0.28	17.1	25.7	5.71
25 May/1800	45	39.7	35.0	47	51	37	11.7	22.2	4.44	13.3	17.7
26 May/0000	50	51.5	54.6	63	65	52	3.00	9.20	26.0	30.0	4.00
26 May/0600	55	53.9	55.0	72	75	74	2.00	0.00	30.9	36.3	34.5
26 May/1200	65	59.8	73.7	77	78	80	8.00	13.3	18.4	20.0	23.0
26 May/1800	70	65.9	84.4	84	88	85	5.85	20.5	20.0	25.7	21.4
27 May/0000	75	72.1	89.6	92	79	82	3.86	19.4	22.6	5.33	9.33
27 May/0600	70	70.5	94.9	74	74	78	0.71	35.5	5.71	5.71	11.4
27 May/1200	70	69.8	85.7	67	71	65	0.28	22.4	4.28	1.42	7.14
27 May/1800	70	67.2	80.4	55	64	62	4.00	14.8	21.4	8.57	11.4
28 May/0000	65	65.1	75.4	51	66	50	0.15	16.0	21.5	1.53	23.0
28 May/0600	65	62.9	70.4	49	64	51	3.23	8.30	24.6	1.53	21.5
28 May/1200	65	61.1	65.1	47	61	57	6.00	0.15	27.6	6.15	12.3
28 May/1800	65	59.4	74.8	45	59	58	8.61	15.0	30.7	9.23	10.7
29 May/0000	60	57.1	70.4	43	55	47	4.83	17.3	28.3	8.33	21.6
29 May/0600	60	55.4	65.4	41	56	42	7.66	9.00	31.6	6.66	30.0
29 May/1200	55	52.8	60.4	47	54	41	4.00	9.81	14.5	1.81	25.4
29 May/1800	50	49.8	55.4	47	54	45	0.40	10.8	6.00	8.00	10.0
30 May/0000	45	47.8	50.1	51	56	49	6.22	11.3	13.3	24.4	8.88
30 May/0600	40	42.9	50.0	33	46	45	7.25	25.0	17.5	15.0	12.5
30 May/1200	40	37.9	45.1	28	39	35	5.25	12.7	30.0	2.50	12.5
Mean							4.43	13.97	19.58	12.25	15.93
RMSE		3.07	10.17	12.91	8.65	10.68					
R^2		0.92	0.21	-0.26	0.43	0.13					

[a] Error Percentage of our proposed method with respect to Best track data provided by the RSME Tokyo.
[b] Error Percentage of ARCHER method with respect to Best track data provided by the RSME Tokyo.
[c] Error Percentage of ADT method with respect to Best track data provided by the RSME Tokyo.
[d] Error Percentage of SATCON method with respect to Best track data provided by the RSME Tokyo.
[e] Error Percentage of AiDT method with respect to Best track data provided by the RSME Tokyo.

Table 4. Numerical results of Very Strong Typhoon Gaemi.

Date/Time	RSMC Tokyo (knots)	Our Method (knots)	ARCHER (knots)	ADT (knots)	SATCON (knots)	AiDT	E_1^1	E_2^2	E_3^3	E_4^4	E_5^5
20 Jul./0000	35	33.7	29.9	29	41	30	3.71	14.5	17.1	17.1	14.2
20 Jul./0600	40	37.9	30.0	32	36	33	5.25	25.0	20.0	10.0	17.5
20 Jul./1200	40	39.1	30.0	34	40	34	2.25	25.0	15.0	0.00	15.0
20 Jul./1800	40	42.7	39.3	37	43	34	6.75	1.75	7.50	7.50	15.0
21 Jul./0000	45	44.5	44.6	43	49	37	1.11	0.88	4.44	8.88	17.7
21 Jul./0600	50	47.1	49.9	45	49	40	5.80	0.20	10.0	2.00	20.0
21 Jul./1200	50	49.4	50.0	49	47	43	1.20	0.00	2.00	6.00	14.0
21 Jul./1800	60	55.9	54.9	41	48	42	6.83	8.50	31.6	20.0	30.0
22 Jul./0000	65	59.2	50.1	49	57	49	8.92	22.9	24.6	12.3	34.6
22 Jul./0600	65	62.7	54.6	53	62	46	3.53	16.0	18.4	4.61	29.2
22 Jul./1200	65	64.2	59.5	63	72	63	1.23	8.46	3.07	10.7	3.07
22 Jul./1800	70	67.8	69.1	72	72	77	3.14	1.28	2.85	2.85	10.0
23 Jul./0000	75	70.5	79.2	79	79	76	6.00	5.60	5.33	5.33	1.33
23 Jul./0600	75	72.2	89.4	84	81	83	3.73	19.2	12.0	8.00	10.6
23 Jul./1200	75	75.0	94.4	84	86	85	0.00	25.8	12.0	14.6	13.3
23 Jul./1800	85	78.9	119.6	84	102	83	7.17	40.0	1.17	20.0	2.35
24 Jul./0000	85	81.7	120.0	102	124	99	3.88	41.1	20.0	45.8	16.4
24 Jul./0600	90	84.2	124.9	115	133	113	6.44	38.7	27.7	47.7	25.5
24 Jul./1200	65	74.5	111.3	112	113	108	14.6	71.2	72.3	73.8	66.1
24 Jul./1800	60	69.4	96.4	107	94	101	15.6	60.6	78.3	56.6	68.3
25 Jul./0000	60	65.0	76.5	72	79	62	8.33	27.5	20.0	31.6	3.33
25 Jul./0600	55	62.7	61.0	67	66	57	14.0	10.9	21.8	20.0	3.63
Mean							5.89	21.15	19.43	19.36	19.16
RMSE		4.55	19.7	17.50	19.22	16.21					
R^2		0.91	-0.63	-0.28	-0.54	-0.09					

Table 5. Numerical results of Severe Tropical Storm Prapiron.

Date/Time	RSMC Tokyo (knots)	Our Method (knots)	ARCHER (knots)	ADT (knots)	SATCON (knots)	AiDT	E_1^1	E_2^2	E_3^3	E_4^4	E_5^5
21 Jul./0000	35	33.5	30.0	31	40	34	4.28	14.2	11.4	14.2	2.85
21 Jul./0600	35	35.9	30.0	37	43	33	2.57	14.2	5.71	22.8	5.71
21 Jul./1200	40	37.8	30.0	45	48	39	5.50	25.0	12.5	20.0	2.50
21 Jul./1800	45	42.9	34.9	51	52	44	4.66	22.4	13.3	15.5	2.22
22 Jul./0000	50	47.5	35.0	53	60	49	5.00	30.0	6.00	20.0	2.00
22 Jul./0600	55	53.2	35.0	65	66	52	3.27	36.3	18.1	20.0	5.45
22 Jul./1200	55	55.9	39.4	74	70	70	1.63	28.3	34.5	27.2	27.2
22 Jul./1800	55	57.8	49.8	79	62	64	5.09	9.45	43.6	12.7	16.3
23 Jul./0000	45	51.9	45.1	67	60	59	15.3	0.22	48.8	33.3	31.1
23 Jul./0600	35	45.8	54.3	61	51	47	30.8	55.1	74.2	45.7	34.2
Mean							7.82	23.55	26.85	23.17	12.97
RMSE		4.41	12.32	15.12	10.85	8.14					
R^2		0.70	-1.33	-2.51	-0.81	-0.02					

Table 6. Numerical results of Tropical Storm Sanvu.

Date/Time	RSMC Tokyo (knots)	Our Method (knots)	ARCHER (knots)	ADT (knots)	SATCON (knots)	AiDT (knots)	E_1^1	E_2^2	E_3^3	E_4^4	E_5^5
20 Apr./1200	35	37.8	39.6	53	51	43	8.00	13.14	51.4	45.7	22.8
20 Apr./1800	40	42.7	44.9	55	49	43	6.75	12.25	37.5	22.5	07.5
21 Apr./0000	40	42.3	45.0	55	48	42	5.75	12.50	37.5	20.0	05.0
21 Apr./0600	45	40.6	45.0	51	50	41	9.77	0.000	13.3	11.1	08.8
21 Apr./1200	40	39.2	45.0	53	47	41	2.00	12.50	32.5	17.5	02.5
21 Apr./1800	35	33.9	40.6	45	42	39	3.14	16.00	28.5	20.0	11.4
Mean							5.90	11.06	33.47	22.80	9.69
RMSE		2.63	4.59	13.14	9.34	4.28					
R^2		0.41	-0.78	-14.23	-6.39	-0.55					

To assess the accuracy of our neural network model, we use four statistical techniques: t-test, error percentage, R-squared, and root mean square error. Using these techniques, we confidently assert that our neural network-based model achieves more efficiency and accuracy than the three traditional methods (ADT, ARCHER, and SATCON) and one artificial approach (Artificial Intelligence-enhanced Dvorak Technique).

The Artificial Intelligence-enhanced Dvorak Technique (AiDT) [37] represents a modern evolution of the traditional Dvorak methodology for estimating tropical cyclone intensity using geostationary satellite data. Traditional Dvorak methods relied on pattern recognition and empirical rules, but AiDT integrates machine learning to enhance objectivity and accuracy. This hybrid approach allows the extraction of intricate satellite features, such as cloud structures and thermal patterns, which are processed by advanced algorithms to improve intensity predictions.

t-test:

The t-test serves as a statistical tool to assess differences in mean values between two datasets, validating the effectiveness of our model against alternative methods. This involves comparing our model's performance metrics or predictions with existing models. By calculating the t-value and corresponding p-value, we determine the statistical significance of mean differences, establishing whether our method outperforms alternatives. This approach objectively evaluates the effectiveness and reliability of our proposed model, providing empirical evidence that supports its accuracy in predicting the desired outcome.

Error Percentage:

Error percentage, also known as error rate or relative error, measures the relative difference between predicted and actual values. It quantifies how far off the model's predictions are from the ground truth. The error percentage can be calculated as follows:

$$\text{Error percentage} = \left(\frac{|\text{Predicted Value} - \text{Actual Value}|}{\text{Actual value}} \right) \times 100$$

A lower error percentage indicates a more accurate model, which signifies that the model's predictions are closer to the actual values.

Root Mean Square Error (RMSE):

RMSE is a widely used metric to assess the accuracy of regression models. It measures the average magnitude of the errors between predicted and actual values. The RMSE is calculated as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n} \times \sum (\text{Predicted Value} - \text{Actual Value})^2}$$

A lower RMSE indicates that the model's predictions are, on average, closer to the actual values. It represents the standard deviation of the model's errors.

R^2 (R-squared):

R^2 is a statistical measure representing the proportion of the variance for a dependent variable explained by an independent variable or variables in a regression model. It ranges from 0 to 1, where:

- $R^2 = 1$ indicates that the regression model predicts the dependent variable perfectly.
- $R^2 = 0$ indicates that the model does not explain any of the variability of the dependent variable around its mean.

3. Result and discussion

The NN-based model proposed for Typhoon intensity estimation has been assessed across fifteen cases. Table 2 presents information on selected Typhoons for validating, detailing their respective categories, seasons, and time periods.

3.1. Typhoon Ewiniar

Typhoon Ewiniar (2024) slammed into the Philippines in May 2024, bringing heavy rains, floods, and landslides. Thousands were displaced and sought shelter in evacuation centers. The typhoon also caused significant damage to homes, infrastructure, and agriculture

Figure 5 compares the performance of our proposed typhoon intensity estimation method to traditional approaches (ARCHER, ADT, and SATCON) and the artificial learning approach, i.e., AiDT using RSMC Tokyo's Best Track data. Table 3 summarizes the results of Typhoon Ewiniar.

For Typhoon Ewiniar, our method achieved a 4.43% error rate, outperforming ARCHER (13.97%), ADT (19.58%), SATCON (12.25%), and AiDT (15.93%). RMSE values are detailed in table 20. A t-test confirms our method's superior performance. Table 21 shows R^2 values of 0.92, 0.12, -0.26, 0.43, and 0.13 for the methods. An

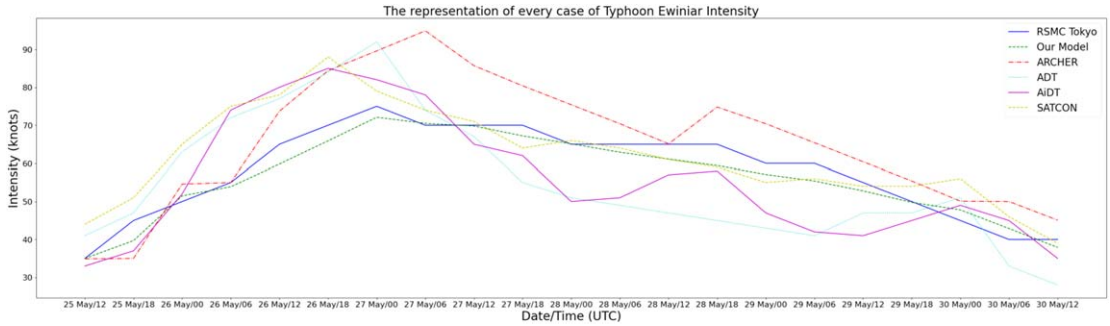


Figure 5. A graphic representation for each method's result for intensity of typhoon ewiniar.

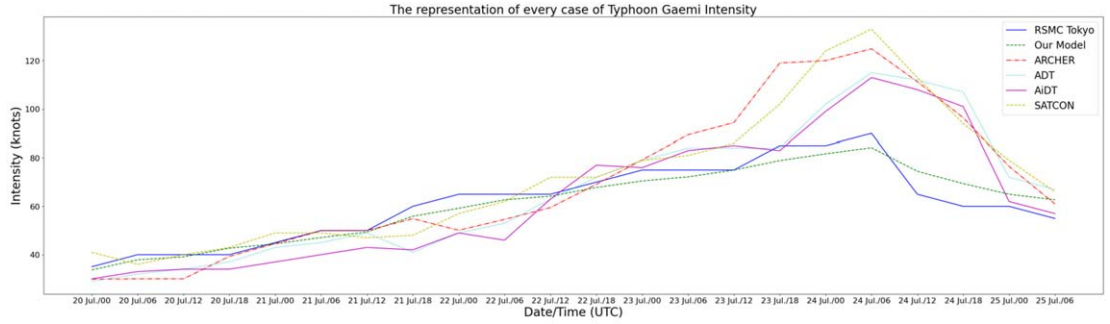


Figure 6. A graphic representation for each method's result for intensity of very strong typhoon gaemi.

R^2 of 0.92 indicates strong predictive accuracy, capturing 92% of data variability. Moderate fits are shown by R^2 values of 0.43, explaining 43% variability. However, the negative R^2 reveals poor performance, suggesting the model is less reliable than using the mean. Interpreting R^2 values in context is crucial for understanding model alignment with data patterns.

3.2. Very strong Typhoon gaemi

Very Strong Typhoon Gaemi (2024) was a powerful typhoon that affected East China, Taiwan, and the Philippines in late July 2024. It caused significant damage and casualties in these regions, including flooding, landslides, and power outages. The typhoon reached its peak intensity, with maximum sustained winds of 165 km/h. Figure 6 compares the performance of our proposed typhoon intensity estimation method with traditional methods: ARCHER, ADT, SATCON, and artificial learning approach, i.e., AiDT against RSMC Tokyo's Best Track data. Table 4 summarizes the quantitative results for Typhoon Gaemi.

For Very Strong Typhoon Gaemi, our method achieves an error rate of 5.89%, outperforming ARCHER (21.15%), ADT (19.43%), SATCON (19.36%) and AiDT (19.16%). The RMSE values are detailed in table 20. A t-test confirms our method's superiority over others.

The R^2 values (table 21) further highlight model performance: 0.91 (our method), -0.63 (ARCHER), -0.28 (ADT), 0.54 (SATCON), and −0.09 (AiDT). An R^2 of 0.91 indicates our method explains 91% of the data variation, representing a good fit. In contrast, a negative R^2 indicates poor performance, suggesting the model underperforms relative to using the mean for predictions. Evaluating R^2 values in context helps assess model alignment with data patterns.

3.3. Severe tropical storm prapiron

Severe Tropical Storm Prapiron (2024) formed in the South China Sea in July 2024. It made landfall in Hainan, China, and Vietnam, bringing heavy rain and strong winds. The storm caused flooding, landslides, and power outages in the affected areas. Figure 7 illustrates the comparative performance of the proposed intensity estimation method relative to conventional approaches (ARCHER, ADT, and SATCON) and also with the artificial approach (AiDT) using Best Track data from RSMC Tokyo. A quantitative summary of Severe Tropical Storm Prapiron is presented in table 5.

The proposed method demonstrates an error rate of 7.82%, outperforming ARCHER (23.55%), ADT (26.85%), SATCON (23.17%), and AiDT (12.97%). RMSE values, detailed in table 20, further validate the

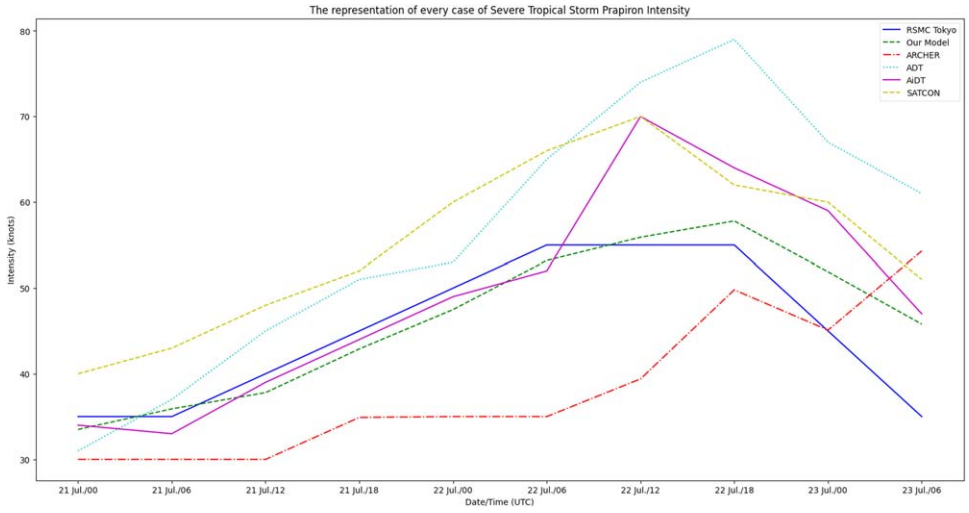


Figure 7. A graphic representation for each method's result for intensity of severe tropical storm prapiron.

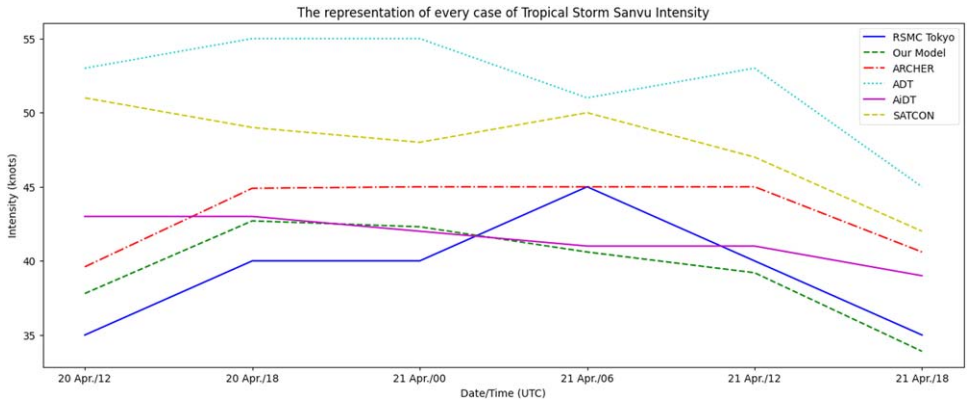


Figure 8. A graphic representation for each method's result for intensity of tropical storm sanvu.

improved accuracy. Statistical analysis via a t-test confirms the statistical significance of the proposed method's superior performance.

The R^2 values (table 21) quantify model efficacy, with the proposed method achieving 0.70, indicating that it accounts for 70% of the variability in the observed data. ARCHER, ADT, SATCON, and AiDT yield a negative R^2, reflecting the models' performance inferior to a baseline prediction using the mean. These results emphasize the better estimative capacity of our proposed method and its alignment with the observed data patterns.

3.4. Tropical storm Sanvu

Tropical Storm Sanvu was the first named storm of the 2023 Pacific typhoon season. It formed in April 2023 and remained a weak tropical storm throughout its lifespan. Sanvu did not make landfall and dissipated over open waters. Figure 8 illustrates the performance of our proposed method for estimating typhoon intensity in comparison to traditional approaches (ARCHER, ADT, and SATCON) and the artificial approach (AiDT) using Best Track data from RSMC Tokyo. Quantitative results for Tropical Storm Sanvu are summarized in table 6.

The proposed method achieves a percentage error of 5.90%, outperforming ARCHER (11.06%), ADT (33.47%), SATCON (22.80%), and AiDT (9.69%). Root Mean Square Error (RMSE) values are provided in table 20, and statistical analysis using a t-test confirms the significant improvement of the proposed method over the others. Table 21 presents the R^2 values for the models: 0.41 (proposed method), -0.78 (ARCHER0, -6.39 (SATCON), -14.23 (ADT), and -0.55 (AiDT). The R^2 value of 0.41 indicates that the proposed method explains 41% of the variance in the intensity data, representing a strong fit. In contrast, the negative R^2 value implies that it performs worse than a baseline prediction using the mean intensity, suggesting limitations in its applicability or alignment with the dataset. These results demonstrate the fitness of the proposed method in accurately modeling typhoon intensity.

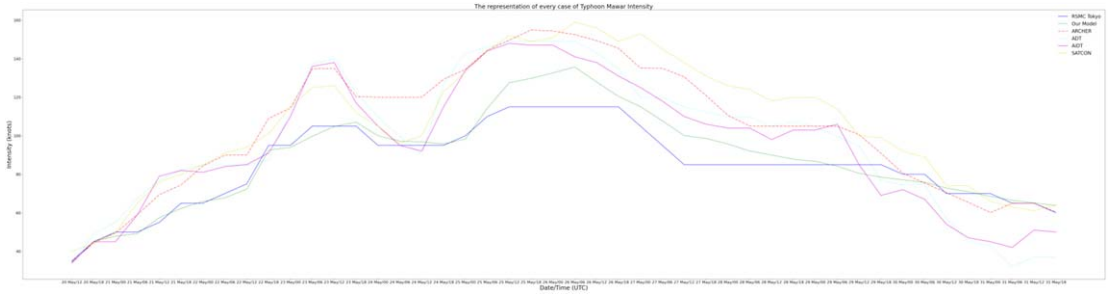


Figure 9. A graphic representation for each method's result for intensity of violent typhoon mawar.

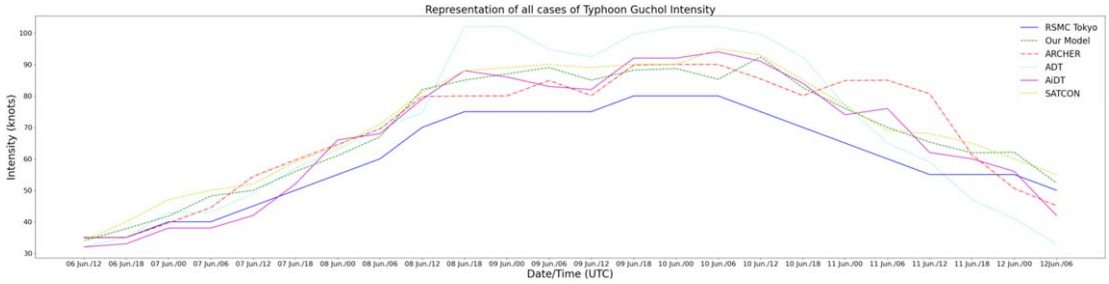


Figure 10. Graphically representation of all results of typhoon guchol.

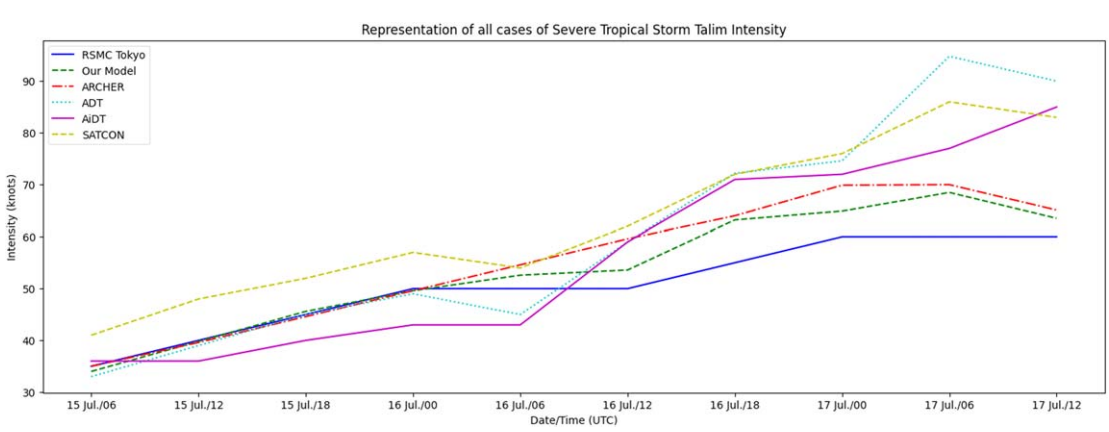


Figure 11. Graphically representation of all results of severe tropical storm talim.

3.5. Violent typhoon mawar

Violent Typhoon Mawar (2023) was one of the strongest Northern Hemisphere tropical storms on record in May and the strongest worldwide in 2023. It caused significant damage to Guam, including widespread power outages, downed trees, and coastal flooding. Figure 9 shows the comparative performance of our proposed typhoon intensity estimation method against established approaches (ARCHER, ADT, SATCON, and AiDT), benchmarked using Best Track data from RSMC Tokyo. A detailed summary of the results of Typhoon Mawar is provided in table 7.

The proposed method achieves a percentage error of 5.75%, demonstrating superior accuracy over ARCHER (18.79%), ADT (24.86%), SATCON (23.02%), and AiDT (20.17 %). Corresponding RMSE values, detailed in table 20, further emphasize this improvement. Statistical validation using a t-test confirms the proposed method's significant advantage over traditional techniques.

The R^2 values, presented in table 21, further distinguish model performance: 0.90 (proposed method), 0.05 (ARCHER), -0.21 (SATCON), 0.06 (ADT), and 0.31 (AiDT). A R^2 value of 0.90 suggests that the proposed method captures 90% of the variance in intensity, indicating a strong model fit. In contrast, ARCHER and SATCON exhibit moderate fits, explaining 5% of the variability, while SATCON's negative R^2 of -0.21 highlights

Table 7. Numerical results of Violent Typhoon Mawar.

Date/Time	RSMC Tokyo (knots)	Our Method (knots)	ARCHER (knots)	ADT (knots)	SATCON (knots)	AiDT (knots)	E_1^1	E_2^2	E_3^3	E_4^4	E_5^5
20 May/1200	35	33.9	34.5	34.0	40	34	3.14	1.42	2.85	14.2	2.85
20 May/1800	45	45.2	44.4	49.0	44	45	0.44	1.33	8.88	2.22	0.00
21 May/0000	50	47.8	49.5	55.0	50	45	04.4	1.00	10.0	0.00	10.0
21 May/0600	50	49.2	59.1	67.4	65	59	01.6	18.2	34.8	30.0	18.0
21 May/1200	55	57.4	69.4	77.0	76	79	4.36	26.1	40.0	38.1	43.6
21 May/1800	65	62.3	74.5	82.2	80	82	4.15	14.6	26.4	23.0	26.1
22 May/0000	65	65.9	84.4	84.8	85	81	1.38	29.8	30.4	30.7	24.6
22 May/0600	70	67.8	89.9	84.8	91	84	3.14	28.4	21.1	30.0	20.0
22 May/1200	75	72.3	90.0	84.8	94	85	3.60	20.0	13.0	25.3	13.3
22 May/1800	95	92.5	109.0	87.4	101	91	2.63	14.7	8.00	6.31	4.21
23 May/0000	95	94.0	114.4	112.4	115	110	1.05	20.4	18.3	21.0	15.7
23 May/0600	105	99.7	134.7	137.4	125	136	5.04	28.2	30.8	19.0	29.5
23 May/1200	105	104.8	135.0	140.0	126	138	0.19	28.5	33.3	20.0	31.4
23 May/1800	105	107.0	120.3	122.2	112	117	1.90	14.5	16.3	6.66	11.4
24 May/0000	95	100.0	120.0	109.8	105	105	5.26	26.3	15.5	10.5	10.5
24 May/0600	95	97.2	120.0	99.6	96	95	2.31	26.3	4.84	1.05	0.00
24 May/1200	95	96.8	120.0	90.0	100	92	1.89	26.3	5.26	5.26	3.15
24 May/1800	95	95.7	129.3	127.0	123	115	0.73	36.1	33.6	29.4	21.0
25 May/0000	100	98.3	134.5	143.0	133	134	1.70	34.5	43.0	33.0	34.0
25 May/0600	110	114.0	144.4	146.0	144	144	3.90	31.2	32.7	30.9	30.9
25 May/1200	115	127.5	149.6	149.0	152	148	10.8	30.0	29.5	32.1	28.6
25 May/1800	115	129.8	154.9	149.0	149	147	12.8	34.6	29.5	29.5	27.8
26 May/0000	115	132.6	154.4	149.0	151	147	15.3	34.2	29.5	31.3	27.8
26 May/0600	115	137.7	152.5	149.0	159	141	18.0	32.6	29.5	38.2	22.6
26 May/1200	115	127.9	149.4	143.0	156	138	11.2	29.9	24.3	35.6	20.0
26 May/1800	115	120.4	145.4	134.8	149	131	4.69	26.4	17.2	29.5	13.9
27 May/0000	105	115.1	135.2	119.8	153	125	9.61	28.7	14.0	45.7	19.0
27 May/0600	95	107.8	135.0	119.8	145	118	13.4	42.1	26.1	52.6	24.2
27 May/1200	85	100.1	130.6	115.0	138	110	17.7	53.6	35.2	62.3	29.4
27 May/1800	85	98.6	120.9	112.4	131	106	16.0	42.2	32.2	54.1	24.7
28 May/0000	85	95.7	110.6	109.8	126	104	12.5	30.1	29.1	48.2	22.3
28 May/0600	85	92.1	105.1	109.8	124	104	8.35	23.6	29.1	45.8	22.3
28 May/1200	85	90.0	105.0	104.6	118	98	5.88	23.5	23.0	38.8	15.2
28 May/1800	85	87.8	105.0	107.2	120	103	3.29	23.5	26.1	41.1	21.1

Table 7. (Continued.)

Date/Time	RSMC Tokyo (knots)	Our Method (knots)	ARCHER (knots)	ADT (knots)	SATCON (knots)	AiDT (knots)	E_1^1	E_2^2	E_3^3	E_4^4	E_5^5
29 May/0000	85	86.7	105.0	104.6	120	103	2.00	23.5	23.0	41.1	21.1
29 May/0600	85	84.2	105.0	99.6	114	106	0.94	23.5	17.1	34.1	24.7
29 May/1200	85	80.5	100.6	94.8	100	85	5.29	18.3	11.5	17.6	0.00
29 May/1800	85	78.5	90.9	77.0	99	69	7.64	6.95	9.41	16.4	18.8
30 May/0000	80	77.1	80.6	74.6	92	72	3.62	0.75	6.75	15.0	10.0
30 May/0600	80	75.8	75.4	74.6	89	67	5.25	5.75	6.75	11.2	16.2
30 May/1200	70	72.8	70.4	57.0	74	54	4.00	0.57	18.5	5.71	22.8
30 May/1800	70	70.9	65.4	45.0	74	47	1.28	6.57	35.7	5.71	32.8
Mean							5.75	18.79	24.86	23.02	20.17
RMSE		6.96	22.24	22.12	25.15	18.84					
R^2		0.90	0.05	0.06	-0.21	0.31					

Table 8. Numerical results of Guchol.

Date/Time	RSMC Tokyo (knots)	Our Method (knots)	ARCHER (knots)	ADT (knots)	SATCON (knots)	AiDT	E_1^1	E_2^2	E_3^3	E_4^4	E_5^5
06 Jun./1200	35	34.0	34.9	32.0	34	32	2.80	0.28	8.57	2.85	8.57
06 Jun./1800	35	37.8	35.0	35.0	40	33	8.00	0.00	0.00	14.2	5.71
07 Jun./0000	40	41.8	39.6	43.0	47	38	4.50	1.00	7.50	17.5	5.00
07 Jun./0600	40	48.2	44.5	43.0	50	38	20.5	11.2	7.50	25.0	5.00
07 Jun./1200	45	50.0	54.4	49.0	52	42	11.1	20.8	8.88	15.5	6.66
07 Jun./1800	50	56.0	59.6	57.0	59	52	12.0	19.2	14.0	18.0	4.00
08 Jun./0000	55	61.0	64.6	63.0	64	66	10.9	17.4	14.5	16.3	20.0
08 Jun./0600	60	67.0	69.5	69.8	71	68	11.6	15.8	16.3	18.3	13.3
08 Jun./1200	70	82.0	79.8	74.6	81	79	17.1	14.0	6.57	15.7	12.8
08 Jun./1800	75	85.0	80.0	102	88	88	13.3	6.66	36.0	17.3	17.3
09 Jun./0000	75	87.0	80.0	102	89	86	16.0	6.66	36.0	18.6	14.6
09 Jun./0600	75	89.0	84.9	94.8	90	83	18.6	13.2	26.4	20.0	10.6
09 Jun./1200	75	85.0	80.1	92.4	89	82	13.3	6.80	23.2	18.6	9.33
09 Jun./1800	80	88.2	89.8	99.6	90	92	10.2	12.2	24.5	12.5	15.0
10 Jun./0000	80	88.7	90.0	102	90	92	10.8	12.5	27.5	12.5	15.0
10 Jun./0600	80	85.3	90.0	102	95	94	6.62	12.5	27.5	18.7	17.5
10 Jun./1200	75	92.4	85.4	99.6	93	91	23.2	13.8	32.8	24.0	21.3
10 Jun./1800	70	82.6	80.1	92.4	85	84	18.0	14.4	32.0	21.4	20.0
11 Jun./0000	65	76.0	84.9	77.0	77	74	16.9	30.6	18.4	18.4	13.8
11 Jun./0600	60	70.0	85.0	65.0	69	76	16.6	41.6	8.33	15.0	26.6
11 Jun./1200	55	65.3	80.7	59.0	68	62	18.7	46.2	7.27	23.6	12.7
11 Jun./1800	55	61.9	61.3	47.0	65	60	12.5	11.4	14.5	18.1	9.09
12 Jun./0000	55	62.1	50.6	41.0	60	56	12.9	8.00	24.4	9.09	1.81
12 Jun./0600	50	52.3	45.1	33.0	55	42	4.60	9.80	34.0	10.0	16.0
Mean							12.97	14.46	19.07	16.74	14.58
RMSE		9.06	11.03	15.30	11.05	9.38					
R^2		0.61	0.42	-0.09	0.42	0.58					

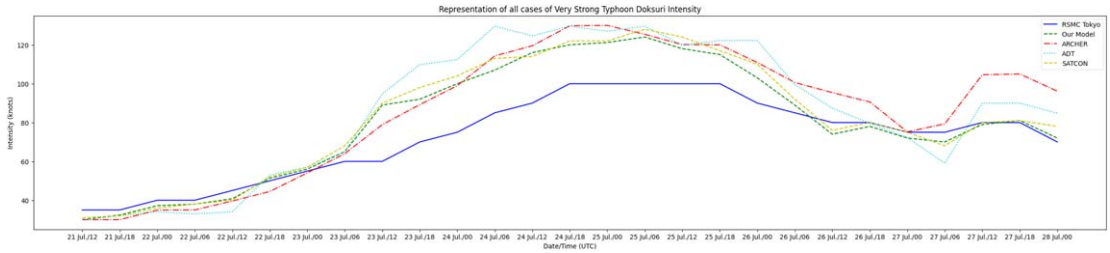


Figure 12. Graphically representation of all results of Very Strong Typhoon Doksuri.

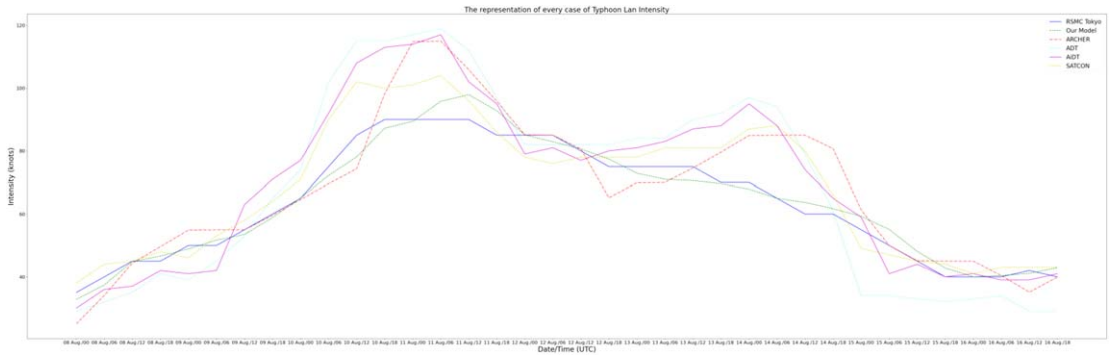


Figure 13. A graphic representation for each method's result for intensity of very strong typhoon lan.

Table 9. Numerical results of Talim.

Date/Time	RSMC Tokyo (knots)	Our Method (knots)	ARCHER (knots)	ADT (knots)	SATCON (knots)	AiDT	E_1^1	E_2^2	E_3^3	E_4^4	E_5^5
15 Jul./0600	35	34.0	35.0	33	41	36	2.85	0.00	5.71	17.1	2.85
15 Jul./1200	40	39.8	39.6	39	48	36	0.50	1.00	2.50	20.0	10.0
15 Jul./1800	45	45.6	44.6	45	52	40	1.33	0.88	0.00	15.5	11.1
16 Jul./0000	50	49.6	49.6	49	57	43	0.80	0.80	2.00	14.0	14.0
16 Jul./0600	50	52.6	54.6	45	54	43	5.20	9.20	10.0	8.00	14.0
16 Jul./1200	50	53.6	59.6	59	62	59	7.20	19.2	18.0	24.0	18.0
16 Jul./1800	55	63.2	64.6	72.2	72	71	14.9	16.3	31.2	30.9	29.0
17 Jul./0000	60	64.9	69.9	74.6	76	72	8.16	16.5	24.3	26.6	20.0
17 Jul./0600	60	68.5	70.0	94.8	86	77	14.1	16.6	58.0	43.3	28.3
17 Jul./1200	60	63.5	65.1	90.0	83	85	5.83	8.50	50.0	38.3	41.6
Mean							6.09	8.91	20.18	23.79	18.90
RMSE		4.43	6.47	16.52	14.51	12.38					
R^2		0.70	0.37	-3.06	-2.13	-1.28					

a failure to perform better than mean-based predictions. These findings underscore the efficacy of the proposed approach in capturing the underlying dynamics of typhoon intensity with higher precision and reliability.

3.6. Typhoon Guchol

Typhoon Guchol (Chedeng) in 2023 was a storm that stayed out at sea, never making landfall. It reached its peak intensity as a typhoon in the Philippine Sea but weakened before reaching the Philippines or Japan. While it brought some heavy rain to western Luzon in the Philippines due to the enhanced southwest monsoon, it caused minimal overall impact. Figure 10 shows how our proposed method for estimating Typhoon intensity is doing well as compared to other traditional methods: ARCHER, ADT, and SATCON concerning the Best track data provided by the RSMC Tokyo. Table 8 provides a quantitative summary of the results for Typhoon Guchol.

Based on Typhoon intensity data from RSMC Tokyo, with an intensity recorded at 60.62 knots, our proposed method shows an error rate of 12.97%. This is better than the error rates of other methodologies:

Table 10. Numerical results of Doksuri.

Date/Time	RSMC Tokyo (knots)	Our Method (knots)	ARCHER (knots)	ADT (knots)	SATCON (knots)	AiDT	E_1^1	E_2^2	E_3^3	E_4^4	E_5^5
21 Jul./1200	35	30.0	30.0	30	31	30	14.2	14.2	14.2	11.4	16.6
21 Jul./1800	35	32.4	30.0	32	32	34	7.42	14.2	8.57	8.57	2.94
22 Jul./0000	40	37.2	34.9	34	36	35	7.00	12.7	15.0	10.0	14.28
22 Jul./0600	40	38.0	35.0	33	38	33	5.00	12.5	17.5	5.00	21.21
22 Jul./1200	45	40.8	39.6	34	40	35	9.33	12.0	24.4	11.1	28.57
22 Jul./1800	50	51.3	44.5	53	52	41	2.60	11.0	6.00	4.00	21.95
23 Jul./0000	55	56.0	54.1	57	57	48	1.81	1.63	3.63	3.63	14.58
23 Jul./0600	60	65.0	64.0	65	68	56	8.33	6.66	8.33	13.3	7.14
23 Jul./1200	60	89.0	78.9	94.8	90	81	48.3	31.5	58.0	50.0	25.92
23 Jul./1800	70	92.0	89.2	109.8	98	94	31.4	27.4	56.8	40.0	25.53
24 Jul./0000	75	100	99.0	112.4	104	100	33.3	32.0	49.8	38.6	25.0
24 Jul./0600	75	107	114.3	129.6	113	119	25.8	34.4	52.4	32.9	28.57
24 Jul./1200	90	116	119.5	124.6	114	118	28.8	32.7	38.4	26.6	23.72
24 Jul./1800	100	120	129.8	129.6	122	128	20.0	29.8	29.6	22.0	21.87
25 Jul./0000	100	121	130.0	127.0	122	124	21.2	30.0	27.0	22.0	19.35
25 Jul./0600	100	124	125.4	129.6	128	125	24.0	25.4	29.6	28.0	20.0
25 Jul./1200	100	118	120.1	119.8	124	116	18.0	20.1	19.8	24.0	13.79
25 Jul./1800	100	115	120.0	122.2	117	116	15.0	20.2	22.2	17.0	13.79
26 Jul./0000	90	103	110.9	122.2	110	114	14.4	23.2	35.7	22.2	21.05
26 Jul./0600	85	89.0	100.6	99.6	92	104	4.70	18.3	17.1	8.23	18.26
26 Jul./1200	80	74.0	95.4	87.4	76	87	7.50	19.2	9.25	5.00	8.04
26 Jul./1800	80	78.0	90.6	79.6	80	78	2.50	13.2	0.50	0.00	2.56
27 Jul./0000	75	72.0	75.2	72.2	75	65	4.00	0.22	3.73	0.00	15.3
27 Jul./0600	75	70.0	79.3	59.0	68	57	6.66	5.73	21.3	9.33	31.5
27 Jul./1200	80	79.0	104.6	90.0	80	85	1.25	30.7	12.5	0.00	5.88
27 Jul./1800	80	81.0	105.0	90.0	81	96	1.25	31.2	12.5	1.25	16.6
28 Jul./0000	70	72.0	96.1	84.8	78	83	2.85	37.2	21.1	11.4	15.6
Mean							13.59	20.29	22.79	15.77	17.77
RMSE		14.19	18.66	21.95	16.35	17.52					
R^2		0.71	-2.02	-2.58	-0.92	-3.24					

Table 11. Numerical results of Very Strong Typhoon Lan.

Date/Time	RSMC Tokyo (knots)	Our Method (knots)	ARCHER (knots)	ADT (knots)	SATCON (knots)	AiDT (knots)	E_1^1	E_2^2	E_3^3	E_4^4	E_5^5
08 Aug./0000	35	32.8	25	29	38	30	6.28	28.5	17.1	8.57	14.2
08 Aug./0600	40	37.4	34.1	32	44	36	6.50	14.7	20.0	10.0	10.0
08 Aug./1200	45	44.9	44.4	35	45	37	0.22	1.33	22.2	0.00	17.7
08 Aug./1800	45	46.6	49.6	41	48	42	3.55	10.2	8.88	6.66	6.66
09 Aug./0000	50	48.9	54.9	39	46	41	2.20	9.80	22.0	8.00	18.0
09 Aug./0600	50	51.7	55.0	45	53	42	3.40	10.0	10.0	6.00	16.0
09 Aug./1200	55	53.5	55.0	53	58	63	2.72	0.00	3.63	5.45	14.5
09 Aug./1800	60	58.9	59.6	65	64	71	1.83	0.66	8.33	6.66	18.3
10 Aug./0000	65	65.1	64.6	74	71	77	0.15	0.61	13.8	9.23	18.4
10 Aug./0600	75	72.2	69.6	102	90	92	3.73	7.20	36.0	20.0	22.6
10 Aug./1200	85	78.1	74.3	115	102	108	8.11	12.5	35.2	20.0	27.0
10 Aug./1800	90	87.2	98.2	115	100	113	3.11	9.11	27.7	11.1	25.5
11 Aug./0000	90	89.4	114.8	117	101	114	0.66	27.5	30.0	12.2	26.6
11 Aug./0600	90	95.8	115.0	119	104	117	6.44	27.7	32.2	15.5	30.0
11 Aug./1200	90	97.9	105.9	112	96	102	8.77	17.6	24.4	6.66	13.3
11 Aug./1800	85	92.8	95.8	97	86	95	9.17	12.7	14.1	1.17	11.7
12 Aug./0000	85	85.1	85.1	82	78	79	0.11	0.11	3.52	8.23	7.05
12 Aug./0600	85	82.9	85.0	82	76	81	2.47	0.00	3.52	10.5	4.70
12 Aug./1200	80	80.6	80.6	82	78	77	0.75	0.75	2.50	2.50	3.75
12 Aug./1800	75	77.4	65.2	82	78	80	3.20	13.0	9.33	4.00	6.66
13 Aug./0000	75	72.9	69.9	84	78	81	2.80	6.80	12.0	4.00	8.00
13 Aug./0600	75	71.0	70.0	84	81	83	5.33	6.66	12.0	8.00	10.6
13 Aug./1200	75	70.6	74.6	90	81	87	5.86	0.53	20.0	8.00	16.0
13 Aug./1800	70	69.6	79.6	92	81	88	0.57	13.7	31.4	15.7	25.7
14 Aug./0000	70	97.9	84.9	97	87	95	3.00	21.2	38.5	24.2	35.7
14 Aug./0600	65	65.0	85.0	94	88	88	0.00	30.7	44.6	35.3	35.3
14 Aug./1200	60	63.7	85.0	79	80	74	6.16	41.6	31.6	33.3	23.3
14 Aug./1800	60	61.7	80.7	61	66	65	2.83	34.5	1.66	10.0	8.33
15 Aug./0000	55	59.3	61.3	34	49	59	7.81	11.4	38.1	10.9	7.27
15 Aug./0600	50	55.1	50.6	34	47	41	10.2	0.00	32.0	6.00	18.0
15 Aug./1200	45	48.1	45.1	33	45	44	6.88	0.22	26.6	0.00	2.22
15 Aug./1800	40	42.8	45.0	32	44	40	7.00	12.5	20.0	10.0	0.00
16 Aug./0000	40	40.0	45.0	33	41	41	0.00	12.5	17.5	2.50	2.50
16 Aug./0600	40	40.4	40.4	34	43	39	1.00	1.00	15.0	7.50	2.50
16 Aug./1200	42	41.1	35.1	29	43	39	2.14	16.4	30.9	2.38	7.14
16 Aug./1800	40	42.9	39.9	29	43	41	7.25	0.25	27.5	7.50	2.50
Mean							3.27	11.5	20.6	12.2	14.4
RMSE		3.27	10.6	15.8	11.6	12.4					
R^2		0.96	0.64	0.20	0.57	0.50					

ARCHER (14.46%), ADT (19.07%), and SATCON (16.74%). The RMSE values for this case are listed in table 20. Statistical analysis using a t-test indicates that our proposed method performs better than the other methods.

The R^2 values (table 21) of 0.61, 0.42, -0.09, and 0.42 provide insight into how well each model explains the variation in the data. A R^2 of 0.61 suggests a good fit, meaning the proposed method captures about 61% of the variability in the dependent variable. Similarly, a R^2 of 0.42 indicates a moderate fit, explaining 42% of the variability.

However, the negative R^2 value of -0.09 is concerning as it suggests the model performs worse than simply using the mean of the dependent variable. This could indicate that the model is not appropriate for the data or that there are fundamental issues with the model's ability to predict accurately. It's essential to interpret R^2 values in context, considering both positive and negative values to assess how well the models align with the underlying data patterns.

3.7. Severe Tropical Storm Talim

Severe Tropical Storm Talim (2023), also known as Tropical Storm Dodong in the Philippines, was a compact tropical cyclone that affected the Philippines, China, and Vietnam in July 2023. The fourth named storm of the 2023 Pacific typhoon season, Talim, originated from a weak monsoon depression east of Manila while moving towards Luzon [38].

Table 12. Numerical results of Severe Tropical Storm Damrey.

Date/Time	RSMC Tokyo (knots)	Our Method (knots)	ARCHER (knots)	ADT (knots)	SATCON (knots)	AiDT (knots)	E_1^1	E_2^2	E_3^3	E_4^4	E_5^5
25 Aug./0000	35	36.4	39.4	35	41	34	4.00	12.5	0.00	17.1	2.85
25 Aug./0600	40	38.9	44.9	33	42	35	2.75	12.2	17.5	5.00	12.5
25 Aug./1200	40	40.7	45.0	33	44	35	1.75	12.5	17.5	10.0	12.5
25 Aug./1800	40	43.8	49.9	33	45	35	9.50	24.7	17.5	12.5	12.5
26 Aug./0000	45	47.8	50.0	35	49	37	6.22	11.1	22.2	8.88	17.7
26 Aug./0600	50	48.2	50.0	43	59	40	3.60	0.00	14.0	18.0	20.0
26 Aug./1200	50	48.9	50.0	45	58	42	2.20	0.00	10.0	16.0	16.0
26 Aug./1800	50	49.3	54.9	33	51	40	1.40	9.80	34.0	2.00	20.0
27 Aug./0000	50	50.8	55.0	39	56	56	1.60	10.0	22.0	12.0	12.0
27 Aug./0600	50	51.6	59.6	39	58	58	3.20	19.2	22.0	16.0	16.0
27 Aug./1200	50	53.1	64.9	39	58	58	6.20	29.8	22.0	16.0	16.0
27 Aug./1800	45	51.5	60.1	39	56	60	14.4	33.5	13.3	24.2	33.3
28 Aug./0000	45	48.0	60.0	39	53	64	6.66	33.3	13.3	17.7	42.2
28 Aug./0600	45	46.2	45.3	37	53	62	2.66	0.66	17.7	17.7	37.7
28 Aug./1200	50	48.7	44.3	39	48	46	2.60	11.4	22.0	4.00	8.00
Mean							4.58	14.72	17.67	13.16	18.63
RMSE		2.55	8.32	9.05	6.63	9.86					
R^2		0.71	-2.02	-2.58	-0.92	-3.24					

Table 13. Numerical results of Typhoon Haikui.

Date/Time	RSMC Tokyo (knots)	Our Method (knots)	ARCHER (knots)	ADT (knots)	SATCON (knots)	AiDT (knots)	E_1^1	E_2^2	E_3^3	E_4^4	E_5^5
29 Aug./0000	40	38.9	44.2	34	44	36	2.75	10.5	15.0	10.0	10.0
29 Aug./0600	45	42.6	54.9	45	45	41	5.33	22.0	0.00	0.00	8.88
29 Aug./1200	45	44.1	50.1	47	42	46	2.00	11.3	4.44	6.66	2.22
29 Aug./1800	45	45.8	54.9	51	44	46	1.77	22.0	13.3	2.22	2.22
30 Aug./0000	45	46.9	50.4	55	48	46	4.22	12.0	22.2	6.66	2.22
30 Aug./0600	45	47.1	45.1	41	48	42	4.66	0.22	8.88	6.66	6.66
30 Aug./1200	45	48.4	49.9	51	49	45	7.55	10.8	13.3	8.88	0.00
30 Aug./1800	50	51.6	50.0	55	60	50	3.20	0.00	10.0	20.0	0.00
31 Aug./0000	55	53.5	50.0	53	66	50	2.72	9.09	3.63	20.0	9.09
31 Aug./0600	55	54.3	59.3	51	64	47	1.27	7.81	7.27	16.3	14.5
31 Aug./1200	55	56.1	64.9	57	61	62	2.00	18.0	3.63	10.9	12.7
31 Aug./1800	55	57.8	65.0	69	67	74	5.09	18.1	25.4	21.8	34.5
01 Sep./0000	60	61.1	69.6	74	74	79	1.83	16.0	23.3	23.3	31.6
01 Sep./0600	60	63.8	74.9	74	78	80	6.33	24.8	23.3	30.0	33.3
01 Sep./1200	65	66.7	75.0	77	76	80	2.61	15.3	18.4	16.9	23.0
01 Sep./1800	65	68.4	79.6	77	78	77	5.23	22.4	18.4	20.0	18.4
02 Sep./0000	70	72.4	84.9	79	80	84	3.42	21.2	12.8	14.2	20.0
02 Sep./0600	75	74.1	85.0	72	74	82	1.20	13.3	4.00	1.33	9.33
02 Sep./1200	75	76.9	80.1	82	76	79	2.53	6.80	9.33	1.33	5.33
02 Sep./1800	85	82.2	89.1	90	94	88	3.29	4.82	5.88	10.5	3.52
03 Sep./0000	80	79.8	99.4	94	112	100	0.25	24.2	17.5	40.0	25.0
03 Sep./0600	65	75.3	104.9	97	85	95	15.8	61.3	49.2	30.7	46.1
03 Sep./1200	55	68.5	81.1	94	82	90	24.5	47.4	70.9	49.0	63.6
03 Sep./1800	50	61.6	75.0	79	70	63	23.2	50.0	58.0	40.0	26.0
04 Sep./0000	45	57.6	75.0	63	58	59	28.0	66.6	40.0	28.8	31.1
04 Sep./0600	45	52.7	61.1	51	51	49	17.1	35.7	13.3	13.3	8.88
04 Sep./1200	40	45.9	51.1	47	52	41	14.7	37.7	17.5	30.0	2.50
04 Sep./1800	35	41.1	30.3	45	45	39	17.4	13.4	28.5	28.5	11.4
Mean							7.50	21.20	19.21	18.16	16.52
RMSE		5.35	14.76	13.89	12.72	13.18					
R^2		0.82	-0.32	-0.17	0.01	-0.05					

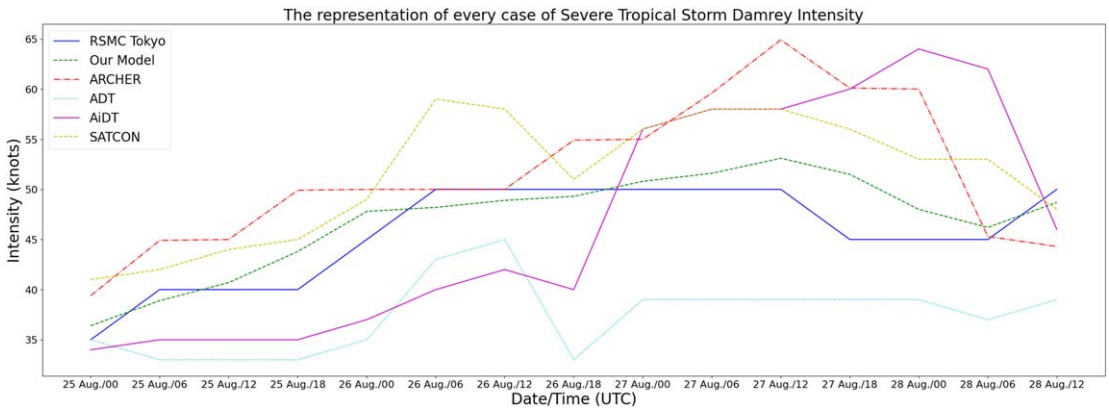


Figure 14. A graphic representation for each method's result for intensity of severe tropical storm damrey.

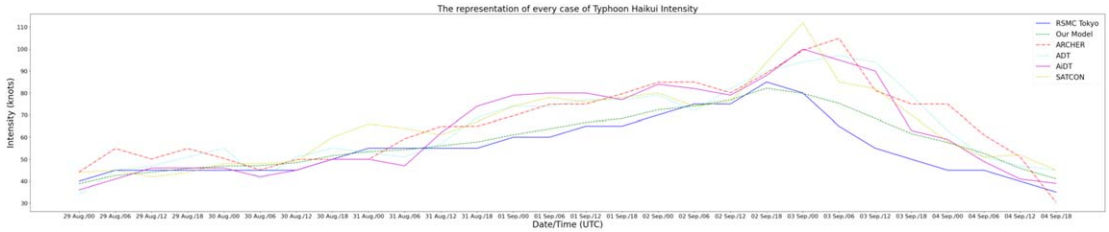


Figure 15. A graphic representation for each method's result for intensity of typhoon haikui.

The storm tracked westward near the northern edge of Luzon before emerging off the coast of Ilocos Norte. Favorable conditions allowed Talim to intensify rapidly into a Severe Tropical Storm rapidly. It maintained this intensity until shortly before landfall near Zhanjiang, Guangdong, China, on July 17th. Talim then quickly weakened and dissipated early on July 18th. Figure 11 shows how our proposed method for estimating Typhoon intensity is doing well compared to other traditional methods: ARCHER, ADT, and SATCON concerning the Best track data provided by the RSMC Tokyo. Table 9 provides a quantitative summary of Severe Tropical Storm Talim results.

Based on Typhoon intensity data from RSMC Tokyo, with an intensity recorded at 58 knots, our proposed method shows an error rate of 6.09%. This is better than the error rates of other methodologies: ARCHER (8.91%), ADT (20.18%), SATCON (23.79%), and AiDT (12.38%). The RMSE values for this case are listed in table 20. Statistical analysis using a t-test indicates that our proposed method performs better than the other methods.

The R^2 values (table 21) of 0.70, 0.37, -3.06, -2.13, and -1.28 clearly show how well each model fits the data. A R^2 of 0.70 indicates that our proposed method explains 70% of the data variability, showing it captures the underlying patterns effectively. On the other hand, a R^2 of 0.37 suggests moderate performance, explaining 37% of the variability but leaving room for improvement.

The negative R^2 values (-3.06, -2.13, and -1.28) are concerning as they imply that these models perform worse than simply using the mean of the data to predict outcomes. This indicates serious issues with these models, suggesting they are unsuitable for accurately representing the relationships within the data. These findings highlight the importance of selecting models with positive R^2 values, as they better explain and predict the observed data variability.

3.8. Very strong Typhoon doksuri

Very Strong Typhoon Doksuri, also known as Super Typhoon Egay in the Philippines. It packed a powerful punch, causing serious trouble in several places. The Philippines bore the brunt with heavy rains, floods, and strong winds, while Taiwan and China also experienced damage from fallen trees, power outages, and landslides. Figure 12 shows how our proposed method for estimating Typhoon intensity is doing well as compared to other traditional methods: ARCHER, ADT, SATCON, and AiDT concerning the Best track data provided by the RSMC Tokyo. Table 10 provides a quantitative summary of Very Strong Typhoon Doksuri results.

Based on Typhoon intensity data from RSMC Tokyo, with an intensity recorded at 72.03 knots, our proposed method shows an error rate of 13.59%. This is better than the error rates of other methodologies: ARCHER (20.29%),

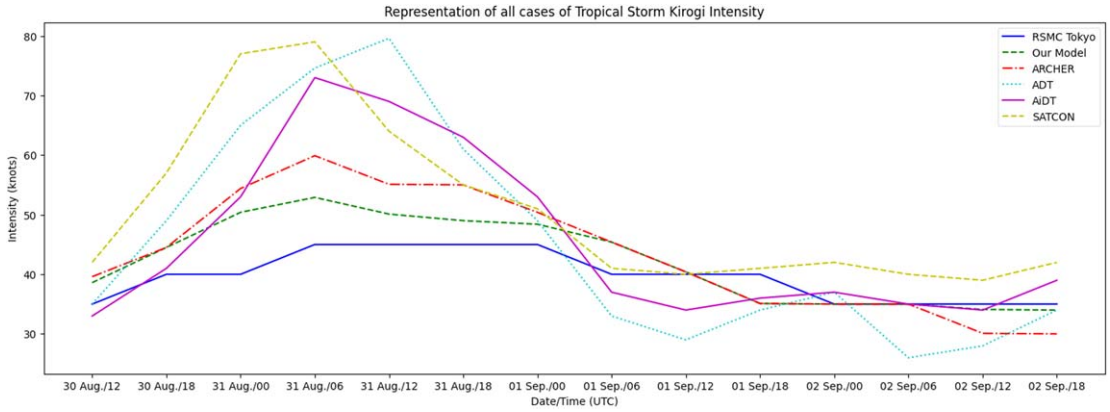


Figure 16. Graphically representation of all results of Tropical Storm Kirogi.

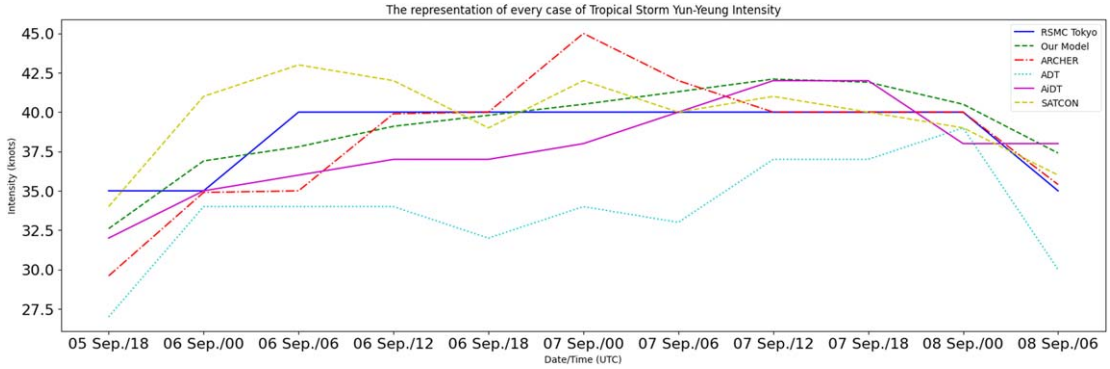


Figure 17. A graphic representation for each method's result for intensity of tropical storm yun-yeung.

Table 14. Numerical results of Kirogi.

Date/Time	RSMC Tokyo (knots)	Our Method (knots)	ARCHER (knots)	ADT (knots)	SATCON (knots)	AiDT	E_1^1	E_2^2	E_3^3	E_4^4	E_5
30 Aug./1200	35	38.6	39.6	35.0	42	33	10.2	13.1	00.0	20.0	5.71
30 Aug./1800	40	44.5	44.5	49.0	57	41	11.2	11.2	22.5	42.5	2.50
31 Aug./0000	40	50.4	54.4	65.0	77	53	26.0	36.0	62.5	92.5	32.5
31 Aug./0600	45	52.9	59.9	74.6	79	73	17.5	33.1	65.7	75.5	62.2
31 Aug./1200	45	50.1	55.1	79.6	64	69	11.3	22.4	76.8	42.2	53.3
31 Aug./1800	45	49.0	55.0	61.0	55	63	08.8	22.2	35.5	22.2	40.0
01 Sep./0000	45	48.4	50.4	49.0	51	53	07.5	12.0	08.8	13.3	17.7
01 Sep./0600	40	45.4	45.4	33.0	41	37	13.5	13.5	17.5	02.5	7.50
01 Sep./1200	40	40.4	40.4	29.0	40	34	01.0	01.0	27.5	00.0	15.0
01 Sep./1800	40	35.1	35.1	34.0	41	36	12.2	12.2	15.0	02.5	10.0
02 Sep./0000	35	35.0	35.0	37.0	42	37	00.0	00.0	05.7	20.0	5.71
02 Sep./0600	35	35.0	35.0	26.0	40	35	00.0	00.0	25.7	14.2	0.00
02 Sep./1200	35	34.1	30.1	28.0	39	34	02.5	14.0	20.0	11.4	2.85
02 Sep./1800	35	34.0	30.0	34.0	42	39	02.8	14.2	02.8	20.0	11.4
Mean							8.89	14.6	27.0	27.5	19.03
RMSE		4.72	7.57	15.81	15.56	11.96					
R^2		-0.39	-2.60	-14.60	-14.10	-7.97					

ADT (22.79%), SATCON(15.77%), and AiDT (17.77%). The RMSE values for this case are listed in table 20. Statistical analysis using a t-test indicates that our proposed method performs better than the other methods.

The R^2 values (table 21) of 0.53, 0.20, -0.10, 0.38, and 0.41 show different levels of how well the models explain the data. A R^2 of 0.53 means that our proposed method does a decent job, explaining 53% of the variability in the data. A R^2 of 0.41 and 0.38 is similar but weaker, explaining 41% and 38% of the variability.

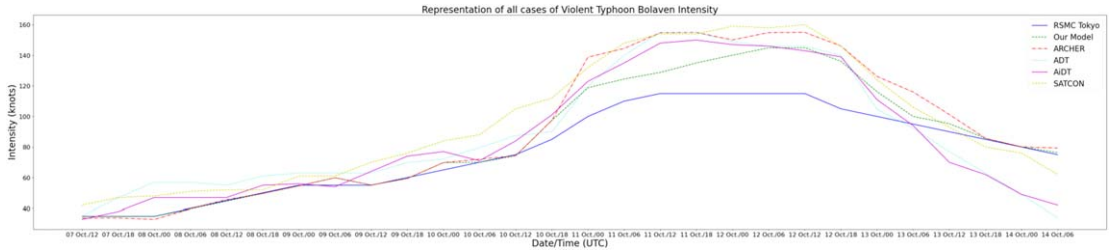


Figure 18. Graphically representation of all results of Violent Typhoon Bolaven.

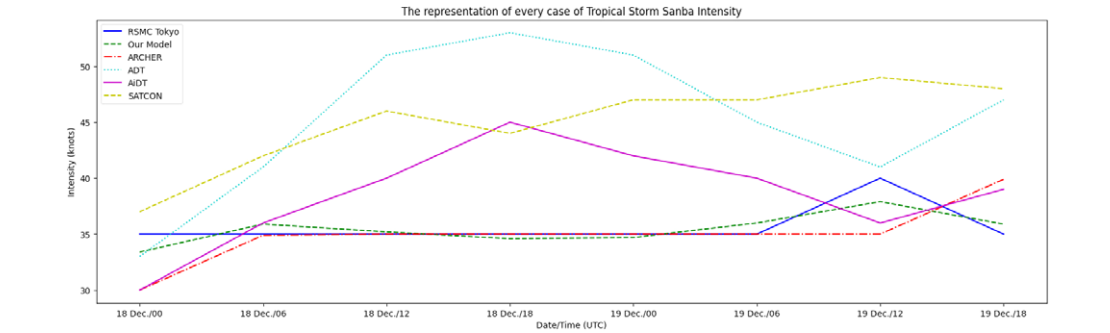


Figure 19. A graphic representation for each method's result for intensity of tropical storm sanba.

Table 15. Numerical results of Tropical Storm Yun-Yeung.

Date/Time	RSMC Tokyo (knots)	Our Method (knots)	ARCHER (knots)	ADT (knots)	SATCON (knots)	AiDT (knots)	E_1^1	E_2^2	E_3^3	E_4^4	E_5^5
05 Sep./1800	35	32.6	29.6	27	34	32	6.85	15.4	22.8	2.85	8.57
06 Sep./0000	35	36.9	34.9	34	41	35	5.42	0.28	2.85	17.1	0.00
06 Sep./0600	40	37.8	35.0	34	43	36	5.50	12.5	15.0	7.50	10.0
06 Sep./1200	40	39.1	39.9	34	42	37	2.25	0.25	15.0	5.00	7.50
06 Sep./1800	40	39.8	40.0	32	39	37	0.50	0.00	20.0	2.50	7.50
07 Sep./0000	40	40.5	45.0	34	42	38	1.25	0.12	15.0	5.00	5.00
07 Sep./0600	40	41.3	42.0	33	40	40	3.25	0.05	17.5	0.00	0.00
07 Sep./1200	40	42.1	40.0	37	41	42	5.25	0.00	7.50	2.50	5.00
07 Sep./1800	40	41.9	40.0	37	40	42	4.75	0.00	7.50	0.00	5.00
08 Sep./0000	40	40.5	40.0	39	39	38	1.25	0.00	2.50	2.50	5.00
08 Sep./0600	35	37.4	35.4	30	36	38	6.85	1.14	14.2	2.85	8.57
Mean							3.92	2.75	12.72	4.35	5.64
RMSE		1.67	2.75	5.47	2.29	2.48					
R^2		0.43	-0.52	-5.04	-0.06	-0.24					

On the other hand, a R^2 of 0.20 means the model only explains 20% of the data, indicating it doesn't capture much of the underlying trend. The R^2 of -0.10 is especially bad, suggesting the model performs worse than just using the data average, indicating it is not a good fit at all.

3.9. Very strong typhoon lan

Very Strong Typhoon Lan (2023) was a powerful tropical cyclone that struck central Japan in mid-August 2023. It caused significant damage and disruption, including flooding, landslides, and power outages. The typhoon reached its peak intensity as a typhoon with sustained winds of 90 kts. Figure 13 compares the performance of the proposed typhoon intensity estimation method with ARCHER, ADT, SATCON, and AiDT, using Best Track data from RSMC Tokyo. Table 11 summarizes Typhoon Lan results.

The proposed method achieves a 3.95% error, outperforming ARCHER (11.5%), ADT (20.6%), SATCON (12.2%), and AiDT (14.4%). Table 20 highlights lower RMSE values, confirming its accuracy. A t-test statistically validates its significant improvement over other methods. Table 21 presents R^2 values: 0.96 (proposed method),

0.64 (ARCHER), 0.57 (SATCON), 0.20 (ADT), and 0.50 (AiDT). The proposed method, with 0.96 R^2, explains 96% of intensity variance, indicating a strong fit. In contrast, ARCHER, ADT, AiDT, and SATCON show weak fits and poor performance. These results emphasize the proposed method's superior precision and reliability in modeling typhoon intensity.

3.10. Severe tropical storm damrey

Severe Tropical Storm Damrey (2023) formed in the western Pacific Ocean in August 2023. It intensified into a Severe Tropical Storm before weakening again. Damrey did not make landfall and dissipated over open waters. Figure 14 compares the performance of the proposed typhoon intensity estimation method with ARCHER, ADT, SATCON, and AiDT, using Best Track data from RSMC Tokyo. Table 12 summarizes Severe Tropical Storm Damrey results.

The proposed method achieves a 4.58% error, outperforming ARCHER (14.72%), ADT (17.67%), SATCON (13.16%), and AiDT (18.63%). Table 20 highlights lower RMSE values, confirming its accuracy. A t-test statistically validates its significant improvement over other methods. Table 21 presents R^2 values: 0.71 (proposed method), -2.02 (ARCHER), -0.92 (SATCON), -2.58 (ADT), and -3.24 (AiDT). The proposed method, with 0.71 R^2, explains 71% of intensity variance, indicating a strong fit. In contrast, ARCHER, ADT, AiDT, and SATCON show weak fits and poor performance. These results emphasize the proposed method's superior precision and reliability in modeling typhoon intensity.

3.11. Typhoon haikui

Typhoon Haikui (2023) was a significant typhoon that impacted Taiwan and mainland China in September 2023. It brought strong winds, heavy rain, and flooding to both regions. The typhoon caused widespread damage, including power outages, downed trees, and landslides. Figure 15 compares the performance of our proposed typhoon intensity estimation method with ARCHER, ADT, SATCON, and AiDT, using Best Track data from RSMC Tokyo. Results for Typhoon Haikui are summarized in table 13.

The proposed method achieves a percentage error of 7.50%, outperforming ARCHER (21.20%), ADT (19.21%), SATCON (18.16%), and AiDT (16.52%). RMSE values (table 20) confirm this accuracy improvement. A t-test validates the statistical significance of the proposed method's advantage over traditional approaches.

R^2 values (table 21) highlight model performance: 0.82 (proposed method), -0.32 (ARCHER), 0.01 (SATCON), -0.17 (ADT), and -0.05 (AiDT). An R^2 of 0.82 indicates the proposed method explains 82% of intensity variance, showing a strong fit. In contrast, SATCON explains only 1%, and ARCHER, ADT, and AiDT show negative R^2, reflecting poor performance compared to mean-based predictions. These results emphasize the proposed method's precision and reliability in modeling typhoon intensity.

3.12. Tropical Strom Kirogi

Tropical Storm Kirogi (2023) formed far east of Guam in the West Pacific Ocean and moved north but remained over water, never making landfall. It reached peak intensity on August 30th with winds of 45 knots before dissipating on September 2*nd*. Figure 16 shows the estimated Typhoon intensity calculated by the proposed method; traditional methodologies like ARCHER, ADT, SATCON, and Best Track data from RSMC Tokyo. Table 14 provides a quantitative summary of the results of Tropical Storm Kirogi.

When evaluating the mean error based on Typhoon intensity data from RSMC, where the intensity is recorded at 39.6 knots, our proposed method shows an error rate of 8.89%. This is better than the error rates of other methodologies: ARCHER (14.6%), ADT (27.0%), SATCON (27.5%), and AiDT (19.03%). The RMSE values for this case are listed in table 20. Statistical analysis using a t-test indicates that our proposed method performs better than the other methods.

3.13. Tropical storm Yun-Yeung

Tropical Storm Yun-Yeung, also known as Tropical Storm Ineng in the Philippines, was a weak tropical storm that affected Japan in early September 2023. It brought heavy rainfall and strong winds to parts of Japan, leading to localized flooding and landslides. Figure 17 illustrates a performance comparison between the proposed typhoon intensity estimation method and existing approaches: ARCHER, ADT, SATCON, and AiDT—using Best Track data from RSMC Tokyo. Key findings for Tropical Storm Yun-Yeung are outlined in table 15.

With a percentage error of 1.92%, the proposed method significantly outperforms ARCHER (2.69%), ADT (12.72%), SATCON (4.35%), and AiDT (5.64%). This improvement is further supported by RMSE values presented in table 20. Statistical analysis via a t-test confirms the proposed method's superior accuracy compared to traditional models. Table 21 reports R^2 values: 0.43 (proposed method), 0.03 (ARCHER), -0.06 (SATCON), -5.04 (ADT), and -0.24 (AiDT). An R^2 of 0.43 demonstrates the proposed method accounts for 43% of intensity variance, indicating a robust model fit. In contrast, ARCHER captures only 3%, while SATCON's

Table 16. Numerical results of Bolaven.

Date/Time	RSMC Tokyo (knots)	Our Method (knots)	ARCHER (knots)	ADT (knots)	SATCON (knots)	AiDT	E_1^1	E_2^2	E_3^3	E_4^4	E_5^5
07 Oct./1200	35	33.9	34.9	35.0	42	33	0.28	3.14	0.00	20.0	5.71
07 Oct./1800	35	34.0	35.0	47.0	47	38	00.0	2.85	34.2	34.2	8.57
08 Oct./0000	35	33.0	35.0	57.0	48	47	00.0	5.71	62.8	37.1	34.28
08 Oct./0600	40	39.8	39.6	57.0	51	47	01.0	0.50	42.5	27.5	17.5
08 Oct./1200	45	45.6	44.6	55.0	52	47	0.88	1.33	22.2	15.5	4.44
08 Oct./1800	50	49.6	49.6	61.0	52	55	0.80	0.80	22.0	4.00	10.0
09 Oct./0000	55	54.6	54.6	63.0	61	56	0.72	0.72	94.5	10.9	1.81
09 Oct./0600	55	59.9	59.9	63.0	61	54	8.90	8.90	14.5	10.9	1.81
09 Oct./1200	55	55.1	55.1	63.0	70	64	0.18	0.18	14.5	27.2	16.36
09 Oct./1800	60	59.4	59.4	69.8	76	74	1.00	1.08	16.3	26.6	23.33
10 Oct./0000	65	69.8	69.8	72.2	84	77	7.38	7.38	11.0	29.2	18.46
10 Oct./0600	70	72.0	70.0	79.6	88	71	0.00	2.85	13.7	25.7	1.42
10 Oct./1200	75	74.3	74.3	87.4	105	84	0.93	0.93	16.5	40.0	12.0
10 Oct./1800	85	97.4	97.4	90.0	112	101	14.5	14.5	5.88	31.7	18.8
11 Oct./0000	100	118.8	138.8	119.8	132	123	18.8	38.8	19.8	32.0	23.0
11 Oct./0600	110	124.5	144.5	140.0	148	135	13.1	31.3	27.2	34.5	22.7
11 Oct./1200	115	128.8	154.8	155.0	154	148	12.0	34.6	34.7	33.9	28.6
11 Oct./1800	115	135.0	155.0	155.0	154	150	17.3	34.7	34.7	33.9	30.4
12 Oct./0000	115	140.1	150.1	149.0	159	147	21.8	30.5	29.5	38.2	27.8
12 Oct./0600	115	144.9	154.9	146.0	158	146	26.0	34.6	26.9	37.3	26.9
12 Oct./1200	115	145.0	155.0	146.0	160	143	26.0	34.7	26.9	39.1	24.3
12 Oct./1800	105	136.1	146.1	140.0	146	139	29.6	39.1	33.3	39.0	32.3
13 Oct./0000	100	116.2	126.2	104.6	124	111	16.2	26.2	4.60	24.0	11.0
13 Oct./0600	95	100.0	116.0	92.4	106	94	5.26	22.1	2.73	11.5	1.05
13 Oct./1200	90	95.3	101.3	77.0	93	70	5.88	12.5	14.4	3.33	22.2
13 Oct./1800	85	85.7	85.7	63.0	80	62	0.82	0.82	25.8	5.88	27.0
14 Oct./0000	80	80.1	80.1	49.0	76	49	0.12	0.12	38.7	5.00	38.7
14 Oct./0600	75	76.3	79.3	34.0	62	42	1.73	5.73	54.6	17.3	44.0
Mean							8.27	14.18	26.62	24.86	19.10
RMSE		13.45	21.94	24.28	24.87	20.25					
R^2		0.76	0.36	0.21	0.18	0.45					

Table 17. Numerical results of Tropical Storm Sanba.

Date/Time	RSMC Tokyo (knots)	Our Method (knots)	ARCHER (knots)	ADT (knots)	SATCON (knots)	AiDT	E_1^1	E_2^2	E_3^3	E_4^4	E_5^5
18 Dec./0000	35	33.4	30.0	33	37	30	4.57	14.2	5.71	5.71	14.2
18 Dec./0600	35	35.9	34.9	41	42	36	2.57	0.28	17.1	20.0	2.85
18 Dec./1200	35	35.2	35.0	51	46	40	0.57	0.00	45.7	31.4	14.2
18 Dec./1800	35	34.6	35.0	53	44	45	1.14	0.00	51.4	25.7	28.5
19 Dec./0000	35	34.7	35.0	51	47	42	0.85	0.00	45.7	34.2	20.0
19 Dec./0600	35	36.0	35.0	45	47	40	2.85	0.00	28.5	34.2	14.2
19 Dec./1200	40	37.9	35.0	41	49	36	5.25	12.5	2.50	22.5	10.0
19 Dec./1800	35	35.9	39.9	47	48	39	2.57	14.0	34.2	37.1	11.4
Mean							2.54	5.13	28.88	26.38	14.46
RMSE		1.11	3.04	11.83	9.95	5.66					
R^2		0.54	-2.38	-50.24	-35.25	-10.74					

Table 18. Comparison between the performance of our method and the CNN model by Yadav *et al* (2024).

Name of Typhoon	Our Method			Yadav *et al* (2024)		
	Error percentage	RMSE	R^2	Error percentage	RMSE	R^2
Ewiniar	4.43	3.07	0.92	14.87	12.76	0.10
Gaemi	5.89	4.55	0.91	22.45	20.21	-0.7
Prapiron	7.82	4.41	0.70	25.87	11.02	0.20
Sanvu	5.90	2.63	0.41	35.63	07.01	-0.1
Mawar	5.75	6.96	0.90	20.10	28.37	0.42
Guchol	12.97	9.06	0.61	19.87	19.73	0.10
Talim	6.09	4.43	0.70	15.78	10.65	-0.8
Doksuri	13.59	14.19	0.53	29.87	24.23	-1.2
Lan	3.95	3.27	0.96	17.80	09.97	0.61
Damrey	4.58	2.55	0.71	08.97	07.87	0.27
Haikui	7.50	5.35	0.82	15.52	12.57	0.51
Kirogi	8.89	4.72	-0.3	20.09	14.42	-3.9
Yun-Yeung	3.92	1.67	0.43	10.72	05.75	-0.9
Bolaven	8.27	13.45	0.76	28.97	26.89	0.17
Sanba	2.54	1.11	0.54	4.89	09.12	-20.4

Table 19. Detail of error percentage of all cases.

Name of Typhoon	Our Method w.r.t. Best Track	ARCHER w.r.t. Best Track	ADT w.r.t Best Track	SATCON w.r.t. Best Track	AiDT w.r.t. Best Track
Ewiniar	4.43	13.97	19.58	12.25	15.93
Gaemi	5.89	21.15	19.43	19.36	19.16
Prapiron	7.82	23.55	26.85	23.17	12.97
Sanvu	5.90	11.06	33.47	22.80	9.69
Mawar	5.75	18.79	24.86	23.02	20.17
Guchol	12.97	14.46	19.07	16.74	12.58
Talim	6.09	8.91	20.18	23.79	18.90
Doksuri	13.59	20.29	22.79	15.77	17.77
Lan	3.95	11.5	20.6	12.2	14.4
Damrey	4.58	14.72	17.67	13.16	18.63
Haikui	7.50	21.20	19.21	18.16	16.52
Kirogi	8.89	14.6	27.0	27.5	19.03
Yun-Yeung	3.92	4.28	12.72	4.35	5.64
Bolaven	8.27	14.18	26.62	24.86	19.10
Sanba	2.54	5.13	28.88	26.38	14.46

negative R^2 of -00.06 reflects poor predictive capability. These results underline the proposed approach's effectiveness in capturing typhoon intensity dynamics with enhanced precision and reliability.

3.14. Violent typhoon bolaven

Bolaven is the powerful Typhoon that formed near the Federated States of Micronesia on October 6th in the Western Pacific Ocean. It moved northwest, impacting areas like Micronesia and the Northern Mariana Islands with heavy rain and strong winds. It is absorbed by another weather system, with remnants bringing rain and wind to parts of Alaska and British Columbia. Figure 18 shows how our proposed method for estimating Typhoon intensity is doing well compared to other traditional methods: ARCHER, ADT, and SATCON concerning the Best track data provided by the RSMC Tokyo. Table 16 provides a quantitative summary of the results for Violent Typhoon Bolaven.

Based on Typhoon intensity data from RSMC Tokyo, with an intensity recorded at 87 knots, our proposed method shows an error rate of 8.27%. This is better than the error rates of other methodologies: ARCHER (14.18%), ADT (26.62%), SATCON (24.86%), and AiDT (19.10%). The RMSE values for this case are listed in table 20. Statistical analysis using a t-test indicates that our proposed method performs better than the other methods.

Our proposed method has the value of R^2 0.76 (table 21), indicating that our proposed method captures most of the data's underlying patterns. On the other hand, other methods have values of 0.36, 0.21, 0.18, and 0.45, which show weaker fits. These lower R^2 values suggest that these models do not capture much of the variability in the data, indicating a need for improvement or alternative approaches to explain the data better.

Table 20. Detail of root mean square error of all cases.

Name of Typhoon	Our Method w.r.t. Best Track	ARCHER w.r.t. Best Track	ADT w.r.t Best Track	SATCON w.r.t. Best Track	AiDT w.r.t. Best Track
Ewiniar	3.07	10.17	12.91	8.65	10.68
Gaemi	4.55	19.76	17.50	19.22	16.21
Prapiron	4.41	12.32	15.12	10.85	8.14
Sanvu	2.63	4.59	13.41	9.34	4.28
Mawar	6.96	22.24	22.12	25.15	18.84
Guchol	9.06	11.03	15.30	11.05	9.38
Talim	4.43	6.47	16.52	14.51	12.38
Doksuri	14.19	18.66	21.95	16.35	17.52
Lan	3.27	10.6	15.8	11.6	12.4
Damrey	2.55	8.32	9.05	6.63	9.86
Haikui	5.35	14.76	13.89	12.72	13.18
Kirogi	4.72	7.57	15.81	15.56	11.96
Yun-Yeung	1.67	2.75	5.47	2.29	2.48
Bolaven	13.45	21.94	24.28	24.87	20.25
Sanba	1.11	3.04	11.83	9.95	5.66

Table 21. Detail of R-squared of all cases.

Name of Typhoon	Our Method w.r.t. Best Track	ARCHER w.r.t. Best Track	ADT w.r.t Best Track	SATCON w.r.t. Best Track	AiDT w.r.t. Best Track
Ewiniar	0.92	0.21	-0.26	0.43	0.13
Gaemi	0.91	-0.63	-0.28	-0.54	-0.09
Prapiron	0.70	-1.33	-2.51	-0.81	-0.02
Sanvu	0.41	-0.78	-14.23	-6.39	-0.55
Mawar	0.90	0.05	0.06	-0.21	0.31
Guchol	0.61	0.42	-0.09	0.42	0.58
Talim	0.70	0.37	-3.06	-2.13	-1.28
Doksuri	0.53	0.20	-0.10	0.38	0.41
Lan	0.96	0.64	0.20	0.57	0.50
Damrey	0.71	-2.02	-2.58	-0.92	-3.24
Haikui	0.82	-0.32	-0.17	0.01	-0.05
Kirogi	-0.39	-2.60	-14.6	-14.1	-7.97
Yun-Yeung	0.43	-0.52	-5.04	-0.06	-0.24
Bolaven	0.76	0.36	0.21	0.18	0.45
Sanba	0.54	-2.38	-50.24	-35.25	-10.74

3.15. Tropical Storm Sanba

Tropical Storm Sanba (2023) was a weak tropical storm that formed in the South China Sea in October 2023. It made landfall in Hainan, China, and brought heavy rainfall to the region. Sanba caused some flooding and disruptions, but its impact was relatively minor compared to other storms in the 2023 Pacific typhoon season. Figure 19 compares the proposed typhoon intensity estimation method with ARCHER, ADT, SATCON, and AiDT, based on Best Track data from RSMC Tokyo. Results for Tropical Storm Sanba are summarized in table 17.

The proposed method shows a percentage error of 2.54%, much lower than ARCHER (5.13%), ADT (28.88%), SATCON (26.38%), and AiDT (14.46%). RMSE values in table 20 confirm its improved accuracy, supported by a t-test analysis.

R^2 values in table 21 show strong performance: 0.54 for the proposed method, compared to -2.38 (ARCHER), -35.25 (SATCON), -50.24 (ADT), and -10.74 (AiDT). An R^2 of 0.54 means the proposed method explains 54% of intensity variation, while other methods show negative R^2 shows poor performance. These results highlight the method's precision and reliability in estimating typhoon intensity.

4. Conclusion

In this paper, we develop a NN-based model to estimate Typhoon intensity, primarily focusing on the West Pacific Ocean. Our proposed method performs better than traditional methodologies, including ARCHER, ADT, and SATCON, as well as the artificial intelligence-based approach AiDT, achieving more accurate estimations of typhoon intensity.

Additionally, we evaluate the performance (table 18) of our NN model against the recently introduced convolutional neural network (CNN) model by Yadav *et al* (2024) [4], and our approach achieves improved accuracy in estimating typhoon intensity.

We substantiate the effectiveness of our proposed method through four statistical tests: the t-test, error percentage, R-squared (R^2), and root mean square error (RMSE). These tests confirm the proposed method's reduced error percentage (as shown in table 19), lower RMSE (as detailed in table 20), R^2 (as shown in table 21) says our proposed method is a good fit and captures a substantial portion of the underlying variability in the data as compared to traditional individual methodologies, thereby validating its improved accuracy and reliability over existing methods.

The implications of our proposed method are significant when estimating various aspects of Typhoons, such as the radius of maximum winds and rapid intensification. These aspects are crucial for understanding the potential impact and behavior of TCs. Our proposed method's design and capabilities remain consistent across different applications, ensuring its robustness and adaptability. The only necessary adjustments in certain scenarios would involve updating the training data and output classifications to tailor the proposed method to specific Typhoon cases.

Accurate estimation of Typhoon intensity enhances our understanding of other Typhoon characteristics and plays a vital role in disaster preparedness and mitigation. By providing more reliable intensity estimates, our proposed method can help authorities and communities minimize the loss of life and property due to Typhoons.

Acknowledgments

The best track data for Typhoons are freely available thanks to the RSMC over the West Pacific Basin for providing the information on their website. The authors thank the editor and reviewers for their valuable input.

Conflict of interest

None of the authors declare a conflict of interest.

Data availability statement

The data cannot be made publicly available upon publication because no suitable repository exists for hosting data in this field of study. The data that support the findings of this study are available upon reasonable request from the authors.

Declarations

The Declarations contains the author statement and detail regarding the funding.

Author statement

Monu Yadav: Research, data curation, formal analysis, investigation, and first draft preparation. Laxminarayan Das: Supervision, Monitoring, Review, and Editing.

Funding

None of the authors received any funding for this research work.

ORCID iDs

Monu Yadav https://orcid.org/0000-0002-1271-0443
Laxminarayan Das https://orcid.org/0000-0002-4072-9766

References

[1] Wu L and Wang B 2004 Assessing impacts of global warming on tropical cyclone tracks *J. Clim.* **17** 1686–98
[2] Wang Y-q and Wu C-C 2004 Current understanding of tropical cyclone structure and intensity changes-a review *Meteorol. Atmos. Phys.* **87** 257–78

[3] Zhang G, Perrie W, Zhang B, Yang J and He Y 2020 Monitoring of tropical cyclone structures in ten years of radarsat-2 sar images *Remote Sens. Environ.* **236** 111449
[4] Yadav M, Das L and Kant S 2024 Estimation of tropical cyclone's radius of maximum wind using ensemble machine learning approach *J. Earth Syst. Sci.* **133** 239
[5] Yadav M and Das L 2024 Formulation and evaluation of the radius of maximum wind of the tropical cyclones over the north indian ocean basin *Theor. Appl. Climatol.* 1–14
[6] Velden C *et al* 2006 The dvorak tropical cyclone intensity estimation technique: a satellite-based method that has endured for over 30 years *Bull. Am. Meteorol. Soc.* **87** 1195–210
[7] Velden C S, Olander T L and Zehr R M 1998 Development of an objective scheme to estimate tropical cyclone intensity from digital geostationary satellite infrared imagery *Weather and Forecasting* **13** 172–86
[8] Olander T L, Velden C and Kossin J 2004 The advanced objective dvorak technique (aodt)-latest upgrades and future directions *26th Conference on Hurricanes and Tropical Meteorology* 294–5
[9] Wimmers A J and Velden C S 2010 Objectively determining the rotational center of tropical cyclones in passive microwave satellite imagery *J. Appl. Meteorology and Climatology* **49** 2013–34
[10] Kidder S Q, Goldberg M D, Zehr R M, DeMaria M, Purdom J F, Velden C S, Grody N C and Kusselson S J 2000 Satellite analysis of tropical cyclones using the advanced microwave sounding unit (amsu) *Bull. Am. Meteorol. Soc.* **81** 1241–60
[11] Velden C S and Herndon D 2020 A consensus approach for estimating tropical cyclone intensity from meteorological satellites: satcon *Weather and Forecasting* **35** 1645–62
[12] Ahmed R, Mohapatra M, Dwivedi S, Giri R K and Kant S 2022 An overview of the satellite consensus (satcon) algorithm to estimate tropical cyclone intensity over the north indian ocean *J. Earth Syst. Sci.* **131** 169
[13] Yadav M and Das L 2024 Analyze the satcon algorithm's capability to estimate tropical storm intensity across the west pacific basin *J. Earth Syst. Sci.* **133** 79
[14] Goodfellow I, Bengio Y and Courville A 2016 *Deep Learning* (MIT Press)
[15] Pradhan R, Aygun R S, Maskey M, Ramachandran R and Cecil D J 2017 Tropical cyclone intensity estimation using a deep convolutional neural network *IEEE Trans. Image Process.* **27** 692–702
[16] Chen B, Chen B-F and Lin H-T 2018 Rotation-blended CNNs on a new open dataset for tropical cyclone image-to-intensity regression *Proceedings of the 24th ACM SIGKDD International Conf. on Knowledge Discovery & Data Mining* 90–9
[17] Combinido J S, Mendoza J R and Aborot J 2018 A convolutional neural network approach for estimating tropical cyclone intensity using satellite-based infrared images *2018 24th International Conference on Pattern Recognition (ICPR), IEEE* 1474–80
[18] Lee J, Im J, Cha D-H, Park H and Sim S 2019 Tropical cyclone intensity estimation using multi-dimensional convolutional neural networks from geostationary satellite data *Remote Sensing* **12** 108
[19] Chen B-F, Chen B, Lin H-T and Elsberry R L 2019 Estimating tropical cyclone intensity by satellite imagery utilizing convolutional neural networks *Weather and Forecasting* **34** 447–65
[20] Higa M, Tanahara S, Adachi Y, Ishiki N, Nakama S, Yamada H, Ito K, Kitamoto A and Miyata R 2021 Domain knowledge integration into deep learning for typhoon intensity classification *Sci. Rep.* **11** 12972
[21] Chen Z and Yu X 2020 A novel tensor network for tropical cyclone intensity estimation *IEEE Trans. Geosci. Remote Sens.* **59** 3226–43
[22] Zhang C-J, Wang X-J, Ma L-M and Lu X-Q 2021 Tropical cyclone intensity classification and estimation using infrared satellite images with deep learning *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **14** 2070–86
[23] Varalakshmi P, Vasumathi N and Venkatesan R 2023 Tropical cyclone intensity prediction based on hybrid learning techniques *J. Earth Syst. Sci.* **132** 28
[24] Griffin S M, Wimmers A and Velden C S 2024 Predicting short-term intensity change in tropical cyclones using a convolutional neural network *Weather and Forecasting* **39** 177–202
[25] Raynaud L, Faure G, Puig M, Dauvilliers C, Trosino J-N and B'ejean P 2024 A convolutional neural network for tropical cyclone wind structure identification in kilometer-scale forecasts *Artificial Intelligence for the Earth Systems* **3** 230059
[26] Velden C and Herndon D 2019 An update on the cimss satellite consensus (satcon) tropical cyclone intensity algorithm *2019 Joint Satellite Conference* AMS
[27] Olander T L and Velden C S 2007 The advanced dvorak technique: Continued development of an objective scheme to estimate tropical cyclone intensity using geostationary infrared satellite imagery *Weather and Forecasting* **22** 287–98
[28] Herndon D, Velden C, Hawkins J, Olander T and Wimmers A 2010 The cimss satellite consensus (satcon) tropical cyclone intensity algorithm *29th Conference on Hurricanes and Tropical Meteorology* D **4**
[29] Velden C and Herndon D 2014 Update on the satellite consensus (satcon) algorithm for estimating tc intensity; 31st conf. on hurricanes and tropical meteorology, sandiego *ca. Amer. Meteor. Soc.* **30**
[30] Goyal S, Mohapatra M, Kumari P, Dube S and Rajendra K 2017 Validation of advanced dvorak technique (ADT) over north indian ocean *Mausam* **68** 689–98
[31] Pedregosa F *et al* 2011 Scikit-learn: machine learning in python *Journal of Machine Learning Research* **12** 2825–30
[32] Shimada S, Pannattee P, Ikei Y, Nishiuchi N and Yem V 2023 High-frequency cybersickness prediction using deep learning techniques with eye-related indices *IEEE Access* **11** 95825–39
[33] Pohjankukka J, Pahikkala T, Nevalainen P and Heikkonen J 2017 Estimating the prediction performance of spatial models via spatial k-fold cross validation *Int. J. Geogr. Inf. Sci.* **31** 2001–19
[34] Nti I K *et al* 2021 Performance of machine learning algorithms with different k values in k-fold cross-validation *International Journal of Information Technology and Computer Science* **13** 61–71
[35] Chollet F *et al* Keras (https://github.com/fchollet/Keras)
[36] Kingma D P and Ba J 2014 *Adam: a Method for Stochastic Optimization* arXiv:1412.6980
[37] Olander T, Wimmers A, Velden C and Kossin J P 2021 Investigation of machine learning using satellite-based advanced dvorak technique analysis parameters to estimate tropical cyclone intensity *Weather and Forecasting* **36** 2161–86
[38] (Wikipedia) 2023 *Tropical Storm Talim* Accesed on 12 June, 2024 (https://en.wikipedia.org/wiki?curid=74354063)

JMEPEG
https://doi.org/10.1007/s11665-025-11185-7

©ASM International
1059-9495/$19.00

ORIGINAL RESEARCH ARTICLE

Evolution of Microstructure, Mechanical Properties, and Tribological Behavior of AA2024/B_4C/Al_2O_3 Nanocomposites Fabricated by Multi-pass FSP

Md Saquib Bin Reyaz, Joginder Singh, Shallender Singh Bhati, Husain Mehdi, and Subhash Mishra

Submitted: 28 November 2024 / Revised: 23 February 2025 / Accepted: 8 March 2025

The objective of this study is to enhance the microstructure, mechanical properties, and tribological performance of a 6 mm-thick AA2024 alloy through the incorporation of B_4C and Al_2O_3 nanoparticles using multi-pass friction stir processing (MPFSP). The hybrid composites are fabricated using six FSP passes at a constant spindle speed of 800 rpm and a traverse speed of 35 mm/min. The research assessed composite performance using metallography, tensile testing, hardness evaluation, wear analysis, and fractographic studies. The microstructural findings indicated that MPFSP facilitated a consistent reduction in the mean grain size, decreasing from 7.18 ± 0.26 μm after the initial pass to 2.82 ± 0.06 μm following six passes. Electron backscatter diffraction results indicated that the fraction of high-angle grain boundaries rose from 76.49 to 95.21% after one to six passes. The ultimate tensile strength of the composite rises from 469 to 507 MPa, while the yield strength increases from 327 to 360 MPa. However, the elongation decreases from 6.4 to 3.1% when subjected to one pass to six passes. The hardness of the composite increases from 155 to 189 HV, while the wear depth decreases from 83 to 59 μm when subjected to one pass to six passes. The enhanced mechanical and tribological performance of the composite is linked to grain refinement, consistent reinforcement distribution, and improved grain boundary transformations following six passes of FSP. The tensile fractured surfaces exhibited dominant ductile mode failure and the worn surface morphology showed abrasive wear, oxidation, and delamination as the primary wear mechanism caused due to thermal effects and repeated sliding cycles.

Keywords AA2024-T351, dynamic recrystallization, MPFSP, severe plastic deformation, hardness, UTS, wear

1. Introduction

One of the key challenges in materials engineering today is developing materials that can meet the demands of sustainable transportation. With the push for lightweight yet durable materials, aluminum matrix composites (AMCs) have garnered significant attention due to their combination of a light matrix and high strength, derived from the inclusion of ceramic particles, such as silicon carbide (SiC), aluminum oxide (Al_2O_3), and boron carbide (B_4C) (Ref 1, 2). These ceramic particles enhance the properties of AMCs through mechanisms such as Orowan strengthening, pinning effect, and grain refinement, which collectively hinder dislocation movement and improve the overall strength and hardness of the composite (Ref 3). The AMCs have potential applications in structural components, especially where low weight and high strength are critical. Recently, research has also been directed toward using these composites as substitutes for aluminum molds in polymer injection processes, aiming to extend the molds' operational lifespan (Ref 4). Despite these benefits, AMCs possesses some challenges in achieving uniform reinforcement distribution, interfacial reactions between matrix and reinforcements, porosity, limited ductility, and high manufacturing costs. To address these issues, researchers have explored innovative approaches (Ref 5, 6). A concept gaining traction over the past decade is the selective reinforcement of the component surface. This approach allows for the strategic placement of reinforcing particles on specific surface regions that are subject to wear, abrasion, and fatigue, without compromising the overall structure. Rather than uniformly dispersing particles throughout the material, the particles are added to the surface after the component has been initially fabricated. This targeted reinforcement can provide durability where it is most needed, while keeping the production cost and weight lower compared to traditional, fully reinforced composites (Ref 7).

Friction stir processing (FSP) has emerged as a promising method for achieving this localized surface reinforcement (Ref 8). FSP allows micro- and nanoparticles to be incorporated into the surface layer of metal matrices, creating what are known as

Md Saquib Bin Reyaz, Department of Mechanical Engineering, National Institute of Technology Patna, Patna, Bihar 800005, India; **Joginder Singh**, Mechanical Engineering Department, Manav Rachna University, Faridabad, Haryana, India; **Shallender Singh Bhati**, Mechanical Engineering Department, Meerut Institute of Engineering and Technology, Meerut, India; **Husain Mehdi**, Mechanical Engineering Department, Meerut Institute of Engineering and Technology, Meerut, India; and Department of Mechanical Engineering, Centre of Research Impact and Outcome, Chitkara University, Rajpura, Punjab, India; and and **Subhash Mishra**, Mechanical Engineering Department, IMS Engineering College Ghaziabad, Ghaziabad, India. Contact e-mail: husainmehdi4u@gmail.com.

Published online: 21 April 2025

surface metal-matrix composites. This process involves using a specially designed FSP tool to stir the particles into a shallow groove cut on the surface of the component, distributing the reinforcing particles in a concentrated area (Ref 9, 10). The FSP technique allows for the creation of a thin surface composite layer, typically several mm deep, on a wrought Al-alloy. This layer is subsequently hardened through heat treatment, providing enhanced durability specifically in high-wear areas (Ref 11). During FSP, the stirring action and localized heating can lead to the dissolution or coarsening of the strengthening precipitates within the Al-alloy matrix, which can reduce hardness in the stirred region. The tribological properties, such as wear and erosion, are influenced by several structural factors, including atomic-scale surface adhesion, nanoscale dislocations, and microscale grain size or secondary phase dimensions. As a result, materials with similar hardness values but different microstructures may exhibit vastly different tribological behaviors, with reinforced surfaces offering increased resistance to abrasion and wear. The unique properties conferred by FSP allow for the integration of hard particles, such as B_4C, as a final surface treatment for heat-treatable Al-alloys. This reinforcement enables the surface to achieve hardness levels comparable to hardened Al-alloys, along with the added benefit of improved wear resistance (Ref 4).

FSP significantly improves the mechanical and tribological properties of aluminum matrix materials when micro- or nanoparticles are incorporated into the matrix. Generally, the matrix materials used in these studies include wrought aluminum from non-heat-treatable series like 1xxx (Ref 12, 13) and 5xxx (Ref 14, 15), and heat-treatable series such as 6xxx and 7xxx (Ref 16-18), though typically in an annealed state. However, little research has applied FSP to hardened materials like 2xxx series, where loss of hardness is seen due to thermal cycles due to intense heat of FSP. Researchers are working to retain or improve its hardness properties by incorporating reinforcing particles utilizing multi-pass friction stir processing (MPFSP). However, achieving uniform dispersion of reinforcements within the matrix by utilizing MPFSP is challenging, as multiple passes can lead to particle clustering, agglomeration, and uneven distribution due to improper material flow and inadequate stirring. Moreover, optimizing the strengthening mechanism is crucial, which involves controlling the grain size, dislocation density, and particle-matrix interface. The repeated thermal cycles during multi-pass FSP may lead to grain coarsening, reducing the overall strengthening effect. The final target is to enhance the mechanical properties of the AMC by achieving a homogeneous reinforcement distribution, refined grain structure, and strong interfacial bonding, ultimately improving strength, hardness, and wear resistance. Ghanbari et al. (Ref 19) fabricated AA2024/SiC-based AMCs, where they find that three pass FSP resulted in a uniform distribution of reinforcement particles and improved hardness properties. A same aluminum surface composite has been fabricated through MPFSP by Acuna et al. (Ref 4). They found a microhardness value of the composite that reached the base metal value, while the wear rate was reduced to 24% after three passes of FSP. Similarly, an aluminum composite of AA2024/Al_2O_2 has been formed by employing three numbers of FSP passes, and they found enhanced ultimate tensile strength (UTS), ductility, and hardness due to the refined grains, homogenous reinforcement dispersion, and enhanced matrix-reinforcement missing due to the increased number of FSP passes (Ref 20). Rana et al. (Ref 21), who developed an AA7075/B4C surface composite through FSP. The composite demonstrated a remarkable 70% increase in hardness over the base Al-alloy in its annealed state (75-80 HV), achieving hardness values close to those of T651-hardened aluminum (175 HV) without undergoing any additional hardening treatments. The aluminum-based hybrid AA7475/B4C/Al_2O_2 composite is fabricated successfully through multi-pass FSP in another study (Ref 22). The authors conducted four passes at a constant tool rotational speed of 1300 rev/min and a feed rate of 45 mm/min, using equal proportions of reinforcements. This resulted in increases in UTS, ductility, and hardness strengths of 22.79%, 69.01%, and 38.12%, respectively. The research conducted by Pranavi et al. (Ref 23) involved the mechanical characteristics of the Al5059 by incorporating B4C and Al_2O_3 particles, whereby the components were incorporated according to a different weight percentage. The investigation elucidated the enhancement in the surface hardness and UTS due to the increase in weight percentage of reinforcements, but there was decrease in elongation properties corresponding to the proportional increase in weight percentage. Moreover, the coefficient of friction was decreased with an increase in the wt.% percentage of the reinforcement particles in the fabricated composites. The wear resistance of aluminum (AA6061) matrix is also improved by incorporating hybrid Al_2O_3 and ZrO_2 reinforcements (Ref 24). Giridharan et al. (Ref 25) enhanced tensile, hardness and fatigue properties of the dissimilar steel-copper welded joints by incorporating SiO_2, GNP and biochar reinforcements via FSP. The maximum ultimate tensile strength of 225 MPa and a fatigue strength of 168 MPa were achieved with 2 wt.% GNPs, whereas the addition of 2 wt.% biochar produced the second highest strength factor. The findings indicate that biochar nanoparticles obtained from natural sources could serve as a substitute for expensive nanofillers in metal joining, achieving over 85% effectiveness compared to the costly fillers for equivalent results. Sezhian et al. (Ref 26) also improved the copper weld properties by incorporating B4C particles to the weld region via FSP. The FSP is conducted using varying weight percentages of B4C particles (ranging from 1 to 4 wt.%), resulting in welds without defects at a spindle speed of 1100 rpm, a travel speed of 30 mm/min, a plunge depth of 0.25 mm, and an axial stress of 5 KN. The findings indicated that incorporating nanoparticles (3 wt.%) into the joint achieved a highest UTS of 203 MPa, hardness of 123 HV, and a fatigue strength of 159 MPa.

Although various research works have been performed by using different reinforcements to produce aluminum-based composites via FSP, an investigation into the fabrication of AA2024/B_4C/Al_2O_2 surface composites using MPFSP is lacking. The objective of this study is to fabricate, characterize, and examine the mechanical and tribological performance of AA2024/B_4C/Al_2O_2 surface composite via MPFSP. The in-depth microstructural analysis of the fabricated composites is carried out using OM, SEM-EDS, EBSD, and XRD that delves into the dissemination of reinforced particles, grain size, grain boundary transformations, and phase structure. Furthermore, the research evaluates the influence of MPFSP on UTS, elongation, and hardness, alongside the tribological behavior of the hybrid composites. The purpose of this work is to enhance the durability and performance of aluminum matrix (AA2024-T351) by incorporating hybrid reinforcements while maintaining a lightweight structure, making them suitable for sustainable applications ranging from automotive to aerospace industries.

2. Materials and Methods

This research employed AA2024-T351 as a substrate material with dimensions of 140 × 100 × 6 mm in thickness to produce nanocomposites via FSP. The chemical composition of the base metal was demonstrated in Table 1 analyzed through spectroscopic process. The mechanical characteristics of the substrate material, derived from experimental results, are presented in Table 2. Nanoparticles such as boron carbide (B_4C) and alumina (Al_2O_3) were employed as reinforcement materials for this study. The mean grain size (MGS) of the B_4C particles was approx 13.8 μm, while the MGS of Al_2O_3 particles was approx 10.1 μm.

Nanocomposites of AA2024/B_4C/Al_2O_3 was fabricated by employing multi-pass FSP on a modified vertical milling machine (Fig. 1a) with a designed clamping fixture setup. The purpose of the specific fixture is to hold the aluminum plate in a proper. The FSP process utilized a tool with the help of profiled pin (Fig. 1b). The tool pin profile featured a taper threaded shape with a 1 mm slot on each face. Before performing MPFSP, the capping pass was performed with pin less tool to compact the hybrid reinforcement particles. To incorporate these nanoparticles into the base alloy, a groove measuring 1 mm in width and 3 mm in depth was cut using wire EDM across the central section of the 140 mm length (Fig. 1c). After compaction, six number of passes were performed by profiled pin. The taper threaded pin tool enhances the mixing of material within the tool pin's area. The FSP tool, made from high-speed H13 steel, had a shoulder diameter of 20 mm and included a square pin measuring 4.5 mm in length and 6 mm in diameter. Table 3 reveals the FSP parameters utilizing the tool rotated at 1300 rpm, a traverse speed of 45 mm/min with a tilt angle of 0° with a plunge depth of 0.12 mm. The six number of FSP passes were performed to facilitate effective mixing of reinforcements with aluminum matrix. The range of FSP factors was determined through trial experiments, experimental conditions, and literature (Ref 20, 27, 28) to ensure the fabrication of defect-free hybrid AA2024/B4C/Al_2O_2 composites. The defect free hybrid AA2024/B_4C/Al_2O_3 surface composites fabricated with MPFSP are revealed in Fig. 1(d).

After the production of composites, tensile and hardness test specimens were prepared by wire electrical discharge machining. Tensile test samples were fabricated as per ASTM E8 standard (refer to Fig. 1e). Tensile tests were conducted using a UTM machine (Tinius Olsen H50KS) with a crosshead speed of 2 mm/min. Hardness measurements of the nanocomposite were performed across the mid cross section of the specimen with a 1 mm gap between two consecutive indents (Fig. 1f) using a microhardness tester (Omnitech YOM-2010), applying a 100-gf load for 10 seconds. Dry sliding wear experiments were conducted on a rotary tribometer using a test pin with dimensions of 4.5 × 4.5 × 10 mm^3 as per ASTM G99-05 standard. The experiments were conducted for a duration of 10 min at a tracking diameter of 45 mm, a vertical burden of 40 N, and a disk rotation speed of 850 rpm. The sliding distance was 1000 m. For metallographic analysis, samples were mirror finished using standard polishing techniques, and then etched in Keller's reagent for 10 to 15 seconds. Microstructural images were captured using an optical microscope (OM) and a scanning electron microscope (SEM). Elemental distribution in the stir zones of the processed samples was analyzed using energy-dispersive spectroscopy (EDS). Phase identification within the treated regions was performed through x-ray diffraction (XRD) using CuKα radiation, covering a 2θ range of 10-90°. Grain size was determined using the line intercept method via ImageJ software, providing detailed insights into the material's phase characteristics (Ref 29).

3. Results and Discussion

3.1 Microstructural Evolution

The properties and performance of base metals and fabricated hybrid surface composites are essential to comprehend through microstructural analysis. SEM and OM generate detailed images of reinforcement particle distribution, grain structure, and surface morphology. EDS is performed to identify the elemental composition, precipitate distribution (if any), and to ensure the dispersion of hybrid reinforcements within the material. Phase analysis is conducted via XRD technique to confirm the intermetallic phases formation, their peak intensity variation with respect to diffraction angle 2θ. EBSD is performed to determine the grain orientations and the grain boundary transformations as a result of FSP. These methods collectively reveal the microstructural evolution and reinforcement distribution, confirm homogeneity and detect defects, relate microstructure to mechanical properties (including strength, hardness, and ductility), optimize fabrication FSP variables, and improve material design and application suitability.

Figure 2(a) depicts the microstructure of the base alloy (AA2024-T351) showing elongated grains oriented along the rolling direction with significant intermetallic phases dispersed in the matrix. The elongated grains contain intermetallic precipitates (Al_2CuMg) both within the grains and along their boundaries. These precipitates, ranging from 3 to 10 μm in diameter, contribute significantly to the alloy's strength and ductility through precipitation hardening (Ref 30). Figure 2(b) illustrates the interface region of the composite, highlighting the stir zone (SZ), thermomechanically affected zone (TMAZ), and the heat-affected zone (HAZ). In MPFSP, the SZ develops from intense plastic deformation and localized thermal input, leading to fine-equiaxed grains through dynamic recrystallization (DRX), in which deformed grains are constantly replaced by new, strain-free grains. The TMAZ arises from a combination of thermal and mechanical effects, while the HAZ is created by the thermal cycles produced by the rotating tool, resulting in localized heating, thermal softening, and microstructural changes without considerable plastic deforma-

Table 1 Specification of base alloy

Elements	Si	Mg	Mn	Zn	Cu	Fe	Ti	Cr	Al
AA2024-T351	0.07	1.4	0.64	0.06	4.57	0.10	0.04	0.12	Balance

Table 2 Mechanical Properties determined experimentally

Base alloy	YS, MPa	UTS, MPa	El, %	Wear depth, μm	Vickers hardness, HV
AA2024-T351	308	421	19.1	105	136

YS Yield strength, *UTS* Ultimate tensile strength, *El* Tensile elongation.

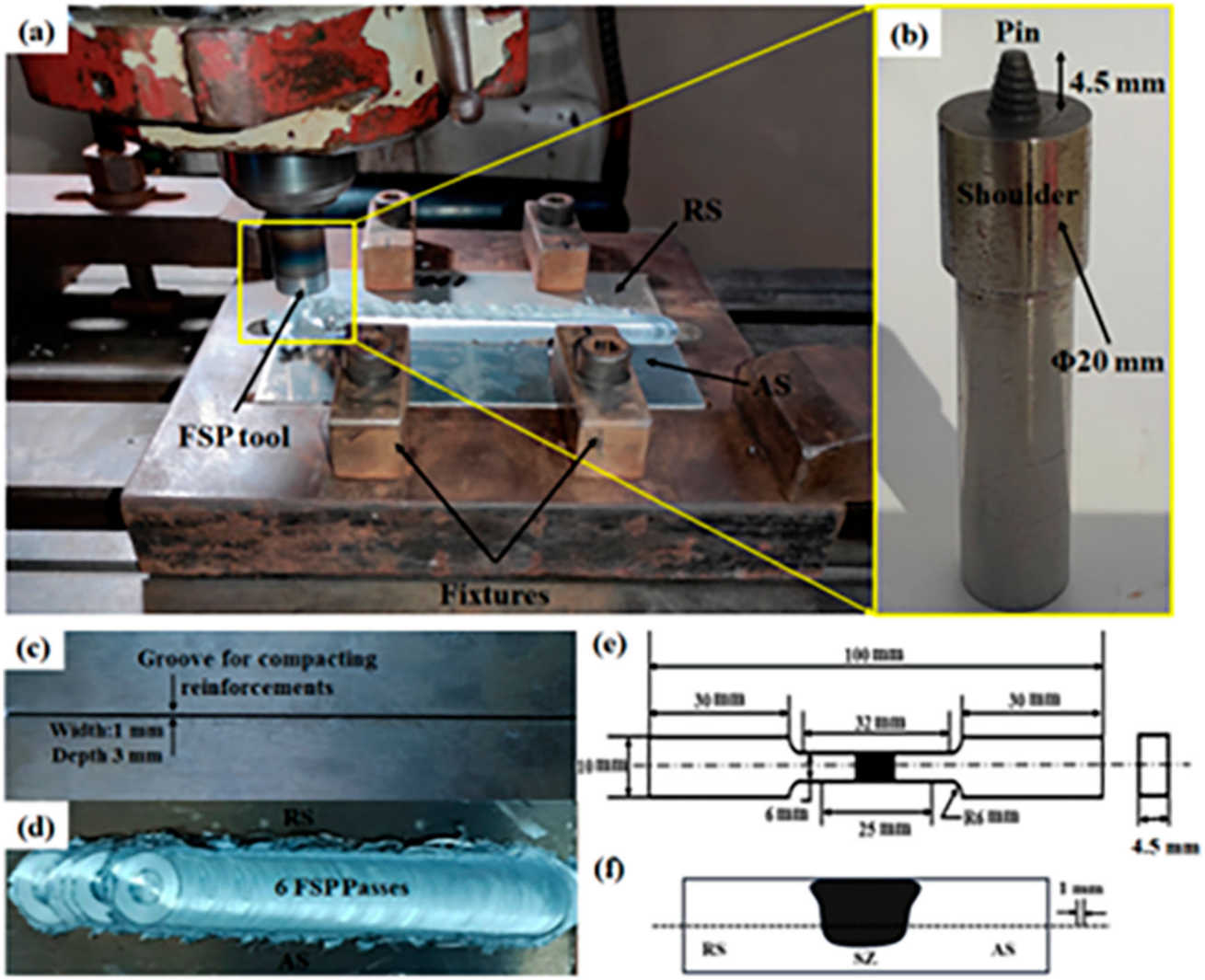


Fig. 1 Fabrication of hybrid AA2024/B_4C/Al_2O_3 surface composites: (a) Vertical milling machine, (b) FSP tool, (c) groove for compacting reinforcements, (d) fabricated nanocomposite, (e) tensile test specimen, and (f) hardness test specimen details

Table 3 FSP parameters

Processing parameters	Values
Tool rotation speed, rpm	800
Tool advancing rate, mm/min	35
Number of passes	6
Reinforcement fraction, %	50% B_4C and 50% Al_2O_3
Tool plunge depth, mm	0.12

tion (Ref 31). Increasing the number of FSP passes significantly influences the grain morphology and reinforcement particle distributions in AA2024/B_4C/Al_2O_2 composites, specifically increasing the number of FSP passes from one pass to six passes refines and homogenizes the grain structures and results in fragmentation and uniform dispersion of particles across the SZ. Figure 2(c), corresponding to the first pass of FSP, the microstructure reveals the presence of non-fragmented B_4C and Al_2O_2 particles, indicating that the reinforcement particles are still in their original form without significant breakdown. The distribution of these particles is likely not uniform, as the FSP has just begun to plastically deform the material, and the stirring effect has not been sufficient to homogenize the matrix. Moving to Fig. 4(d), which shows the second pass of FSP, it is observe that the composite material undergoes further refinement. The stirring effect has increased, leading to a better

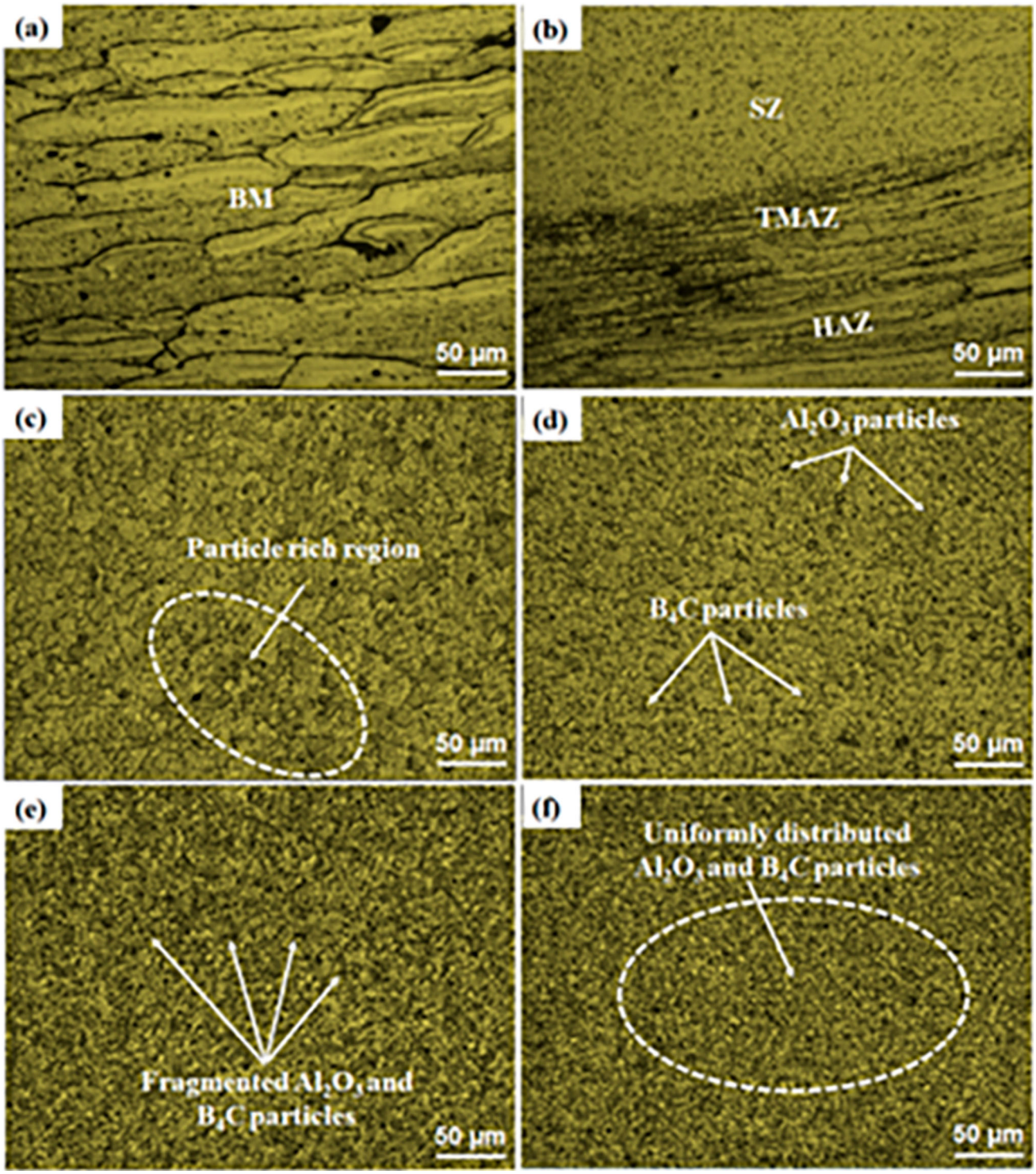


Fig. 2 OM images of the BM and AA2024/B_4C/Al_2O_3 nanocomposites for different FSP passes: (a) AA2024-T351, (b) Interface region, (c) 1 Pass, (d) 2 Pass, (e) 4 Pass, and (f) 6 Pass

mixing of the aluminum matrix and the hybrid reinforcement particles. However, compared to subsequent passes, the distribution may still not be entirely uniform, and some clusters or agglomerations of B_4C and Al_2O_3 is still persist. In Fig. 4(e), corresponding to the fourth pass of FSP, the microstructure shows significant improvement in terms of particle distribution. The stirring action from multiple passes enhances the breaking down of the reinforcement particles, resulting in a finer and more uniformly distributed particle phase. At this stage, fragmentation of B_4C and Al_2O_2 is more apparent, contributing to a more homogeneous microstructure. Figure 4(f), which depicts the microstructure after six passes of FSP, shows the most refined and uniform distribution of the B_4C and Al_2O_2 particles. The particles appear to be thoroughly dispersed throughout the matrix, with no signs of significant agglomeration. The repeated FSP passes have led to the uniform

fragmentation and incorporation of the reinforcement particles, suggesting that the material has achieved a more optimized and homogeneous microstructure for enhanced mechanical properties.

Figure 3 shows the grain size distribution the BM and hybrid composites for different FSP passes. The base AA2024-T351 alloy exhibited a mean grain size (MGS) of 65.64 ± 1.94 μm (Fig. 3a) and the MGS consistently decreases with an increasing FSP passes. The MGS starts at around 7.18 ± 0.26 μm after the first pass, around 4.46 ± 0.14 μm after the fourth pass (Fig. 3c), and reduces to 2.82 ± 0.06 μm after six passes (Fig. 3d). FSP induces severe plastic deformation and high temperatures in the material, leading to DRX. With each pass, new grains nucleate and grow, replacing the existing larger grains. Repeated passes enhance the recrystallization process, reducing the MGS (Ref 32). Each FSP pass applies mechanical deformation through the rotating tool, further breaking down the grain structure. The more deformation a material undergoes, the finer the grains tend to become as the original grains are fragmented and replaced by smaller grains (Ref 33). After six passes, the grain size seems to stabilize at about 2.8 μm signifying that the grains may attain an equilibrium size and further deformation does not dramatically diminish size. The stabilization of grain size after MPFSP is primarily governed by mechanisms such as dynamic recrystallization (DRX), Orowan strengthening, Zener pinning, and particle-stimulated nucleation (Ref 34). DRX plays a crucial role by continuously refining grains through strain-induced nucleation and growth, ultimately leading to a stable ultra-fine or nano-grained structure. Orowan strengthening occurs as dispersed second-phase particles hinder dislocation motion, maintaining grain size stability. Zener pinning further contributes by restricting grain boundary migration due to the presence of finely distributed precipitates, preventing excessive grain coarsening. Additionally, the particle-stimulated nucleation mechanism enhances grain refinement by promoting nucleation sites around dispersed particles, ensuring a balanced microstructure. The interplay of these mechanisms ensures that grain growth remains limited, leading to a stabilized fine-grained structure after multiple FSP passes.

Figure 4 illustrates the SEM images of the hybrid AA2024/B_4C/Al_2O_3 surface composites fabricated with six different passes. In Fig. 4(a), a single FSP pass notably reduces the coarse dendritic structure of the substrate alloy and transformed it into a fine equiaxed grains structure. During subsequent passes, material mixing improves, leading to a reduction in hybrid nanoparticle aggregation within the surface composite. The B_4C and Al_2O_3 nanoparticles undergo significant and progressive fragmentation following the second and fourth FSP passes (Fig. 4b and c), correlating with the decreased B_4C and Al_2O_3 nanoparticles aggregation observed as the FSP passes increase (Ref 35). The area percentage of B_4C increases with each FSP pass, directly linked to the grain refinement and precipitates fragmentation process. By the six passes, the grains become more homogenous and finer (Fig. 4d) due to DRX processes associated with the Zener–Holloman parameter (Ref 36). The MPFSP process consequently yields a refined and uniform microstructure, primarily driven by lower thermal input and accumulated strain. Compared to the one-pass FSP process, the four-pass FSP process demonstrates superior

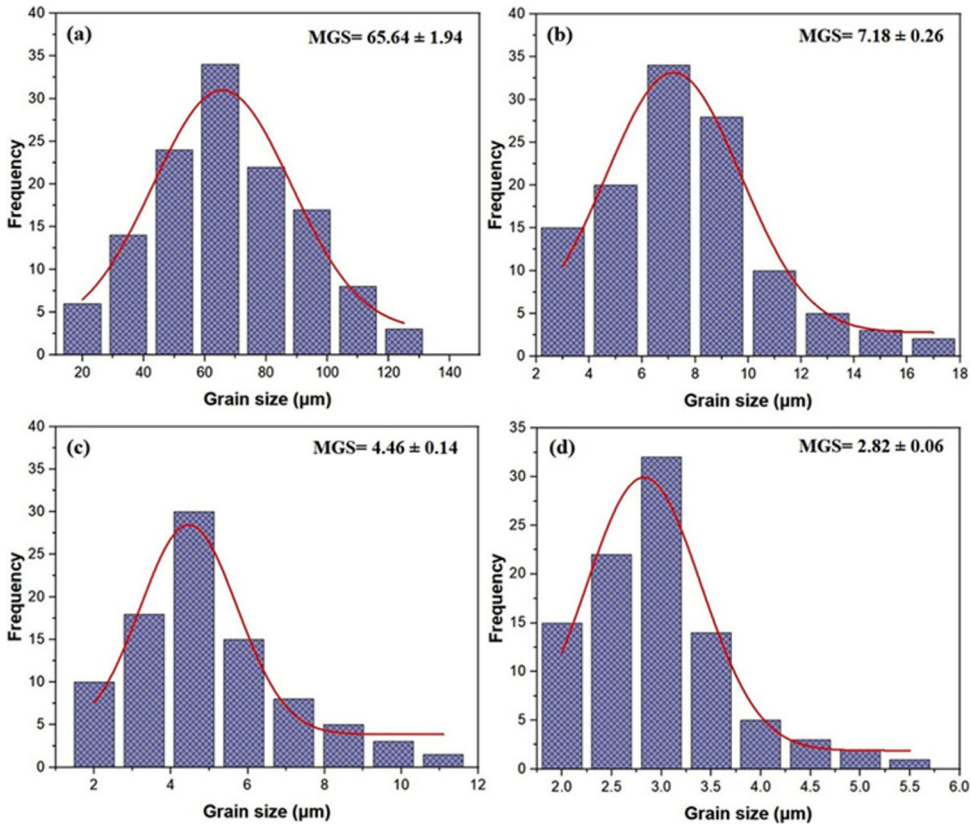


Fig. 3 Grain size distribution of the AA2024/B_4C/Al_2O_3 nanocomposites for different FSP passes: (a) BM, (b) 1 Pass, (c) 4 Pass, and (d) 6 Pass

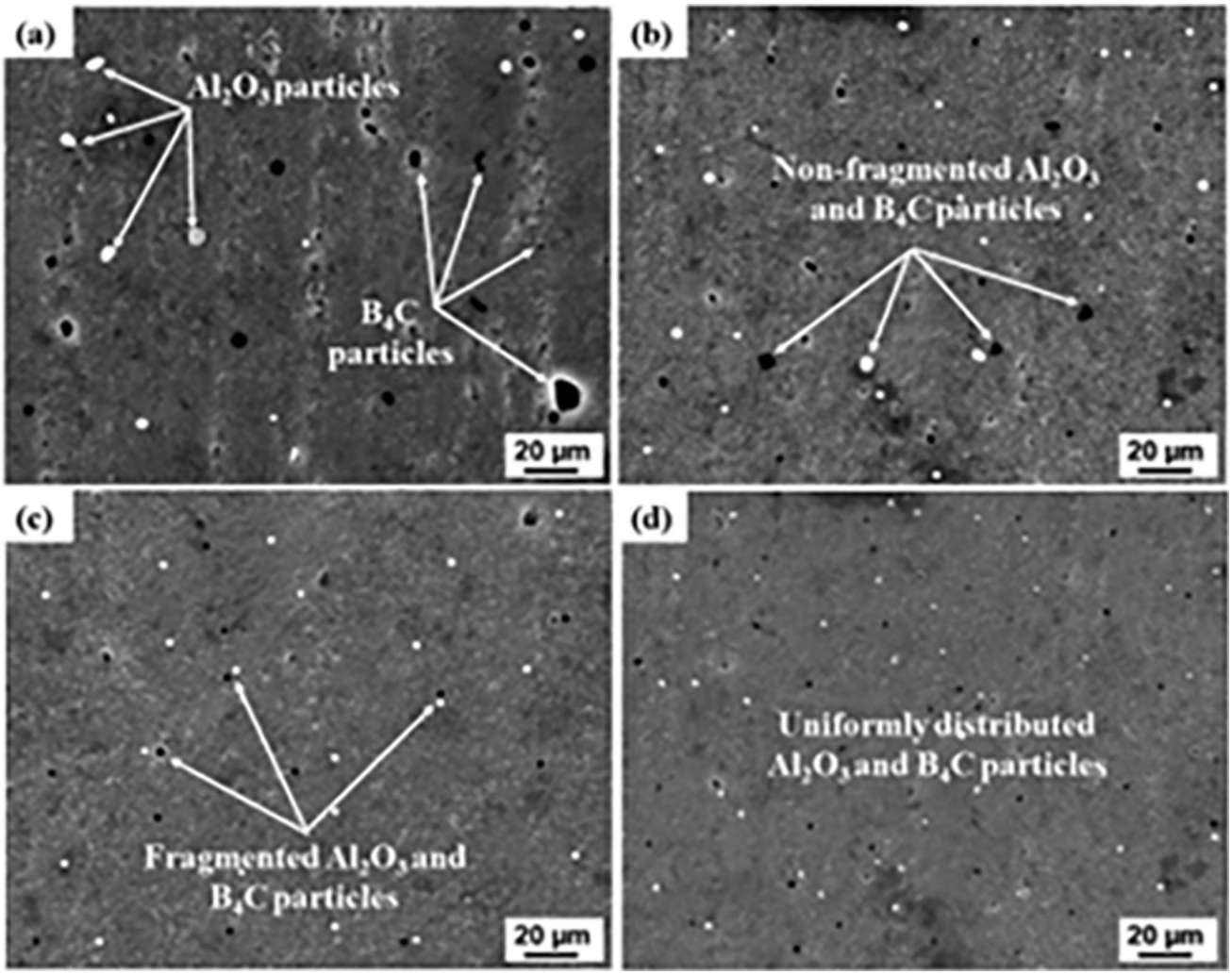


Fig. 4 SEM images of the AA2024/B_4C/Al_2O_3 nanocomposites for different FSP passes: (a) 1 Pass, (b) 2 Pass, (c) 4 Pass, and (d) 6 Pass

mechanical properties of hybrid AA2024/B_4C/Al_2O_3 surface composites (Ref 37]. Moreover, the size of B_4C and Al_2O_3 nanoparticles reduces significantly in the six-pass FSP compared to the one-, two-, and four-pass FSP processes. MPFSP leads to more uniform particle dispersion and contributes to breaking up SiC nanoparticle clusters due to cumulative strain. This process facilitates the creation of a large number of Al-B_4C-Al_2O_3 bonds and optimizes the distribution of nanoparticles, improving the pinning effect that is essential for strengthening and structural stability. Further, dislocation sources emerge during MPFSP, enhancing the DRX process and producing a refined grain structure (Ref 38). Moreover, as the FSP passes increases, B_4C and Al_2O_3 nanoparticles aggregation decreases, with single-pass FSP highlighting inadequate adhesion between highly agglomerated particles and the Al-alloy matrix. Defects such as fractures and discontinuities reduce around the agglomerated B_4C particles with additional FSP passes (Fig. 6a-d). During the plunging phase, the FSP tool applies compression to both B_4C and Al_2O_3 nanoparticles and the substrate metal matrix, causing lateral and upward particle motion that establishes a process zone (Ref 39). Notably, no separation or voids are detected between both B_4C and Al_2O_3 nanoparticles and the substrate matrix, suggesting that the six-pass FSP specimens exhibit higher UTS than those with fewer passes.

Figure 5 illustrates the elemental mapping results of the SZ for the six-pass hybrid surface composites, which include dispersed particles within the composite. The red areas indicate a high concentration of aluminum (Al-Kα), suggesting the existence of an AMCs within the composite. The gold color indicates the region with concentration of copper (Cu-Kα), blue area highlights regions with a concentration of oxygen (O-Kα), which may suggest the presence of aluminum oxides or other compounds that contain oxygen and copper in the material. The cyan regions indicate the integration of boron (B-Kα)-based reinforcements within the matrix. The green area suggests that the presence of carbon (C-Kα) could imply the existence of boron carbide reinforcements. This elemental distribution pattern demonstrates the effective mixing and integration of reinforcement particles during the FSP, resulting in a hybrid composite structure where multiple elements are distributed throughout the SZ, contributing to the material's enhanced properties through the combination of different reinforcement mechanisms. The EDS report displays the weight percentages of elements found in the single-pass FSP hybrid composite. The primary components are Al (91.98%), Cu (3.45%) and

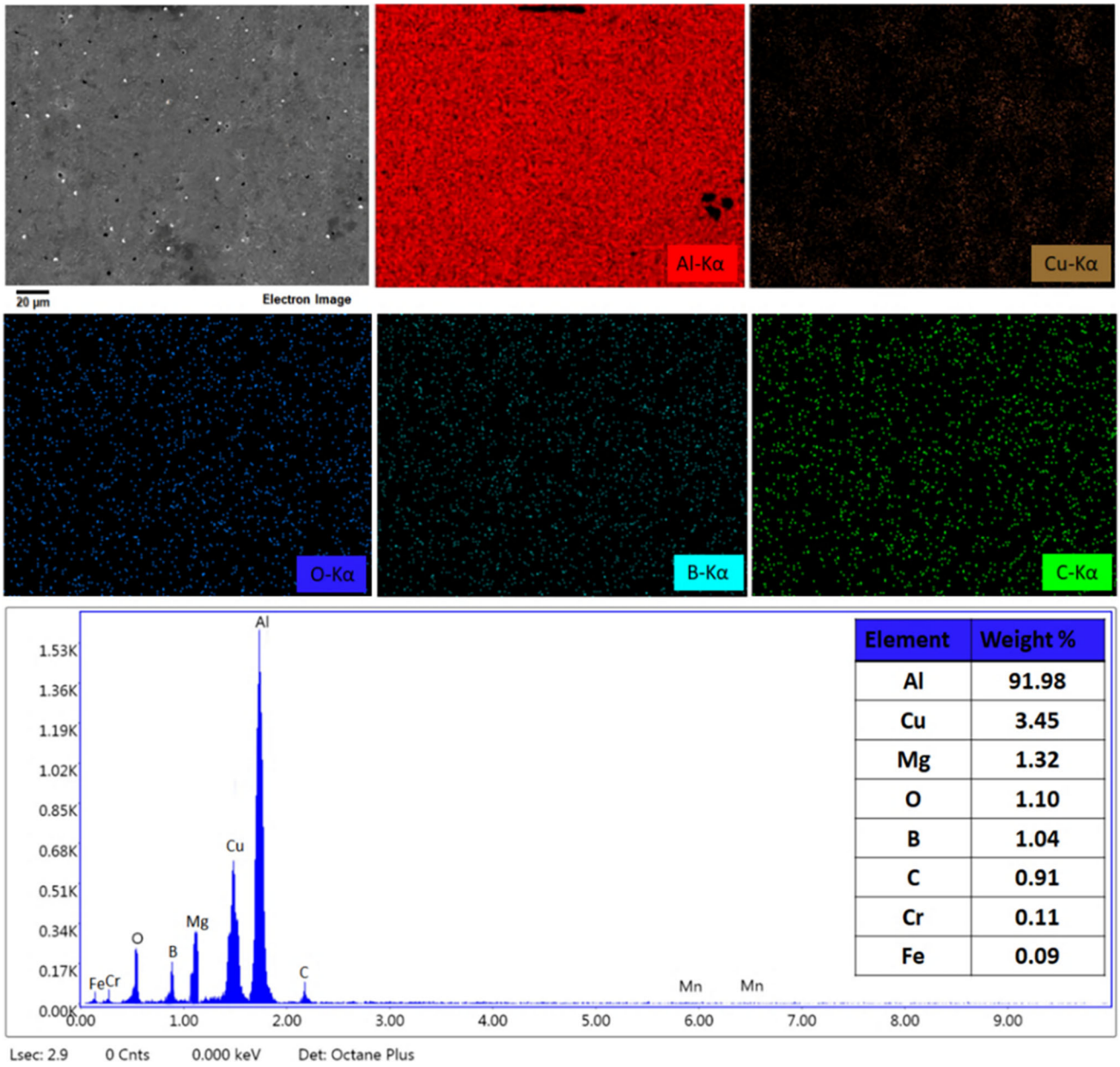


Fig. 5 Elemental mapping with EDS report of the AA7475/B_4C/Al_2O_3 nanocomposite with six pass FSP

magnesium (1.32%), indicating that the material is mainly an Al-alloy with significant amount of copper and magnesium elements may form Al_2CuMg phase. The reduced magnesium content exhibits that the Al_2CuMg phase has been reduced after MPFSP. While, reduced oxygen at 1.10%, boron at 1.04%, and carbon at 0.91% contents indicate that the multiple passes might have contributed to the reduction of aluminum oxides (Al_2O_3) and boron carbides (B_4C) or improved the dispersion of the reinforcement particles. The decrease in oxygen levels and enhanced homogenization indicate that MPFSP enhances the distribution and minimizes oxides in the composite. The results suggest that several FSP passes can improve the quality of hybrid surface composites by enhancing the uniformity of reinforcements (Ref 40).

The XRD patterns for the multi-pass hybrid AA2024/B_4C/Al_2O_3 surface composites in the stir region with Cu-Kα radiation at a wavelength of 1.518 Å are shown in Fig. 6. There are four main phases such as α-Al, Al_2CuMg, Al_2O_3 and B_4C are identified in the hybrid surface composites. No adverse reactions were identified between the Al_2O_3 and B_4C and the substrate metal matrix, demonstrating their compatibility during the MPFSP process. The phases identified in the XRD analysis are in agreement with EDS findings, which also shows the significant contents of the four alloying elements, namely Mg, Zn, B, and C. Moreover, the intensity of phases decreased with increasing FSP passes, signifying the uniform distribution and dispersion of second phase particles within the aluminum matrix, resulting in improved mechanical properties and microstructural homogeneity. This trend well aligns with findings of the previous study (Ref 41), that increasing FSP passes enhances Al_2O_3 particle dispersion and fragmentation and resulted in improved mechanical properties. These phases, when finely dispersed and exhibiting a coherent or semi-coherent interface with the surrounding matrix, can contribute

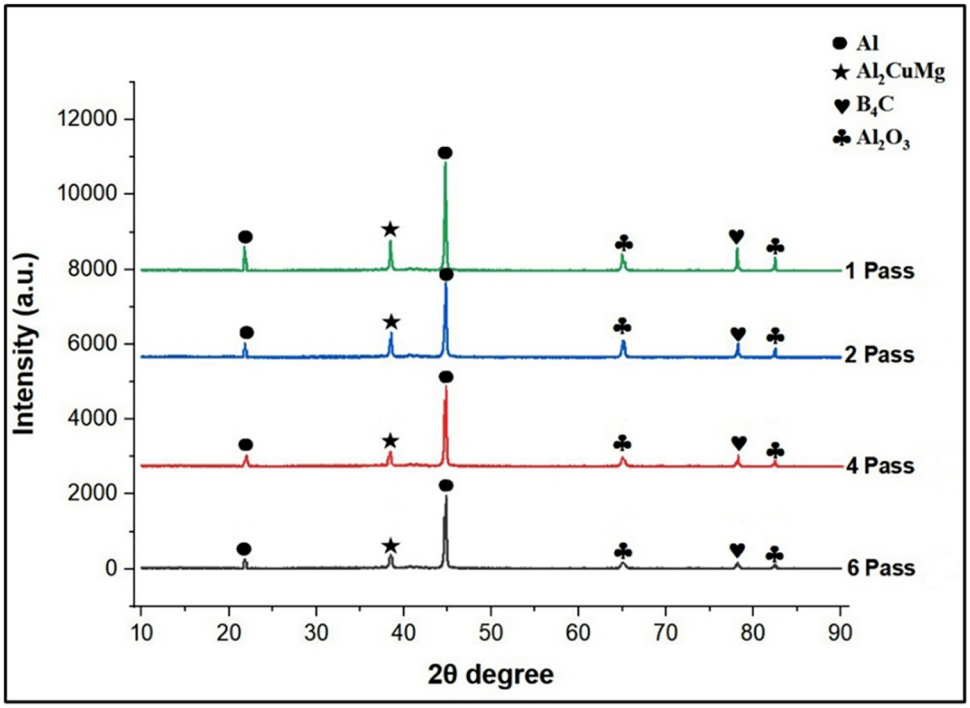


Fig. 6 XRD pattern of the AA2024/B_4C/Al_2O_3 nanocomposites for different FSP passes

to the composite's strengthening through a process known as precipitation hardening. The variations in the phase intensity of different phases can be attributed to several factors, including the heat input, plastic deformation, and material flow during FSP (Ref 42).

Figure 7 displays the EBSD images showing inverse pole figure (IPF) maps, grain boundary (GB) maps, and the distribution of low and high-angle grain boundary in the stir zone of the hybrid AA2024/B_4C/Al_2O_3 composites compared along with the base alloy. The colors in the IPF maps represent the crystallographic orientation of the grains, while the misorientation angle distributions highlight the grain boundary transformations resulting from the MPFSP. Low-angle grain boundaries (LAGBs) are defined by a misorientation angle between 2 and 15°, while high-angle grain boundaries (HAGBs) have angles greater than 15 degrees (Ref 10). Crystallographic orientation refers to the angular relationship between a grain's crystal lattice and a fixed reference frame, such as the sample surface or a standard coordinate system. It describes how individual grains are aligned in a polycrystalline material and helps in understanding texture evolution, slip systems, and grain boundary characteristics, which are crucial for evaluating material behavior due to thermomechanical process. They reveal the evolution of grain refinement, texture development, and recrystallization mechanisms. Additionally, the grain boundary transformations from LAGBs to HAGBs signify a more pronounced dynamic recrystallization process. This transformation indicates that the material has undergone significant plastic deformation, leading to the formation of new, strain-free grains with high-angle boundaries. The increased fraction of high-angle grain boundaries suggests a stronger texture (a significant number of grains have similar crystallographic orientations), which can contribute to improved mechanical properties (Ref 43). Black dots observed among the grains in the IPF map signify the existence of B_4C and Al_2O_3 particles within the stir region. In the IPF map of the composite fabricated with single pass, an equiaxed grain morphology (agreed with OM and SEM results) is observed with relatively coarser grains, as shown in Fig. 7(a), with the grains are equiaxed. The misorientation angle represents the difference in crystallographic orientation between neighboring grains. In the AA2024/B_4C/Al_2O_3 composites, the fraction of HAGBs is 76.49% for the single pass FSP (Fig. 7a) while the fraction of HAGBs is increased as the FSP pass is increased. The composite fabricated with double-pass FSP exhibited a higher fraction of HAGBs of 84.03%, as illustrated in Fig. 7(b). However, the fraction of HAGBs is 91.05% for the composite fabricated with four numbers of FSP passes (Fig. 7c), and it reaches 95.21% for the composite fabricated with six numbers of FSP passes (Fig. 7d). The composite fabricated with six FSP passes exhibited the highest grain boundary transformations with significant grain refinement. The role of Zener pinning and particle-stimulated nucleation (PSN) in grain refinement during MPFSP is significant due to the presence of second-phase particles such as B4C and Al_2O_3 reinforcements. These particles act as obstacles to grain growth through the Zener pinning effect, which restricts the migration of grain boundaries. Additionally, PSN contributes to grain refinement by providing nucleation sites for recrystallization. The IPF map of the composite after six FSP passes demonstrates that B4C and Al_2O_2 reinforcements effectively hinder grain growth, leading to the formation of ultrafine grains. The grain refinement is further supported by the occurrence of continuous DRX during FSP, which facilitates the transformation of coarse grains into finer grains. Similar observations were reported in the previous study (Ref 44), where significant grain refinement in the AMCs is achieved by the combined effect of Zener pinning and PSN. These combined mechanisms play a crucial role in achieving ultrafine grains in hybrid composites fabricated via MPFSP.

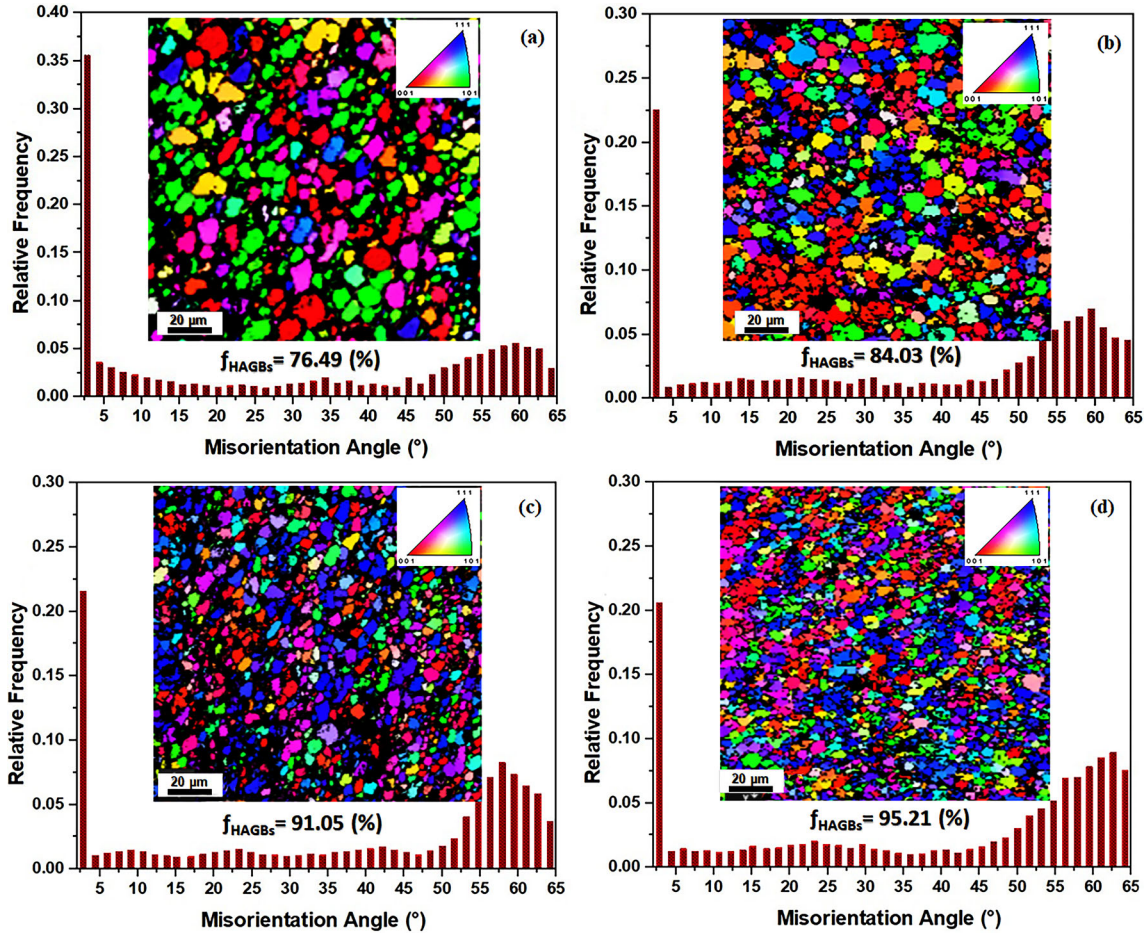


Fig. 7 EBSD micrographs (IPF + GB) reveal the misorientation angle distribution in the SZ of the AA2024/B_4C/Al_2O_3 nanocomposites for different FSP passes: (a) 1 Pass, (b) 2 Pass, (c) 4 Pass, and (d) 6 Pass

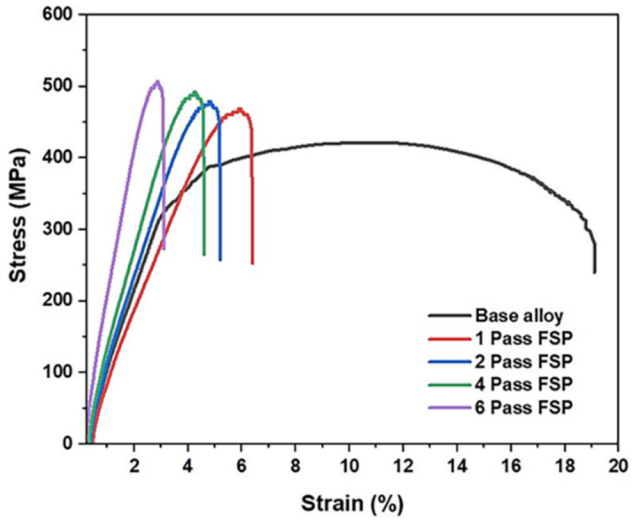


Fig. 8 Engineering stress-strain curve for different FSP passes

3.2 Tensile Properties

Tensile tests are performed to evaluate the tensile performance of the fabricated material, which is crucial in structural and industrial applications (Ref 45). The addition of B_4C and Al_2O_3 nanoparticles to the AA2024-T351 matrix is expected to enhance its strength, ductility, and toughness. Figure 8 presents the engineering stress-strain curves derived from tensile testing of hybrid nanocomposites, revealing the mechanical response of these materials after MPFSP. The data show that the UTS and YS of all samples significantly increased after MPFSP. However, these enhancements in strength were accompanied by a noticeable reduction in tensile elongation (El) as the FSP passes increased. The present findings are well agreed with prior findings (Ref 20, 46, 47), which showed that the MPFSP improved hardness, UTS, and YS while simultaneously diminishing El. The bar graphs in Fig. 9 quantitatively illustrate the changes in YS, UTS, and elongation of the hybrid nanocomposites, while Table 2 highlights that the base alloy retains the highest elongation compared to the processed composites. The mechanical behavior induced by FSP can be attributed to the synergistic effects of severe plastic deformation (SPD) heat-induced microstructural stability, and emerging strain-hardening mechanisms. SPD is a critical mechanism in enhancing material properties, particularly strength and toughness, by refining microstructures at the grain level. The process involves extensive deformation, which introduces a high density of dislocations and promotes grain refinement through dynamic recrystallization (DRX). However, the corresponding frictional and deformation heat plays a vital role in controlling the growth of these recrystallized grains. This heat, coupled with strain-induced

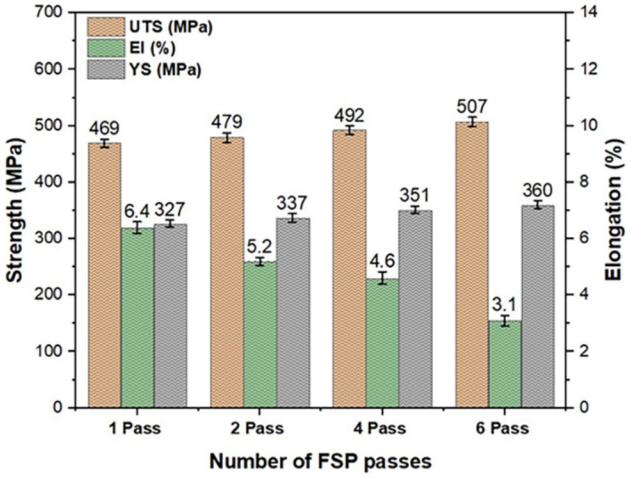


Fig. 9 Bar graphs showing UTS, yield, and elongation properties with different FSP passes

grain fragmentation, effectively inhibits excessive grain coarsening, ensuring the retention of ultra-fine or nanostructured grains that contribute to superior mechanical properties. Additionally, strain hardening mechanisms emerge as a result of this refined microstructure, where dislocation entanglement and interaction with grain boundaries enhance material strength while maintaining ductility. This effect is particularly important in overcoming the conventional strength-ductility trade-off, which typically limits the improvement of both properties simultaneously (Ref 48, 49). The fine-grain microstructure ensures improved grain boundary strengthening (Hall-Petch effect), while controlled dislocation motion facilitates strain hardening, thereby balancing strength and toughness efficiently (Ref 50). The synergy between SPD, heat-induced microstructural stability, and emerging strain-hardening mechanisms provides unique insights into the optimization of structural materials for high-performance applications. The alloy employed in this investigation is a precipitate-hardened Al-alloy. The Al_2CuMg phase is observed to nucleate along grain boundaries, as shown in Fig. 2(a). This phase behaves like secondary particles that impede dislocation motion, significantly contributing to the strength of the alloy. Furthermore, FSP promotes the formation of DRX in the Al_2CuMg phase within the aluminum matrix, along with the uniform dispersion of the reinforcement particles, namely B_4C and Al_2O_3. In this study, mechanical properties of the base alloy showed substantial improvements post-FSP, particularly after six passes. YS and UTS are increased by 16.88% and 20.42%, respectively, due to grain refinement, improved DRX, and enhanced phase distribution. On the other hand, ductility dropped markedly to 83.76% after six FSP passes, reflecting the trade-off between strength and ductility inherent in such processing techniques. The tensile test results presented in this study provide valuable insights into the nanoparticle dispersion and interfacial bonding within the composite. The improved YS and UTS observed with increasing FSP passes suggest enhanced nanoparticle dispersion and interfacial bonding. Specifically, the uniform distribution of nanoparticles and the strong bonding between the nanoparticles and the aluminum matrix contribute to the improved mechanical properties (Ref 51). The tensile test results imply that the MPFSP process effectively disperses the nanoparticles and promotes interfacial bonding, leading to enhanced composite performance.

Figure 10 shows the SEM images of the tensile fractured surface of the AA2024/B_4C/Al_2O_3 composite subjected to different passes, highlighting features such as dimples, tear ridges, facets, and micropores (Ref 50). With the increase in the number of FSP passes from one to six, the size of the dimples diminished (Fig. 10a–d), aligning with the microstructural characteristics that influence the failure modes. The fractured surface of the composite samples made with six FSP passes showed very fine and equiaxed dimples, leading to a ductile mode of failure. Fine dimples enhance the composite's strength due to the robust bonding between the materials, leading to considerable necking prior to failure and ultimately resulting in cup-cone failure (Ref 52). The cup-cone fracture demonstrates the ductile nature of failure, characterized by a fractured surface plane that is inclined at 45° to the tensile loading axis. Nonetheless, the fractured surface of the composite samples created using single-pass FSP displayed several coarse dimples along with rough cleavage facets, leading to a ductile-brittle fracture mode of failure. The transition from ductile to brittle fracture modes in the composite is affected by various factors. The existence of coarse dimples on the fractured surface of single-pass FSP composite samples points to a ductile fracture mode, whereas the presence of rough cleavage facets indicates a brittle fracture component. The failure mode characterized by ductile-brittle fracture can be linked to the interaction between the aluminum matrix and the reinforcement particles. The elements influencing this transition encompass the distribution and size of the ceramic particles, the bonding at the matrix-particle interface, and the strain rate experienced during deformation. Non-fragmented reinforcement particles may serve as stress concentrators, which can promoted brittle fracture, whereas the aluminum matrix is capable of plastic deformation, resulting in ductile fracture. All composite samples were found to fracture within the HAZ region, except the composites fabricated with a single pass, in which the failure occurs across the SZ region. In the HAZ region, localized softening takes place due to intense thermal cycles induced by MPFSP, which can result in the dissolution of precipitates, grain coarsening, and a lack of hard ceramic particles (Ref 2). The localized HAZ softening could lead to failure and diminish the tensile performance. The study reveals that MPFSP incorporating B4C and Al_2O_2 nanoparticles produces finer and more uniform microstructures, increases tensile features, and controls fracture location, thereby enhancing performance and durability of hybrid surface composites.

3.3 Hardness Distribution

The Vickers hardness distribution curves with respect to the distance from the center of hybrid AA2024/B4C/Al_2O_3 surface composites are illustrated in Fig. 11. It can be clearly observed that as the FSP passes increases from one to four, a notable increase in hardness values is observed (Ref 47). This improvement in hardness correlates with refined grain size and increased nanoparticle dispersion within the composite structure, suggesting that each additional pass leads to enhanced material properties. In particular, Vickers hardness values observed along the SZ showed a progressive increase: from 155 HV in the one-pass FSP to 164 HV in the two-pass, 172 HV in the four-pass, and finally to 189 HV in the six-pass FSP (see Fig. 13). This gradual improvement contrasts with the base metal AA2024, which had a lower average hardness of 136 HV (see Table 2). Each pass in the FSP process contributed to

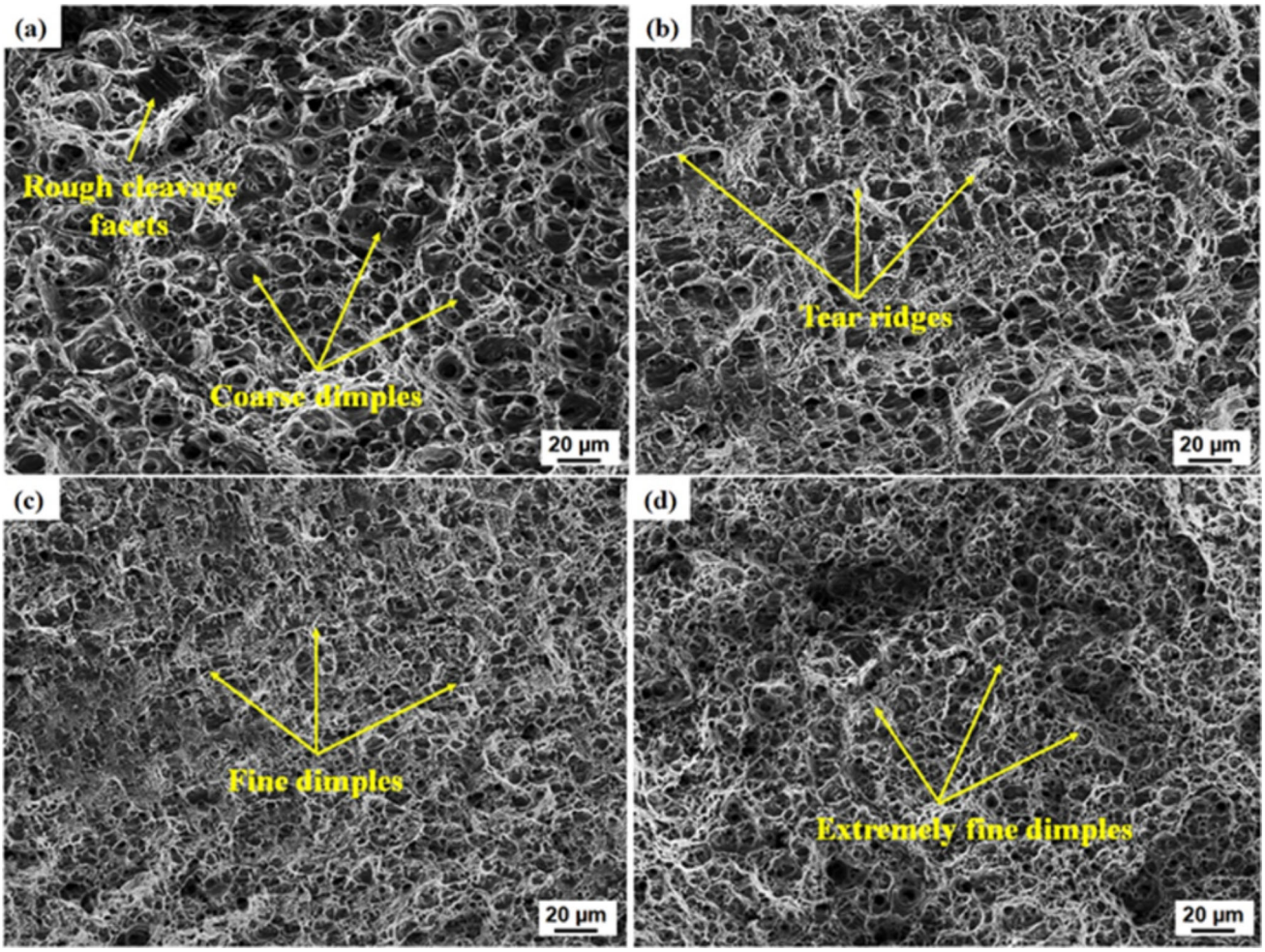


Fig. 10 SEM images of the tensile fractured surfaces of the AA2024/B_4C/Al_2O_3 nanocomposites for different FSP passes: (a) 1 Pass, (b) 2 Pass, (c) 4 Pass, and (d) 6 Pass

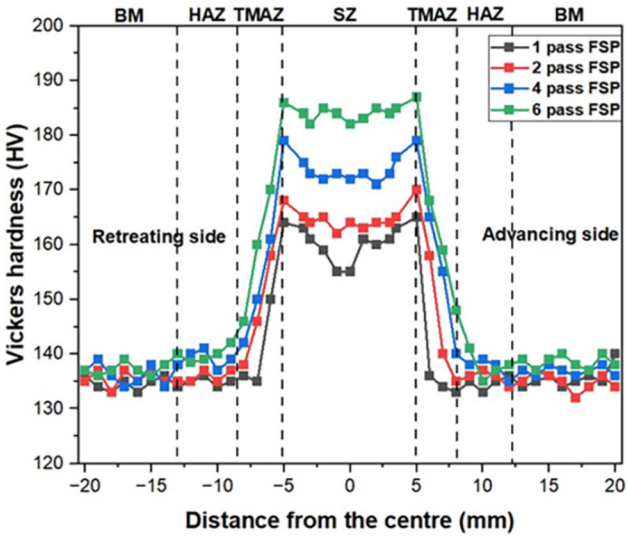


Fig. 11 Hardness distribution of the AA2024/B_4C/Al_2O_3 nanocomposites for different FSP passes

improved hardness through mechanisms like grain fragmentation, refinement of B4C and Al_2O_3 nanoparticles, and the incorporation of B4C and Al_2O_3 particles. These particles are hard, reinforcing agent, enhances composite microhardness due to its inherent structural properties. In the SZ across all MPFSP configurations, a low fluctuation in hardness reflects effective mixing and dispersion of these nanoparticles, which contributes to the uniformity of the SZ microstructure. The hardness increment seen in MPFSP with B4C and Al_2O_3 nanoparticles infill can be attributed to the grain refinement and increased dislocation density as a result of the intense plastic deformation and heat generated by the rotating FSP tool. DRX occurring during MPFSP produces very fine grains by fragmenting B4C and Al_2O_3 nanoparticles and helping migrate grain boundaries. This effect is especially pronounced in the six-pass FSP with B4C and Al_2O_3 nanoparticles, where extremely fine-grained structures emerge compared to those seen in fewer-pass FSPs or the base metal. The B4C and Al_2O_3 particles distribution, alongside the grain boundary migration, facilitates an even finer granular structure, which contributes significantly to hardness.

This refined grain structure and improved particle dispersion are further facilitated by DRX, where B4C and Al_2O_3 nanoparticles in the substrate alloy act as effective nucleation points for new grains. These nanoparticles, therefore, enable optimal formation of fine, equiaxed grains, particularly in the six-pass FSP, resulting in improved hardness properties (Ref 53, 54). This relationship is best explained by the Hall–Petch principle, where a refined grain size leads to an increase in strength and hardness by limiting dislocation movement. Consequently, with each additional FSP pass, the composite structure becomes increasingly resilient due to cumulative microstructural enhancements. The increase in microhardness within the SZ compared to the TMAZ and the HAZ can be attributed to particle size. The SZ benefits from the smallest particle size due to the consistent application of FSP, while microhardness diminishes in the TMAZ and HAZ because of precipitate behavior. Specifically, in the TMAZ, the decline in hardness is tied to the dissolution of precipitates, whereas, in the HAZ, it stems from precipitate coarsening. These changes reflect underlying transformations at the microstructural level that result in a hardness gradient across different regions of the composite. One of the notable observations is the coarsening of reinforcing precipitates in the HAZ, coupled with the dissolution of the Guinier–Preston zones, both of which contribute to a decrease in microhardness in this region (Ref 55). In contrast, MPFSP leads to a finer grain structure, superior particle dispersion, and improved reinforcement-matrix mixing within the SZ, leading to increased hardness properties.

3.4 Tribological Behavior

Figure 12 examines the changes in the coefficient of friction (COF) over sliding distances and as influenced by the number of FSP passes in AA2024/B_4C/Al_2O_3 nanocomposites. Figure 13 specifically illustrates COF fluctuations across various sliding distances, showing a clear trend where COF variability intensifies as the sliding distance increases. This effect is likely due to the fracture of oxide layers, which results from elevated temperatures generated by prolonged sliding. Shafiei et al. (Ref 56) also observed this phenomenon, noting that oxide fracture under high thermal conditions contributes to increased surface roughness,

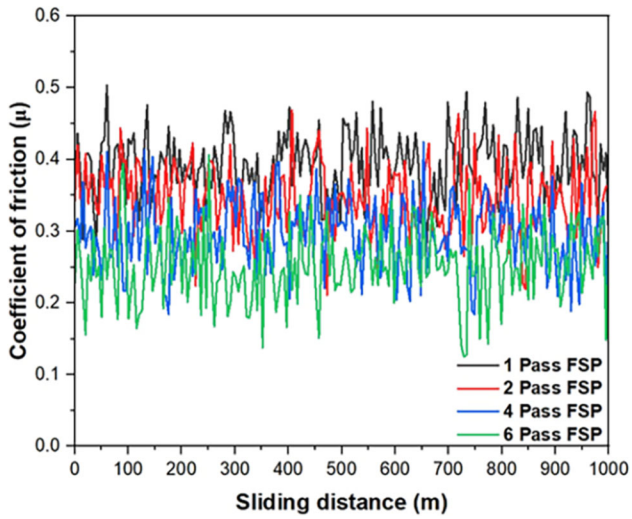


Fig. 12 Graphs showing COF versus sliding distance of the AA2024/B4C/Al_2O_3 nanocomposites for different FSP passes

signaling a shift from mild to more aggressive wear. This progression toward severe wear becomes particularly evident after a significant portion of the sliding distance has been covered. Figure 12 highlights the average COF (μ) for these nanocomposites, demonstrating a consistent decrease in COF as the FSP passes increases. This reduction in COF can be attributed to the effects of DRX and grain refinement induced by the FSP process, which causes SPD of the material. With each additional FSP pass, the grains become finer, leading to a reduction in frictional resistance. Furthermore, finer grain structures increase the hardness of the material, as described by the Hall–Petch relationship (Ref 57, 58), which states that smaller grains enhance material strength. This hardness increase, in turn, leads to a lower wear rate, a trend illustrated in Fig. 13. In Fig. 13, the wear depth is shown to be most significant in the nanocomposites subjected to only a single FSP pass. This greater wear depth indicates lower hardness and increased susceptibility to wear in single-pass samples. As the FSP passes increase, however, the wear depth diminishes substantially, which can be explained by the enhancement in hardness due to grain refinement and DRX. These findings are consistent with previous observations by Yuvaraj et al. (Ref 59), who noted that multiple FSP passes enhance the material's hardness and reduce wear depth. In nanocomposites subjected to a single FSP pass, the COF and wear rate are notably higher. The increased wear rate is due to adhesive wear mechanism involves the transfer of aluminum matrix material to the counter face, often due to the strong bonding between the composite pin and the disk surface, resulting in more material loss. This ultimately resulted in additional resistance, elevating both COF and wear rates.

Figure 15 presents SEM images depicting the worn surfaces of AA2024/B4C/Al_2O_2 nanocomposites after wear testing. The worn surfaces showed abrasive wear, oxidation, and delamination as the primary wear mechanism (Ref 4). In AMCs the abrasive wear mechanism occurs in aluminum surface composites due to the presence of hard reinforcement particles, which can act as abrasive agents and cause micro-cutting and plowing of the counter face, leading to material removal and wear. Concurrently, oxidation of the aluminum matrix and reinforcement phases takes place, forming a thin oxide layer that can delaminate under repeated sliding contact, resulting in material loss and exposure of fresh surfaces. Delamination of the oxide layer and the underlying material can also occur due to the thermal and mechanical stresses generated during sliding, leading to a complex wear mechanism that involves the degradation of both the matrix and reinforcement phases. Similar mechanisms are found in the previous studies (Ref 19, 60), where aluminum-based composites are subjected to MPFSP. The single-pass specimen in Fig. 15(a) has an oxide coating on its worn surface, possibly from frictional sliding heat. Frictional heat, especially under larger loads, raises nanocomposites' temperatures, promoting surface oxidation and oxidative wear (Ref 61). Under sliding wear, plastic deformation creates pressure and stress zones in the front and back. Deformation causes cracks below the surface to migrate outward, detaching material and forming wear debris. Oxide layer forms in some areas imply that debris interacts with the worn surface to cause oxidation wear (Ref 62). The single-pass worn surface (Fig. 15a) has wider grooves than the six-pass specimen (Fig. 15(d)), suggesting higher material loss during sliding wear. This shows that single-pass FSP has broader wear tracks than MPFSP because of decreased hardness and microstructure refinement. In contrast, Fig. 15(b) and (c) shows

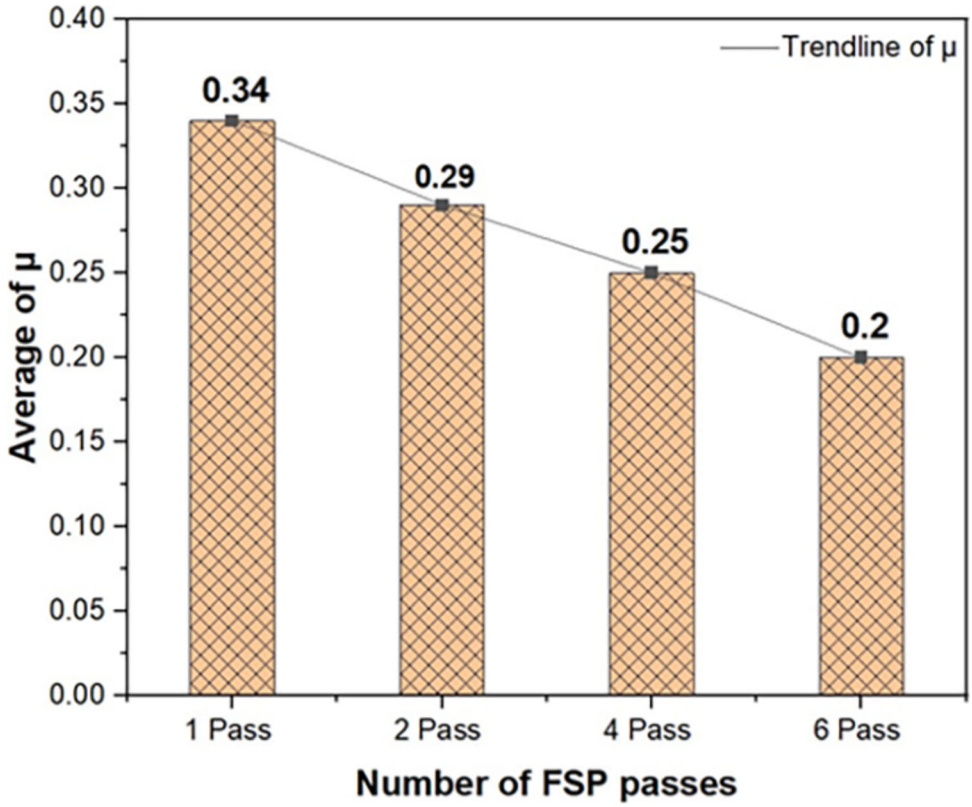


Fig. 13 Bar graph showing trendline of average coefficient of friction (μ) with FSP passes

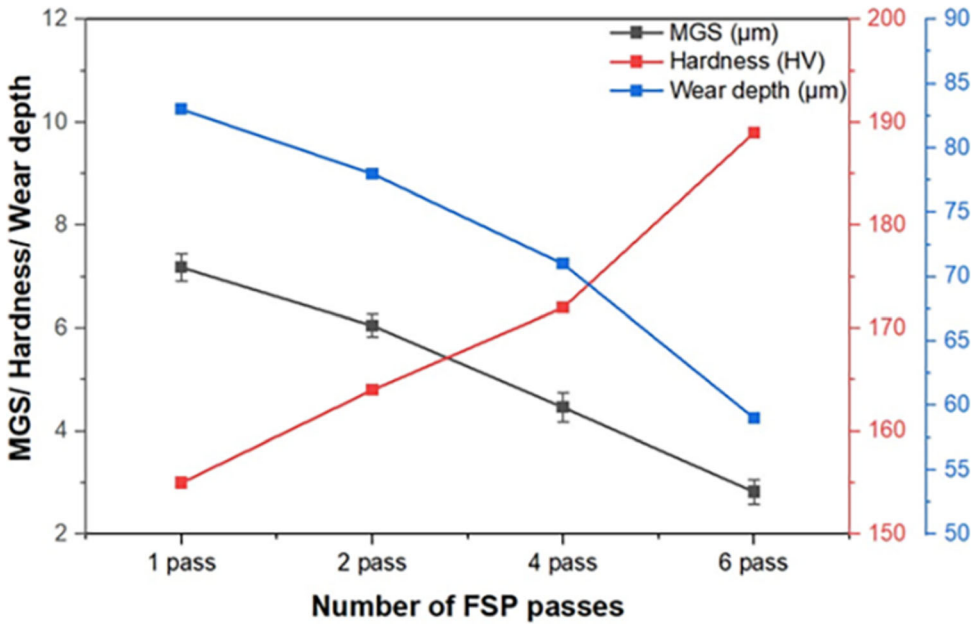


Fig. 14 Correlation between mean grain size (MGS), hardness, and wear depth with FSP passes

worn surfaces from distinct wear track sections in MPFSP specimens. More FSP passes increase wear resistance, as shown by finer wear grooves and less delamination than the single-pass sample. Figure 14(d) shows plowing and material delamination, but fewer deep wear grooves, suggesting less severe adhesive wear in multi-pass samples. In Fig. 15(d), particles distributed throughout the surface indicate oxidized areas, demonstrating oxidative wear in these nanocomposites. Comparing single-pass and six-pass FSP samples shows that more passes reduce groove depth and breadth. The six-pass sample had narrower grooves than the single-pass sample, showing less material was lost during wear testing. Hardness improvements from DRX and grain refining with repeated FSP reduce groove size and wear. Figure 16 shows elements

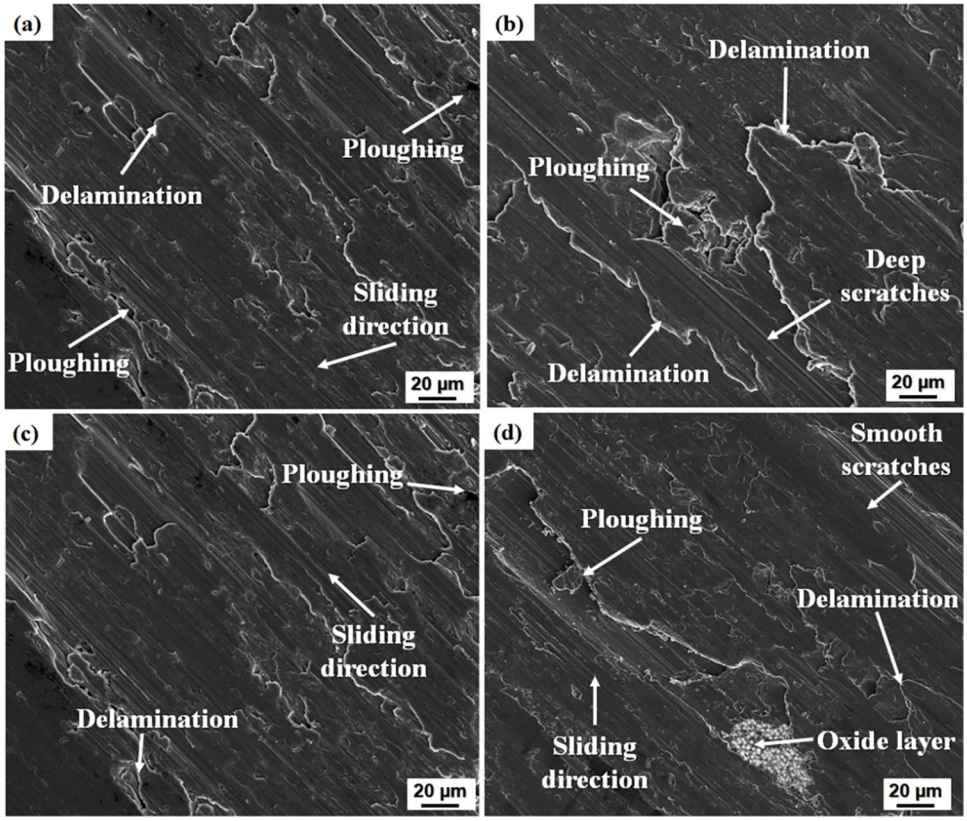


Fig. 15 SEM images of the worn-out surfaces of the AA2024/B_4C/Al_2O_3 nanocomposites for different FSP passes: (a) 1 Pass, (b) 2 Pass, (c) 4 Pass, and (d) 6 Pass

mapping of the worn-out composite surfaces manufactured at six-pass FSP, which include scattered particles. High concentrations of aluminum (Al-Kα) in red regions reflect the presence of AMCs. The gold color represents copper concentration (Cu-Kα), whereas the blue area suggests oxygen concentration (O-Kα), perhaps indicating aluminum oxides or other compounds containing oxygen and copper. In the matrix, blue patches represent the integration of boron (B-Kα)-based reinforcements. The presence of carbon (C-Kα) in the green region may indicate boron carbide reinforcements. The high Al (85.79%) in the EDS data indicates an aluminum-rich matrix. The elemental makeup includes oxygen (5.71%), boron (3.08%), carbon (2.83%), and copper (1.63%) after aluminum. The SEM picture shows a large oxide layer, which may explain the high oxygen levels. This oxide layer may improve composite wear resistance by preventing material deterioration.

The results of the present investigation are compared with prior studies (Ref 4, 20), where AA2024-based composites were fabricated through MPFSP. Compared to the findings of Moustafa (Ref 20), which reported a maximum ultimate tensile strength (UTS) of approximately 308 MPa, elongation of 2.4%, and hardness of 109 HV, with an average mean grain size (MGS) of 5.1 μm in the stir zone, the current study demonstrates a remarkable grain refinement of 82.1%, along with improvements in UTS by 64.6%, ductility by 29.16%, and hardness by 72.4%. Similarly, when compared to Acuna's study (Ref 4), which reported a maximum wear depth of 297 μm, the present research achieves a substantial reduction in wear depth by approximately 80%. The improved mechanical and tribological performance can be attributed to multiple factors (Ref 63). Firstly, the increased number of FSP passes facilitates extensive grain refinement and fragmentation of reinforcement particles due to the pinning effect, which limits grain growth by reducing the heat input during FSP. Secondly, the thermal expansion mismatch between the matrix and reinforcement particles generates a high dislocation density at the reinforcement-matrix interfaces, activating the Orowan strengthening mechanism. Lastly, the intrinsic hardness of the reinforcement phases significantly contributes to the improved overall hardness. The findings of the present study not only validate the effectiveness of the fabricated hybrid composites through employing MPFSP but also the advancement in the fidelity of AMCs to achieve superior performance.

4. Conclusions

The influence of multi-pass friction stir processing (MPFSP) on the microstructure, mechanical and tribological performance of AA2024-T351 surface composites reinforced with 50% B_4C and 50% Al_2O_3 nanoparticles was effectively investigated at a tool rotation speed of 800 rpm and a traverse speed of 35 mm/min. The findings may be summarized as follows:

- The microstructural analysis revealed that MPFSP effectively induced significant grain refinement and homogeneous distribution of reinforcement particles. Notably, the mean grain size (MGS) in the stir region decreased substantially from 7.18 ± 0.26 to 2.82 ± 0.06 μm after six passes,

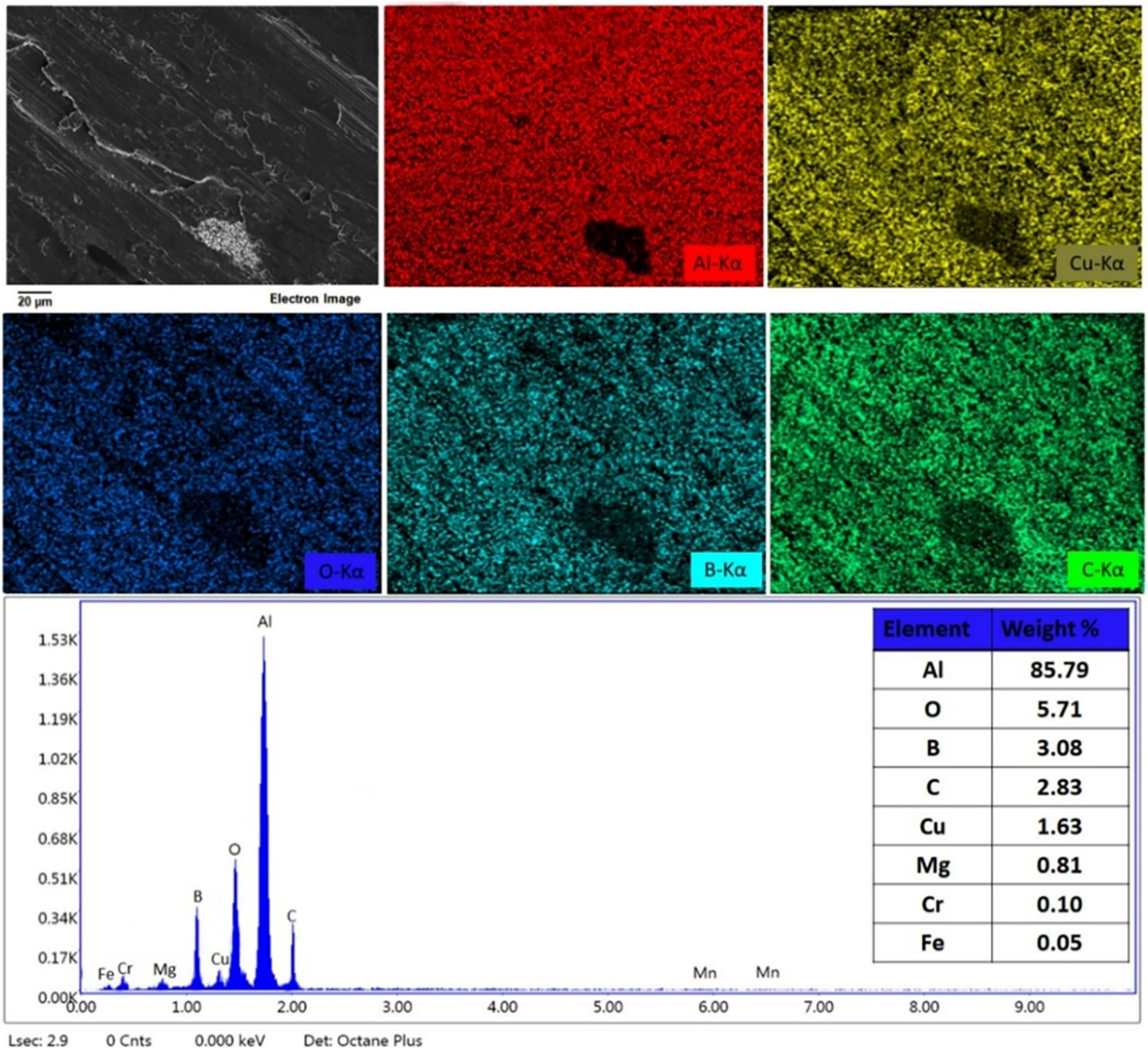


Fig. 16 Elemental mapping with EDS report of the worn-out surface of the nanocomposite with six pass FSP

compared to one pass. Furthermore, XRD results showed a consistent reduction in peak intensity of α-Al, Al_2CuMg, B_4C, and Al_2O_3 phases with increasing number of passes, indicating fragmentation of secondary phase particles.

- Electron backscatter diffraction (EBSD) results revealed a notable increase in the fraction of high-angle grain boundaries (HAGBs) from 76.49 to 95.21% after six passes of FSP. This significant rise in fraction of HAGBs effectively hindered grain boundary dislocation motion, thereby providing substantial resistance to deformation under tensile loading, ultimately yielding improved tensile performance.
- The base alloy AA2024-T351 exhibited an ultimate tensile strength (UTS) of 421 MPa, a yield strength (YS) of 307 MPa, and an elongation (El) of 19.1%. However, the hybrid composite showed significant improvements in tensile properties after MPFSP. Specifically, the UTS increased from 469 to 507 MPa, the YS increased from 327 to 360 MPa, while the El decreased from 6.4 to 3.1% after six passes of FSP, compared to one pass.
- The results demonstrated that increasing the number of FSP passes from one to six led to significant improvements in the hardness and tribological properties of the hybrid composite. Specifically, the hardness increased by 22% from 155 to 189 HV, while the coefficient of friction (COF) decreased by 41% from 0.34 to 0.20, and the wear rate decreased by 29% from 83 to 59 μm. These enhancements in surface properties can be attributed to the formation of ultrafine grain structures and the uniform dispersion of reinforcements within the matrix.
- The tensile fractured surfaces revealed a predominantly ductile mode of failure, characterized by the presence of dimples and shear lips. In contrast, the worn surface morphology exhibited a complex wear mechanism, involving abrasive wear, oxidation, and delamination, which were likely caused by the syner-

gistic effects of thermal softening and repeated sliding cycles, leading to material degradation and removal.

Acknowledgments

The authors would like to thank the CRF, Indian Institute of Technology Kharagpur, West Bengal, for facilitating the EBSD characterization and the Precision Manufacturing Lab, Delhi Technological University, Delhi, for assisting with the wear characterization of the composite samples.

Conflict of interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. G. Huang, J. Wu, W. Hou, L.H. Ahmad Shah, A.R.H. Midawi, F. Cao, A. Gerlich, Y. Shen, and F. Meng, A Novel Two-Step Method to Prepare Fine-Grained SiC/Al-Mg-Sc-Zr Nanocomposite: Processing, Microstructure and Mechanical Properties, *Mater. Sci. Eng. A*, 2021, **823**, p 141764. https://doi.org/10.1016/j.msea.2021.141764
2. A. Mazahery, H. Abdizadeh, and H.R. Baharvandi, Development of High-Performance A356/nano-Al_2O_3 Composites, *Mater. Sci. Eng. A*, 2009, **518**, p 61–64. https://doi.org/10.1016/j.msea.2009.04.014
3. A. Sharma, V.M. Sharma, S. Mewar, S.K. Pal, and J. Paul, Friction Stir Processing of Al6061-*SiC*-Graphite Hybrid Surface Composites, *Mater. Manuf. Processes*, 2018, **33**, p 795–804. https://doi.org/10.1080/10426914.2017.1401726
4. R. Acuña, M.J. Cristóbal, C.M. Abreu, and M. Cabeza, Microstructure and Wear Properties of Surface Composite Layer Produced by Friction Stir Processing (FSP) in AA2024-T351 Aluminum Alloy, *Metall. Mater. Trans. A*, 2019, **50**, p 2860–2874. https://doi.org/10.1007/s11661-019-05172-6
5. M. Gui, S.B. Kang, and K. Euh, Influence of Spraying Conditions on Microstructures of Al-SiC p Composites by Plasma Spraying, *Metall. and Mater. Trans. A.*, 2005, **36**, p 2471–2480. https://doi.org/10.1007/s11661-005-0121-4
6. M. Bashirzadeh, F. Azarmi, C.P. Leither, and G. Karami, Investigation on Relationship Between Mechanical Properties and Microstructural Characteristics of Metal Matrix Composites Fabricated by Cold Spraying Technique, *Appl. Surf. Sci.*, 2013, **275**, p 208–216. https://doi.org/10.1016/j.apsusc.2012.12.166
7. R. Butola, L. Tyagi, R.M. Singari, Q. Murtaza, H. Kumar, and D. Nayak, Mechanical and Wear Performance of Al/SiC Surface Composite Prepared Through Friction Stir Processing, *Mater Res Express.*, 2021, **8**, p 016520. https://doi.org/10.1088/2053-1591/abd89d
8. B. Bagheri and M. Abbasi, Development of AZ91/SiC Surface Composite by FSP: Effect of Vibration and Process Parameters on Microstructure and Mechanical Characteristics, *Adv. Manuf.*, 2020, **8**, p 82–96. https://doi.org/10.1007/s40436-019-00288-9
9. M.S. Bin Reyaz and A.N. Sinha, Fabrication of Tungsten Inert Gas-Welded Dissimilar Aluminium Joints Using Different Friction Stir Processing Tools Through Response Surface Method, *Adv. Mater. Process. Technol.*, 2024 https://doi.org/10.1080/2374068X.2024.2341528
10. M.S. Reyaz and A.N. Sinha, Modeling the Effects of Tool Pin Configurations and Friction Stir Processing Parameters on Tungsten Inert Gas Welded Dissimilar Aluminum Alloy Joints, *J. Mater. Eng. Perform.*, 2024 https://doi.org/10.1007/s11665-024-09444-0
11. M.S. Bin Reyaz and A.N. Sinha, Investigation on the Influence of FSP Parameters on TIG Welded Dissimilar Aluminum Joints Using RSM and Desirability Function Approach, *Proc. Inst. Mech. Eng. C J. Mech. Eng. Sci.*, 2024 https://doi.org/10.1177/09544062241231165
12. M. Sarkari Khorrami, M. Kazeminezhad, and A.H. Kokabi, The effect of SiC nanoparticles on the friction stir processing of severely deformed aluminum, *Mater. Sci. Eng. A*, 2014, **602**, p 110–118. https://doi.org/10.1016/j.msea.2014.02.067
13. E.R.I. Mahmoud, M. Takahashi, T. Shibayanagi, and K. Ikeuchi, Wear Characteristics of Surface-Hybrid-MMCs Layer Fabricated on Aluminum Plate by Friction Stir Processing, *Wear*, 2010, **268**, p 1111–1121. https://doi.org/10.1016/j.wear.2010.01.005
14. M. Kianezhad and A. Honarbakhsh Raouf, Effect of Nano-Al_2O_3 Particles and Friction Stir Processing on 5083 TIG Welding Properties, *J. Mater. Process. Technol.*, 2019, **263**, p 356–365. https://doi.org/10.1016/j.jmatprotec.2018.08.010
15. Q. Huang, R. He, C. Wang, and X. Tang, Microstructure, Corrosion and Mechanical Properties of TiC Particles/Al-5Mg Composite Fillers for Tungsten Arc Welding of 5083 Aluminum Alloy, *Materials.*, 2019, **12**, p 3029. https://doi.org/10.3390/ma12183029
16. R. Ande, P. Gulati, D. Kumar Shukla, and H. Dhingra, Microstructural and Wear Characteristics of Friction Stir Processed Al-7075/SiC Reinforced Aluminium Composite, *Mater Today Proc.*, 2019, **18**, p 4092–4101. https://doi.org/10.1016/j.matpr.2019.07.353
17. D. Aruri, K. Adepu, K. Adepu, and K. Bazavada, Wear and Mechanical Properties of 6061–T6 Aluminum Alloy Surface Hybrid Composites [(SiC+Gr) and (SiC+Al_2O_3)] Fabricated by Friction Stir Processing, *J. Market. Res.*, 2013, **2**, p 362–369. https://doi.org/10.1016/j.jmrt.2013.10.004
18. K.S. Sandhu, H. Singh, and G. Singh, Influence of Sub-Zero Coolant Circulation and Additive Nanoparticles on Performance Characteristics of FSPed Al6061 Alloy, *Aust. J. Mech. Eng.*, 2024 https://doi.org/10.1080/14484846.2024.2400804
19. D. Ghanbari, M. Kasiri Asgarani, K. Amini, and F. Gharavi, Influence of heat Treatment on Mechanical Properties and Microstructure of the Al2024/SiC Composite Produced by Multi–pass Friction Stir Processing, *Measurement*, 2017, **104**, p 151–158. https://doi.org/10.1016/j.measurement.2017.03.024
20. E. Moustafa, Effect of Multi-Pass Friction Stir Processing on Mechanical Properties for AA2024/Al_2O_3 Nanocomposites, *Materials.*, 2017, **10**, p 1053. https://doi.org/10.3390/ma10091053
21. H. Rana and V. Badheka, Elucidation of the Role of Rotation Speed and Stirring Direction on AA 7075-B_4 C Surface Composites Formulated by Friction Stir Processing, *Proc. Inst. Mech. Eng., Part L: J. Mater.: Des. Appl.*, 2019, **233**, p 977–994. https://doi.org/10.1177/1464420717736548
22. S. Kumar and A.N. Sinha, Microstructure and Mechanical Properties of Aluminum Surface Composites Fabricated Via Multi-pass FSP with B_4C and Al_2O_3 Reinforcements, *J. Adhes. Sci. Technol.Adhes Sci. Technol.*, 2025 https://doi.org/10.1080/01694243.2025.2464046
23. U. Pranavi, P. Venkateshwar Reddy, S. Venukumar, and M. Cheepu, Evaluation of Mechanical and Wear Properties of Al 5059/B4C/Al_2O_3 Hybrid Metal Matrix Composites, *J. Compos. Sci.*, 2022, **6**, p 86. https://doi.org/10.3390/jcs6030086
24. S.J. James, M. Ganesan, P. Santhamoorthy, and P. Kuppan, Development of Hybrid Aluminium Metal Matrix Composite and Study of Property, *Mater. Today Proc.*, 2018, **5**, p 13048–13054. https://doi.org/10.1016/j.matpr.2018.02.291
25. K. Giridharan, P. Sevvel, R. Ramadoss, and B. Stalin, Friction Stir Processing of Nanofiller Assisted AISI 1010 Steel-CDA 101 Copper Dissimilar Welds: A Strength Factor Approach, *Metall. Res. Technol.*, 2022, **119**, p 505. https://doi.org/10.1051/metal/2022065
26. M.V. Sezhian, G. Chakravarthi, K. Giridharan, B. Stalin, B.Y. Kumar, P. Sureshkumar, J. Vairamuthu, and R. Krishnaraj, Investigation of Friction Stir-Welded B4C Particles-Reinforced Copper Joint: Mechanical, Fatigue, and Metallurgical Properties, *Adv. Mater. Sci. Eng.*, 2022, **2022**, p 1–11. https://doi.org/10.1155/2022/1667041
27. J. Gunasekaran, P. Sevvel, I. John Solomon, and J. Vasanthe Roy, Multi Objective Optimization of Parameters During FSW of AZ80A - AZ31B Mg Alloys Using Grey Relational Analysis, *J. Mech. Sci. Technol.*, 2024, **38**, p 4971–4982. https://doi.org/10.1007/s12206-024-0832-3
28. J. Solomon, P. Sevvel, J. Gunasekaran, and S.J. Rajarajan, Establishment of Empirical Relations Amidst Mechanical Attributes of Friction Stir Welded Distinctive Alloys of Mg and Optimized Process Parameters, *Mater. Res. Express.*, 2023, **10**, p 066502. https://doi.org/10.1088/2053-1591/acd5ae
29. M.S. Reyaz, A.N. Sinha, H. Mehdi, and Q. Murtaza, Effect of Pulsed TIG Welding Parameters on the Microstructural Evolution and Mechanical Properties of Dissimilar AA6061-T6 and AA7075-T6 Weldments, *Arab. J. Sci. Eng.*, 2023 https://doi.org/10.1007/s13369-023-08563-5
30. M.S. Reyaz and A.N. Sinha, Effect of Heat Input on Microstructure and Mechanical Properties of Automated Tungsten Inert Gas-Welded

Dissimilar AA6061-T6 and AA7075-T6 Joints, *J. Mater. Eng. Perform.*, 2023 https://doi.org/10.1007/s11665-023-09026-6
31. R.S. Mishra and Z.Y. Ma, Friction Stir Welding and Processing, *Mater. Sci. Eng. R. Rep.*, 2005, **50**, p 1–78. https://doi.org/10.1016/j.mser.2005.07.001
32. M.S. Reyaz and A.N. Sinha, Improving Tensile Strength, Impact Toughness and Wear Resistance of TIG Welded Dissimilar AA7075-T6/AA6061-T6 Joints Through Friction Stir Processing, *Int. J. Interact. Des. Manuf. (IJIDeM).*, 2024 https://doi.org/10.1007/s12008-023-01691-0
33. M.S. Reyaz and A.N. Sinha, Parametric Optimization and Characterization of AA6061-T6 and AA7075-T6 Tungsten Inert Gas Welding Joints Subjected to Friction Stir Processing, *J. Mater. Eng. Perform.*, 2024 https://doi.org/10.1007/s11665-023-09111-w
34. A. Sharma, Y. Morisada, and H. Fujii, Interfacial Microstructure and Strengthening Mechanisms of SPSed Al/GNP Nanocomposite Subjected to Multi-pass Friction Stir Processing, *Mater CharactCharact.*, 2023, **197**, 112652. https://doi.org/10.1016/j.matchar.2023.112652
35. M. Paidar, O.O. Ojo, H.R. Ezatpour, and A. Heidarzadeh, Influence of Multi-pass FSP on the Microstructure, Mechanical Properties and Tribological Characterization of Al/B4C Composite Fabricated by Accumulative Roll Bonding (ARB), *Surf. Coat. Technol.*, 2019, **361**, p 159–169. https://doi.org/10.1016/j.surfcoat.2019.01.043
36. C.I. Chang, C.J. Lee, and J.C. Huang, Relationship Between Grain Size and Zener-Holloman Parameter During Friction Stir Processing in AZ31 Mg Alloys, *Scr. Mater.*, 2004, **51**, p 509–514. https://doi.org/10.1016/j.scriptamat.2004.05.043
37. R. Senthilkumar, M. Prakash, N. Arun, and A.A. Jeyakumar, The Effect of the Number of Passes in Friction Stir Processing of Aluminum alloy (AA6082) and its Failure Analysis, *Appl. Surf. Sci.*, 2019, **491**, p 420–431. https://doi.org/10.1016/j.apsusc.2019.06.132
38. M. Srivastava, S. Rathee, A.N. Siddiquee, and S. Maheshwari, Investigation on the Effects of Silicon Carbide and Cooling Medium during Multi-Pass FSP of Al-Mg/ SiC Surface Composites, *SILICON*, 2019, **11**, p 2149–2157. https://doi.org/10.1007/s12633-018-0037-4
39. M. Manjunath, S. Subramanian Pillappan, and S. Kurse, Process Parameter Analysis of Double Pass Friction Stir Composite Welds, *Mater. Manuf. Processes*, 2023 https://doi.org/10.1080/10426914.2023.2190380
40. Y. Chen, H. Ding, J. Li, Z. Cai, J. Zhao, and W. Yang, Influence of Multi-pass Friction Stir Processing on the Microstructure and Mechanical Properties of Al-5083 Alloy, *Mater. Sci. Eng. A*, 2016, **650**, p 281–289. https://doi.org/10.1016/j.msea.2015.10.057
41. H. Mehdi, A. Mehmood, A. Chinchkar, A.W. Hashmi, C. Malla, and P. Mohapatra, Optimization of Process Parameters on the Mechanical Properties of AA6061/Al_2O_3 Nanocomposites Fabricated by Multi-pass Friction Stir Processing, *Mater. Today Proc.*, 2022 https://doi.org/10.1016/j.matpr.2021.11.333
42. M.S. Bin Reyaz and A.N. Sinha, Exploring the Influence of Pin Geometry on Metallurgical and Mechanical Characteristics of Dissimilar Aluminum TIG Welds Via Friction Stir Processing, *J. Adhes. Sci. Technol.Adhes. Sci. Technol.*, 2024, **38**, p 1–24. https://doi.org/10.1080/01694243.2024.2335306
43. M. Raturi and A. Bhattacharya, Microstructure and Texture Correlation of Secondary Heating Assisted Dissimilar Friction Stir Welds of Aluminum Alloys, *Mater. Sci. Eng. A*, 2021, **825**, p 141891. https://doi.org/10.1016/j.msea.2021.141891
44. A.J. Ansari and M. Anas, Effect of Varying Volume Fraction of Reinforcement on the Microstructure, Hardness and Tribological Behaviour of Aluminium Metal Matrix Composites Developed Through FSP Technique, *J. Inst. Eng. (India): Ser. D.*, 2024, **105**, p 263–273. https://doi.org/10.1007/s40033-023-00476-9
45. S.B. Reyaz and S.K. Parida, About the Design, Development of an Automated SMAW Setup and the Weldment Analysis in Air and Wet Working Environment, *IOP Conf. Ser. Mater. Sci. Eng.*, 2019, **653**, 012039. https://doi.org/10.1088/1757-899X/653/1/012039
46. M. Khalili, A.R. Eivani, S.H. Seyedein, and H.R. Jafarian, The Effect of Multi-pass Friction Stir Processing on Microstructure and Mechanical Properties of Dual-Phase Brass Alloy, *J. Market. Res.*, 2022, **21**, p 1177–1195. https://doi.org/10.1016/j.jmrt.2022.09.069
47. H. Mehdi and R.S. Mishra, Consequence of Reinforced SiC Particles on Microstructural and Mechanical Properties of AA6061 Surface Composites by Multi-pass FSP, *J. Adhes. Sci. Technol.Adhes. Sci. Technol.*, 2022, **36**, p 1279–1298. https://doi.org/10.1080/01694243.2021.1964846
48. Y. Xie, X. Meng, Y. Chang, D. Mao, Y. Yang, Y. Xu, L. Wan, and Y. Huang, Ameliorating Strength-Ductility Efficiency of Graphene Nanoplatelet-Reinforced Aluminum Composites Via Deformation-Driven Metallurgy, *Compos. Sci. Technol.*, 2022, **219**, 109225. https://doi.org/10.1016/j.compscitech.2021.109225
49. D. Mao, Y. Xie, X. Meng, X. Ma, Z. Zhang, X. Sun, L. Wan, K. Volodymyr, and Y. Huang, Strength–Ductility Materials by Engineering a Coherent Interface at Incoherent Precipitates, *Mater. Horiz.Horiz.*, 2024, **11**, p 3408–3419. https://doi.org/10.1039/D4MH00139G
50. D. Mao, X. Ma, Y. Xie, X. Meng, N. Wang, Z. Zhang, X. Sun, and Y. Huang, In-situ Solid-State Deformation-Driven Rapid Reaction Towards Higher Strength-Ductility Al-CuO Composites, *Compos. Part A Appl. Sci. Manuf.*, 2024, **182**, 108174. https://doi.org/10.1016/j.compositesa.2024.108174
51. M. Singh, K. Goyal, and D. Bhandari, Influence of TiO_2 and Y_2O_3 Nanoparticles on Mechanical Properties of Aluminium Matrix Hybrid Nanocomposites Fabricated by Vacuum Die Casting, *World J. Eng.*, 2024, **21**, p 837–850. https://doi.org/10.1108/WJE-01-2023-0032
52. A. Garg, M. Raturi, and A. Bhattacharya, Strength, Failure and Microstructure Development for Friction Stir Welded AA6061-T6 Joints with Different Tool Pin Profiles, *CIRP J. Manuf. Sci. Technol.*, 2020, **29**, p 99–114. https://doi.org/10.1016/j.cirpj.2020.03.001
53. H.A. Deore, A. Bhardwaj, A.G. Rao, J. Mishra, and V.D. Hiwarkar, Consequence of Reinforced SiC Particles and post process artificial ageing on Microstructure and Mechanical Properties of Friction Stir Processed AA7075, *Def. Technol.*, 2020, **16**, p 1039–1050. https://doi.org/10.1016/j.dt.2019.12.001
54. S.S. Mirjavadi, M. Alipour, A.M.S. Hamouda, A. Matin, S. Kord, B.M. Afshari, and P.G. Koppad, Effect of Multi-pass Friction Stir Processing on the Microstructure, Mechanical and Wear Properties of AA5083/ZrO2 Nanocomposites, *J. Alloys Compd.*, 2017, **726**, p 1262–1273. https://doi.org/10.1016/j.jallcom.2017.08.084
55. T. Azimzadegan and S. Serajzadeh, An Investigation into Microstructures and Mechanical Properties of AA7075-T6 during Friction Stir Welding at Relatively High Rotational Speeds, *J. Mater. Eng. Perform.*, 2010, **19**, p 1256–1263. https://doi.org/10.1007/s11665-010-9625-1
56. A. Shafiei-Zarghani, S.F. Kashani-Bozorg, and A.Z. Hanzaki, Wear Assessment of Al/Al_2O_3 Nano-Composite Surface Layer Produced Using Friction Stir Processing, *Wear*, 2011, **270**, p 403–412. https://doi.org/10.1016/j.wear.2010.12.002
57. M.S. Bin Reyaz and A.N. Sinha, An Experimental Investigation on Mechanical Characteristics and Wear Behaviour of TIG Welded Dissimilar Aluminum Alloys, *J. Adhes. Sci. Technol.Adhes. Sci. Technol.*, 2023, **38**(8), p 1–36. https://doi.org/10.1080/01694243.2023.2251782
58. M.S. Bin Reyaz and A.N. Sinha, Analysis of Mechanical Properties and Optimization of Tungsten Inert Gas Welding Parameters on Dissimilar AA6061-T6 and AA7075-T6 by a response Surface Methodology-Based Desirability Function Approach, *Eng. Optim.Optim.*, 2023, **56**(12), p 1–32. https://doi.org/10.1080/0305215X.2023.2230133
59. N. Yuvaraj and S. Aravindan, Vipin: Wear Characteristics of Al5083 Surface Hybrid Nano-composites by Friction Stir Processing, *Trans. Indian Inst. Met.*, 2017, **70**, p 1111–1129. https://doi.org/10.1007/s12666-016-0905-9
60. H.S. Arora, H. Singh, and B.K. Dhindaw, Wear behaviour of a Mg alloy subjected to friction stir processing, *Wear*, 2013, **303**, p 65–77. https://doi.org/10.1016/j.wear.2013.02.023
61. C.N. Shyam Kumar, R. Bauri, and D. Yadav, Wear Properties of 5083 Al–W Surface Composite Fabricated by Friction Stir Processing, *Tribol. Int.. Int.*, 2016, **101**, p 284–290. https://doi.org/10.1016/j.triboint.2016.04.033
62. L.H. Mair, T.A. Stolarski, R.W. Vowles, and C.H. Lloyd, Wear: Mechanisms, Manifestations and Measurement Report of a workshop, *J. Dent.*, 1996, **24**, p 141–148. https://doi.org/10.1016/0300-5712(95)00043-7
63. A. Benamor, N. Chiker, M. Salhi, A. Haddad, Y. Hadji, H. Benamor, B. Hamzaoui, R. Badji, and M. Hadji, Microstructure, Hardness, and Wear Resistance of AA5083/Ti3SiC2 Composites Fabricated Via Friction Stir Processing, *Int. J. Adv. Manuf. Technol.*, 2024, **130**, p 5979–5998. https://doi.org/10.1007/s00170-024-13051-7

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Materials Research Proceedings 49 (2025) 29-38 https://doi.org/10.21741/9781644903438-4

Exploring multi-material structures: An overview of arc additive manufacturing

Dhruv Kumar[1a*], Vipin[1,b], Reeta Wattal[1,c]

[1]Department of Mechanical Engineering, Delhi Technological University, Delhi, India – 110042

[a]dhruvsingh.nitw12@gmail.com, [b]vipin@dce.ac.in, [c]reetawattal@dtu.ac.in

Keywords: WAAM, Bi-Metallic Materials, Additive Manufacturing and Functionally Graded Materials

Abstract. Arc welding-based additive manufacturing techniques offer promising prospects for metal additive manufacturing, particularly in the economic production of large-sized components with relatively high deposition rates. This review article provides an introduction to wire arc additive manufacturing (WAAM) and offers an overview of its fundamental principles. Additionally, it delves into the fabrication of bi-metallic structures and functionally graded materials/structures using WAAM, exploring their applications across various industries.

1. Introduction

Wire Arc Additive Manufacturing (WAAM) is a promising and more popular technique for creating metal objects using additive manufacturing. The procedure employs an electric arc as the heat source and filler wires as the material for deposition, enabling the creation of medium to large-scale structures using a step-by-step layering method. The WAAM technique has proven effective in producing structures using various alloys like aluminum, nickel, titanium, and steel [1]. Weldability plays a crucial role in governing the production of metallic structures using WAAM, given its reliance on stacking successive layers of weld beads. Functional Graded Materials (FGMs) represent an innovative and evolving category of high-strength materials ideal for diverse industrial applications, enabling varied functional performances within a single part. Traditionally, FGMs are crafted through post-processing methods such as fusion welding techniques, diffusion bonding, or friction stir welding [2].

Functionally graded materials (FGMs) differ from traditional homogeneous materials in that their composition and microstructure gradually change in spatial direction, leading to continuous or discontinuous alteration of mechanical, physical, and chemical properties along single or multiple directions [3]. FGMs possess distinctive features including enhanced mechanical properties, resistance to heat shock, and corrosion. Currently, FGMs find extensive applications in aerospace, medical, optoelectronics, and energy sectors [4]. Various methods are employed to fabricate FGMs, including centrifugal casting, discharge plasma sintering, vapor deposition, and self-propagating synthesis [5]. Nonetheless, traditional deposition techniques for FGMs suffer from several drawbacks, such as limited capability to produce complex structures and insufficient efficiency. SS321 and Inconel 625 are commonly used alloys in the fabrication of bi-metallic structures for aircraft, power generation, gas turbines, high-temperature bolts, petrochemical, and marine applications. This is due to their outstanding corrosion resistance and mechanical properties, particularly at temperatures below 850°C [1]. Utilizing the WAAM process for manufacturing FGMs offers the potential to expedite the creation of new materials due to the ability to incorporate varying compositions and mechanical properties.

Additive manufacturing (AM) has experienced rapid advancement recently and is poised to play a significant role in the future [2]. Furthermore, AM techniques offer a novel approach to fabricating FGMs. In contrast to traditional subtractive manufacturing methods, AM employs innovative processing techniques to create parts using computer-aided geometric models,

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 license. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. Published under license by Materials Research Forum LLC.

constructing them layer by layer from the bottom up. AM not only enables the direct fabrication of intricate structural components but also enhances production efficiency. By altering alloy composition, process parameters, and stacking paths, AM methods can produce materials with gradients in one or multiple dimensions. Laser AM, electron beam AM, and wire arc additive manufacturing (WAAM) are among the AM technologies employed, widely utilized for producing functionally gradient metals such as aluminum and titanium alloys [3].

2. Bi-metallic structures

Utilizing the wire arc additive manufacturing (WAAM) technique to deposit two metallic wires enables the creation of bimetallic structures (BMS), which can optimize the individual properties of each metal while mitigating their respective drawbacks. The aim of manufacturing BMS is to meet the specific property requirements at designated locations within parts and to reduce costs compared to fabricating entire components using a single, high-cost material. Ahsen et al. [6, 7] fabricated a bimetallic additively manufactured structure of ER70S6, a low carbon steel and SS 316L using a GMAW based WAAM and studied the effect of heat treatment on mechanical properties. The heat treatment technique resulted in a 35% improvement in ultimate tensile strength and a 250% increase in elongation compared to the as-deposited structure. An abrupt rise in hardness was found at the interface of the manufactured bimetallic material. The interface resulting from depositions was determined to be defect-free. The analysis of the bimetallic structure's microstructure showed predominantly dendritic grains with a minor presence of equiaxed grains. The hardness test findings indicated a significant and sudden rise in the hardness value at the bimetallic interface. The samples subjected to tensile tests exhibited a ductile fracture and failed in the SS 904L region, displaying fine dimples and ripping ridges.

Wenhao et al. [8] studied the effect of deposition sequence of 316L and Inconel 625 on the microstructure and mechanical properties of the fabricated bimetallic structures. The results showed no effect of the deposition sequence on the microstructure types. The interface was bonded well but due to element diffusion and Laves phase there were cracks at the interface. It was also revealed that hardness at the interface was within the hardness range of the two used alloys. Yinbao et al. [9] studied the effects of subsequent deposited layers of aluminium alloy on the microstructure characterization of the reaction layers of the Ti-Al bimetallic structure fabricated using cold metal transfer based WAAM. $TiAl_3$,TiAl and Ti_3Al were the Ti-Al intermetallic compounds formed in the subsequent layers.

3. Functionally graded materials/structures

Functionally graded materials (FGMs) are those materials in which chemical composition and microstructure changes gradually which leads to have different properties in one or different directions. The variation in chemical compositions and microstructures throughout the structure results in variation of mechanical properties, which makes the functionally graded materials a suitable candidate for the wide range of applications like aerospace engineering, nuclear power generation, sensor, biomedical implants, optoelectronic devices, and absorption systems [10]. There are several conventional or traditional methods which has been used to manufacture functionally graded materials, but it is somewhat difficult to make with significant composition changes. Conventional methods are not suitable to manufacture FGMs with high structural strength and thus require post treatment, which makes the process costly and uneconomical [11].

Recently, Additive manufacturing has evolved as a better option to manufactured functionally graded materials with different combinations of materials, with a wide scope of eliminating disadvantages of the traditional methods. Additive manufacturing methods to manufacture FGMs include Laser AM, electron beam AM and wire arc additive manufacturing technologies. Meanwhile WAAM has the advantages of low cost, easy operation, high deposition rate and the ability to deposit large-sized components [5]. The depicted WAAM system in **Figure 1** is

Materials Research Proceedings 49 (2025) 29-38 https://doi.org/10.21741/9781644903438-4

employed for the production of functionally graded titanium-aluminide alloys. It comprises a gas tungsten arc welding (GTAW) torch and power supply, a trailing argon shielding mechanism, a double-wire feeder, and a preheating system.

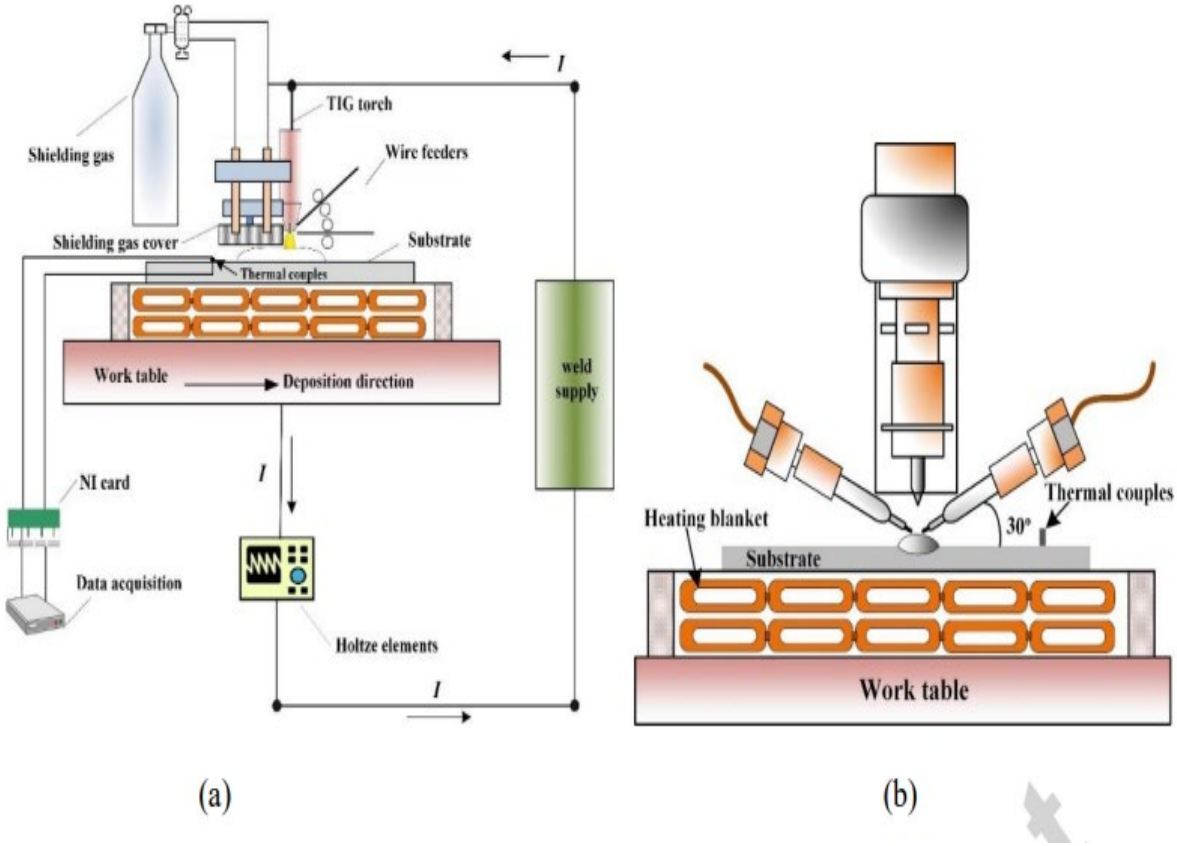


Figure 1 *Schematic diagram of* ***(a)*** *WAAM process* ***(b)*** *Double-wire feeding [5]*

3.2 Methods of fabrication of FGMs:

FGMs' processing methods are broadly categorized into two groups based on constructive processing and mass transport principles [10]. **Figure 2** depicts the various fabrications methods of FGMs. The constructive processing approach involves the deliberate development of gradients in space to build FGMs layer by layer. This approach provides the benefit of generating an infinite amount of gradients. In contrast, within the mass transport category, gradients within components are dependent on natural transport phenomena such as fluid movement, atomic diffusion, or heat conduction. Advancements in automation technology in the last twenty years have made constructive gradation procedures both technologically and economically feasible.

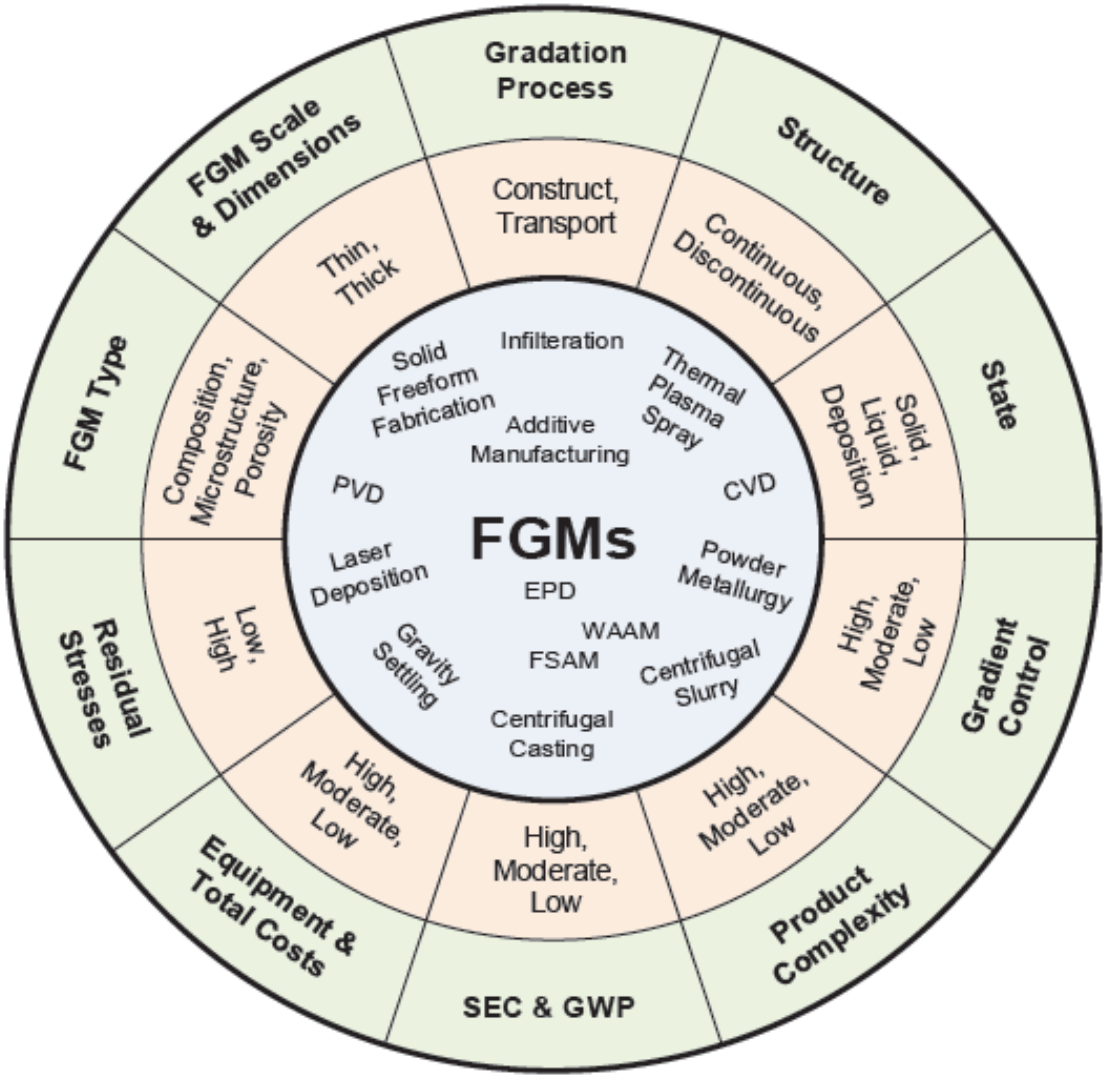


Figure 2. *Methods of fabrication of FGMs [11]*

3.2.1 Vapor deposition technique

There are several vapour deposition (VDT) methods accessible, such as plasma-enhanced CVD, sputter deposition, chemical vapour deposition (CVD), and physical vapour deposition (PVD). Vapour materials are converted to solid solids by condensation, chemical reactions, or other processes during vapour deposition [10-11]. By using these techniques, coatings that modify substrates' mechanical, electrical, thermal, optical, corrosion-resistant, and wear characteristics can be produced. Functionally graded surface coatings, which have good microstructure for thin surface coatings, are deposited using vapour deposition techniques. Although vapour deposition methods consume a lot of energy, they also produce toxic gases as a byproduct.

3.2.2 Centrifugal Casting

Centrifugal casting (CS) utilizes gravitational force by rotating the mold to produce functionally graded material, a technique designed primarily for manufacturing cylindrical parts. Liquid metal is introduced into a rotating mold and allowed to solidify while the rotation persists. The primary advantage of centrifugal casting is its ability to enhance metal density, elevate actual mechanical properties of castings by 10–15%, and ensure a uniform metallurgical structure [2, 10].

Compositional gradients are mainly achieved through variations in centrifugal force arising from differences in density between molten metal and solid particles. Literature suggests there's a limit to the type of gradient that can be produced as it occurs naturally through processes like centrifugal force and density differences.

3.2.3 Powder metallurgy

The powder metallurgy (PM) technique is utilized for producing functionally graded material and involves four fundamental steps: powder preparation, weighing and mixing of powder according to pre-designed spatial distribution, stacking and ramming of premixed powders (forming operations), and finally compacting and forming a solid mass of material (sintering) [1, 11]. Multiple techniques, including chemical processes, electrolytic deposition, grinding, pulverization, atomization, centrifugal disintegration, and solid-state reduction, are used to prepare powders. Accurate measurement of quantities and the even distribution of combined powders are critical factors in powder processing that impact the characteristics of the structure. Forming procedures include compressing powders into certain shapes, usually done at normal ambient temperature. The strength of the compressed, unsintered portion depends on its capacity to be compacted. Sintered parts are produced as a solid mass without visible particle arrangement, although there may be some porosity, which can be reduced through additional processes.

Solid freeform fabrication (SFF) is a leading technology for directly producing physical objects from computer-generated part information. Among its notable applications, fabricating functionally graded materials stands out due to its ability to represent varying material distribution in geometry and control internal component composition. SFF is an additive manufacturing process offering numerous advantages, including higher production speed, lower energy consumption, maximal material utilization, and the ability to produce intricate designs and shapes. Laser-based processes are predominantly employed in SFF for fabricating functionally graded materials.

3.2 Applications of FGMs

In the recent years, the use of functionally gradient materials (FGMs) has grown due to their numerous advantages over composite materials. As functionally graded material to have graded variation in the composition, thus the properties of FGMs changes significantly and continuously from one location to another. This variation in chemical composition resulting in mechanical properties eliminates the interface problems like stress concentrations and poor interfacial bonding.

Nowadays, FGMs have attracted great interest in many fields, especially in some extreme environments, including aerospace engineering, nuclear power generation, marine, sensor, biomedical implants, optoelectronic devices, and absorption systems [1, 11]. **Figure 3** depicts the applications of FGMs in various sectors.

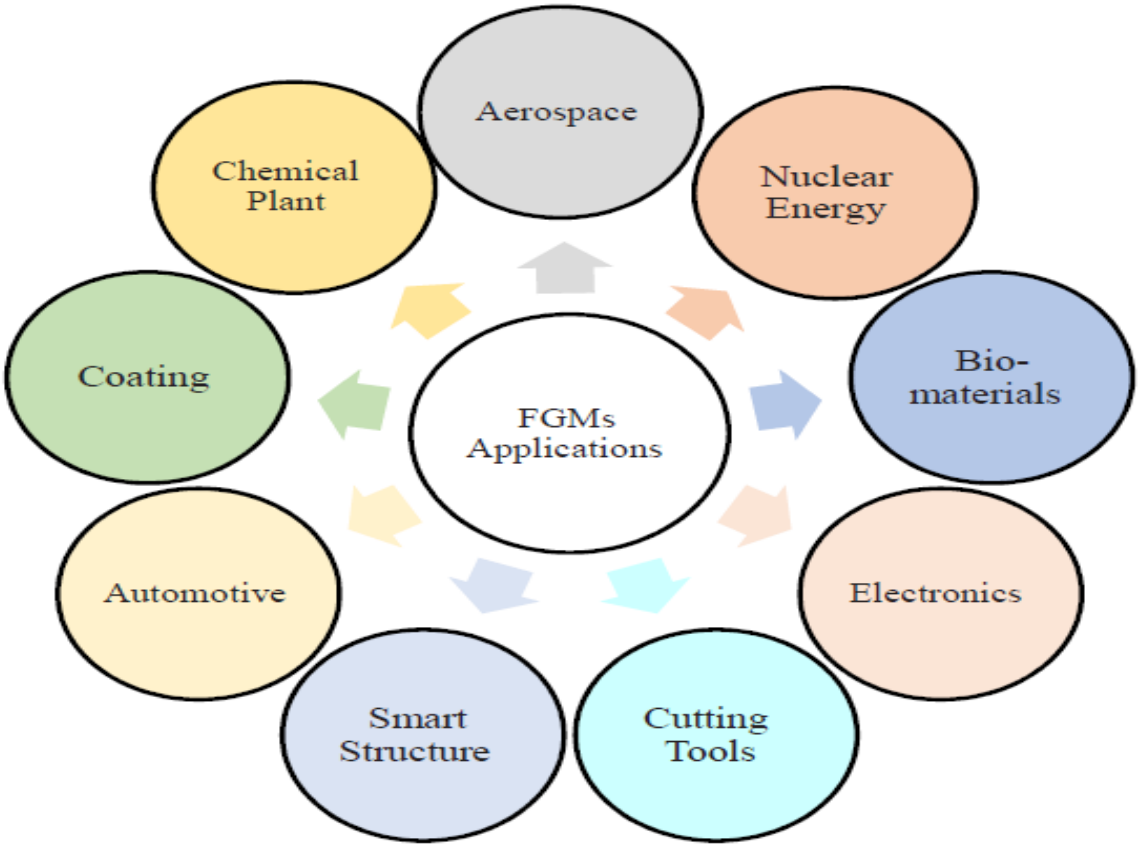


***Figure 3**. Applications of FGMs [11]*

3.2.1 Application in aerospace

Since FGMs reduce thermal tensions between exterior and interior surfaces, their true application began with spacecraft. Nevertheless, because of their many benefits, FGMs have found several additional uses. A rocket engine, the spacecraft gear structure, heat exchange plates, and some structures, like reflectors, solar panels, camera bunks, turbine wheels, turbine blade coating, nose covers, and the front edge of missiles and space shuttles, are among the most common parts of spacecraft and aircraft that currently use FGMs with gradient properties [12, 13].

The FGMs with graded characteristics provide this protection since these components are subjected to high temperatures or intense deterioration in particular places. Engine components and rocket nozzles are made of FG Al2O3/Al alloy, which has great resistance to heat and corrosion. This alloy was developed fairly recently. The manufactured FG polymer–ceramic composites in a different investigation to examine the impact of process variables on the mechanical and microstructure characteristics for use in aircraft [11].

3.2.2 Application in automobile engineering

Numerous vehicle parts require grading in heat, pressure, or stress, including flywheels, driveshafts, combustion chambers, race car brakes, and diesel engine pistons and cylinder liners. FGMs with gradient features are therefore ideal for these components. Furthermore, there is a chance to use FGMs with varying qualities in car body coatings. FGMs are quite expensive, therefore their usage in automotive applications is limited to crucial components of the vehicle [14].

To investigate the influence of graded structure on the microstructure behavior and physical as well as thermal parameters for automotive applications. According to the findings, the graded structure of FGMs improved the mechanical and thermal properties of vehicle brake rotors. A separate study investigated the influence of procedural parameters on the microstructure and mechanical properties of cylinder block liners in automobiles. The results indicated that the properties of FGM composites were superior to those of conventional composites in meeting the requirements of the intended application [15].

3.2.3 Application to fabricate machine parts

Functionally graded materials (FGMs) with graded structures are extensively utilized in manufacturing various engineering components, including cutting tools, forming molds, and machine engine blocks. These materials enhance the strength, thermal performance, and resistance to wear and corrosion of the components. Recently, FG WC/Co and TiCN–WC/Co have been employed in producing cutting tools such as turning tools, grinding and milling cutters, and twist drills. These materials improve surface wear resistance and hardness, thereby extending tool life, increasing feed rates, and reducing machining time [10, 11]. Additionally, FGMs are increasingly used in sports equipment like golf clubs, tennis rackets, and skis due to advancements in additive manufacturing processes. These developments enable the production of FGMs with graded properties, which reduce weight and friction, enhance durability, improve strength-to-weight ratios, and increase hardness and wear resistance.

3.2.4 Application in defense industries

In the defense sector, there is a growing trend to reduce vehicle weight and prevent crack propagation by utilizing functionally graded materials (FGMs). These graded structures are employed as penetration-resistant materials in the production of armor plates, bulletproof vests, and various defensive components such as guide rods, precision rollers, shafts, tubes, latches, axle housings, and firing pins. C. Huang and Y. Chen applied the FGM principle to create lightweight, high-impact-resistant, and cost-effective armor from Al2O3–ZrO2 composites using powder metallurgy. They conducted simulations to validate their experimental findings, demonstrating that FG Al2O3/ZrO2 composites with graded properties offer significant advantages over conventional composites, making them highly promising for defense applications [16].

4. Future scope and research gap

In previous sections, the methods of manufacturing and applications for FGMs were discussed, highlighting their advantages in creating smooth gradations and enhancing mechanical properties and wear resistance. This section addresses the research gaps and future directions for FGMs based on existing literature. FGMs offer significant advantages, making them crucial materials with a promising future. Researchers aim to refine production methods to expand their applications in line with the fourth industrial revolution. FGM processing methods can reduce thermal stresses, control deformation, corrosion, and wear, and enable the manufacture of complex structures with graded properties. However, challenges include mass production, quality control of the graded area, and cost, as well as enhancing production accuracy.

Future recommendations focus on five key areas: materials, processes, and mass production. These trends aim to optimize graded layers, incorporate new materials, refine manufacturing

methods, and expand applications. Particularly, mathematical modeling of FGMs is vital for understanding and predicting their properties. Future development will likely emphasize additive manufacturing to produce cost-effective, high-performance FGMs for automotive, aerospace, and biomedical applications. Despite current use, further research is needed to ensure reliability, especially in applications requiring continuous gradation. Dual gradient and hybrid FGMs via additive manufacturing are expected to be the next advancements.

5. Conclusions

FGMs exhibit graded properties from one point to another in a specific direction, making them an advanced form of conventional composites and it is concluded as:

- FGMs are evolving rapidly due to their superior mechanical and physical properties and that's why it has diverse applications across engineering, defence, petroleum, and transport sectors.
- This review paper explores the manufacturing methods, process parameters, and applications of FGMs. It identifies research gaps and future directions for FGM development.
- Despite recent advancements, certain aspects of FGMs still require further study. The paper outlines challenges and opportunities in the field, providing trends and suggestions for researchers.
- It serves as a comprehensive resource for researchers, academicians, and professionals to understand FGM concepts and complexities.
- The paper details key parameters and development sequences for achieving optimal graded properties. It offers study recommendations for new researchers and benefits all interested in FGM manufacturing, saving time and effort in the research process.

References

[1] Han, J., Lu, L., Xin, Y., Chen, X., Zhang, G., Cai, Y. and Tian, Y., 2022. Microstructure and mechanical properties of a novel functionally graded material from Ti6Al4V to Inconel 625 fabricated by dual wire+ arc additive manufacturing. Journal of Alloys and Compounds, 903, p.163981. https://doi.org/10.1016/j.jallcom.2022.163981

[2] Huang, J., Liu, G., Yu, X., Wu, H., Huang, Y., Yu, S. and Fan, D., 2022. Microstructure regulation of titanium alloy functionally gradient materials fabricated by alternating current assisted wire arc additive manufacturing. Materials & Design, 218, p.110731. https://doi.org/10.1016/j.matdes.2022.110731

[3] Rodrigues, T.A., Bairrão, N., Farias, F.W.C., Shamsolhodaei, A., Shen, J., Zhou, N., Maawad, E., Schell, N., Santos, T.G. and Oliveira, J.P., 2022. Steel-copper functionally graded material produced by twin-wire and arc additive manufacturing (T-WAAM). Materials & Design, 213, p.110270. https://doi.org/10.1016/j.matdes.2021.110270

[4] Mohan Kumar, S., Rajesh Kannan, A., Pravin Kumar, N., Pramod, R., Siva Shanmugam, N., Vishnu, A.S. and Channabasavanna, S.G., 2021. Microstructural features and mechanical integrity of wire arc additive manufactured SS321/Inconel 625 functionally gradient material. Journal of Materials Engineering and Performance, 30, pp.5692-5703. https://doi.org/10.1007/s11665-021-05617-3

[5] Wang, J., Pan, Z., Ma, Y., Lu, Y., Shen, C., Cuiuri, D. and Li, H., 2018. Characterization of wire arc additively manufactured titanium aluminide functionally graded material:

microstructure, mechanical properties and oxidation behaviour. Materials Science and Engineering: A, 734, pp.110-119. https://doi.org/10.1016/j.msea.2018.07.097

[6] Ahsan, M.R.U., Tanvir, A.N.M., Ross, T., Elsawy, A., Oh, M.S. and Kim, D.B., 2020. Fabrication of bimetallic additively manufactured structure (BAMS) of low carbon steel and 316L austenitic stainless steel with wire+ arc additive manufacturing. Rapid Prototyping Journal, 26(3), pp.519-530. https://doi.org/10.1108/RPJ-09-2018-0235

[7] Ahsan, M.R., Tanvir, A.N.M., Seo, G.J., Bates, B., Hawkins, W., Lee, C., Liaw, P.K., Noakes, M., Nycz, A. and Kim, D.B., 2020. Heat-treatment effects on a bimetallic additively-manufactured structure (BAMS) of the low-carbon steel and austenitic-stainless steel. Additive Manufacturing, 32, p.101036. https://doi.org/10.1016/j.addma.2020.101036

[8] Zhang, W., Lei, Y., Meng, W., Ma, Q., Yin, X. and Guo, L., 2021. Effect of deposition sequence on microstructure and properties of 316L and Inconel 625 bimetallic structure by wire arc additive manufacturing. Journal of Materials Engineering and Performance, 30(12), pp.8972-8983. https://doi.org/10.1007/s11665-021-06137-w

[9] Tian, Y., Shen, J., Hu, S., Chen, X., Cai, Y. and Han, J., 2022. Effect of deposition layer on microstructure of Ti-Al bimetallic structures fabricated by wire and arc additive manufacturing. Science and Technology of Welding and Joining, 27(1), pp.22-32. https://doi.org/10.1080/13621718.2021.1996850

[10] Gupta, A. and Talha, M., 2015. Recent development in modeling and analysis of functionally graded materials and structures. Progress in Aerospace Sciences, 79, pp.1-14. https://doi.org/10.1016/j.paerosci.2015.07.001

[11] Saleh, B., Jiang, J., Fathi, R., Al-hababi, T., Xu, Q., Wang, L., Song, D. and Ma, A., 2020. 30 Years of functionally graded materials: An overview of manufacturing methods, Applications and Future Challenges. Composites Part B: Engineering, 201, p.108376. https://doi.org/10.1016/j.compositesb.2020.108376

[12] Kumar, A., Singh, R.C. and Chaudhary, R., 2023. Investigation of microstructure and several quality characteristics of AA7075/Al2O3/coconut shell ash hybrid nano composite prepared through ultrasonic assisted stir-casting. Journal of Materials Engineering and Performance, 32(20), pp.9263-9278. https://doi.org/10.1007/s11665-022-07780-7

[13] Duraisamy, R., Kumar, S.M., Kannan, A.R., Shanmugam, N.S., Sankaranarayanasamy, K. and Ramesh, M.R., 2020. Tribological performance of wire arc additive manufactured 347 austenitic stainless steel under unlubricated conditions at elevated temperatures. Journal of Manufacturing Processes, 56, pp.306-321. https://doi.org/10.1016/j.jmapro.2020.04.073

[14] Kumar, A., Singh, V.P., Singh, R.C., Chaudhary, R., Kumar, D. and Mourad, A.H.I., 2024. A review of aluminum metal matrix composites: fabrication route, reinforcements, microstructural, mechanical, and corrosion properties. Journal of Materials Science, pp.1-68. https://doi.org/10.1007/s10853-024-09398-7

[15] Nemani, A.V., Ghaffari, M. and Nasiri, A., 2020. Comparison of microstructural characteristics and mechanical properties of shipbuilding steel plates fabricated by conventional rolling versus wire arc additive manufacturing. Additive Manufacturing, 32, p.101086. https://doi.org/10.1016/j.addma.2020.101086

[16] Huang, C.Y. and Chen, Y.L., 2016. Design and impact resistant analysis of functionally graded Al2O3-ZrO2 ceramic composite. Materials & Design, 91, pp.294-305. https://doi.org/10.1016/j.matdes.2015.11.091

RSC Advances

PAPER

View Article Online
View Journal | View Issue

Check for updates

Cite this: *RSC Adv.*, 2025, **15**, 9320

Exploring the dermal safety of green-synthesized Ag–TiO_2 nanocomposites for topical applications†

Archana Rana, *[ab] Shweta Parashar,[c] Diksha Singh,[c] Kavita Singh,[c] Debabrata Chanda,[bc] Anirban Pal,[bc] Ritu Srivastava [ab] and Shailesh Narain Sharma*[d]

We investigated Ag–TiO_2 nanocomposites (NCs) synthesized using leaf extracts of *Azadirachta indica* and *Mangifera indica* for topical applications. The Ag–TiO_2 NCs were first characterized by their spherical shapes, with sizes ranging from 20–26 nm to 5–6 nm, and a zeta potential value between −27 and −23 mV. DLS analysis revealed average particle sizes of 671 nm and 573 nm for Ag–TiO_2 NCs synthesized from *A. indica* and *M. indica*, respectively. The MICs of the nanocomposites were determined *via* dilution in both Gram-positive and Gram-negative bacteria to determine the optimal concentration for dermal applications. The cytotoxicity assay (MTT) of ATN and ATM compounds at MICs of 312.5 μg mL^{-1}, 625 μg mL^{-1}, and 1250 μg mL^{-1} showed that they were nontoxic to fibroblast cells. Further assessments of acute and subacute dermal safety were conducted on Charles Foster rats with NCs applied at 625 μg mL^{-1}, 3125 μg mL^{-1}, and 6250 μg mL^{-1} concentrations. Observations were made for any signs of dermal toxicity using behavioural and physical indices. In acute dermal toxicity, the NCs were applied once, and in subacute dermal toxicity, NCs were applied once daily for 28 days and observed for any sign of dermal toxicity using observation indices like behavioural changes, edema scores, and erythema scores. Post-experiment analyses of body weight, serum biochemistry, oxidative stress, and hematological profiles revealed that the nanocomposites exhibited significant antimicrobial activity. Notably, the safety evaluations indicated no adverse changes, suggesting these NCs are well-tolerated for dermal applications and show great promise for future topical applications.

Received 19th November 2024
Accepted 7th March 2025

DOI: 10.1039/d4ra08199d

rsc.li/rsc-advances

Open Access Article. Published on 27 March 2025. Downloaded on 4/8/2025 6:43:27 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

1. Introduction

Nanomaterials are an emerging technology with various applications in science, particularly in the field of topical formulations. Cosmetic and skin-related products have become an essential part of many people's daily routines owing to growing awareness of the harmful biological effects caused by ultraviolet (UV) radiation, skin allergies and infections. Exposure to the sun can lead to serious issues, such as skin cancer, photoaging, DNA damage, erythema, and increased reactive oxygen species in the skin.[1] In today's world, sunscreens are expertly crafted with UV filters—both organic and inorganic—to provide essential protection against harmful UV radiations. Organic UV filters, such as octocrylene, oxybenzone, and avobenzone, have been on the market for decades. However, they can trigger severe skin allergies due to the accumulation of unwanted substances in the adipose tissue. Conversely, inorganic filters stand out for their remarkable benefits, including broad spectrum protection, exceptional photostability, and a unique whitening effect. These qualities not only enhance their effectiveness but also affirm their safety for all skin types. Available in an array of different forms—oils, lotions, gels, and convenient roll-on applications—these products not only safeguard the skin but also ensure a luxurious, enjoyable experience in the sun.[2] TiO_2 NPs have previously been recognized as an inert material used in the cosmetics industry, and they are considered safe for use on the skin, as authorized by the EU's Scientific Committee on Consumer Safety (SCCS), with a maximum concentration of 25%. This safety is likely owing to the high surface area of TiO_2 NPs, which enhances transparency in visible light while reflecting and scattering UV rays.[3] Thus, while NPs have gained popularity in sunscreen formulations, mounting concerns about their potential toxicity have emerged, sparking ongoing discussions about their safety on human skin.[4] Recent studies have highlighted the remarkable potential of titanium dioxide (TiO_2) and AgNPs, which are traditionally synthesized through a variety of physical and chemical methods involving reducing agents.[5] However, there is a growing interest towards more sustainable and cost-effective approaches—

[a]*CSIR-National Physical Laboratory, Dr K. S. Krishnan Marg, New Delhi 110012, India. E-mail: microarchana91@gmail.com*
[b]*Academy of Scientific and Innovative Research (AcSIR), Ghaziabad 201002, India*
[c]*CSIR-Central Institute of Medicinal and Aromatic Plants (CSIR-CIMAP), Lucknow, India*
[d]*Delhi Technological University, Shahbad Daulatpur Village, Rohini, 110042, India. E-mail: shaileshnarainsharma@dtu.ac.in*

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d4ra08199d

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 27 March 2025. Downloaded on 4/8/2025 6:43:27 AM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

known as green methods—that not only reduce energy waste but also eliminate the risks associated with toxic substances.

This innovative shift toward green synthesis not only enhances the safety of nanoparticle production, but also unveils the potential for these nanoparticles in medical treatments, owing to the abundance of rich phytochemicals that they harness.[6] This evolution in synthesis techniques promises a more sustainable future for nanotechnology, merging environmental stewardship with cutting-edge medical applications. Green-synthesized NPs have shown remarkable effectiveness in topical applications.

Kwek *et al.* studied the effect of extracted TiO_2 NPs from sunscreens tested *in vitro* and *ex vivo* on human epidermal skin cells, showing lower cytotoxic activity than in the pristine form.[7]

Similar studies have been conducted by Nilavuk Karasi *et al.* to synthesize TiO_2 NPs using *Capparis zeylanica* leaf extract and assess their effect on wound healing in Swiss albino mice. The results demonstrated that a higher dose of TiO_2 NPs concentration exhibited 99% wound healing on the 13th day. The histopathological studies have not shown any side effects, but produced more cells on the epithelium layer.[8] Similar studies have been conducted for heparin–polyvinylalcohol@TiO_2 nanocomposite bandages to treat burn injuries. The incorporation of the TiO_2 NPs porosity, degradation rate, and antimicrobial activity increases the cell viability to enhance the wound healing properties.[9] AgNPs are well known for their strong antimicrobial properties due to their ability to release silver ions, which can bind to and disrupt the cell walls and membranes of bacteria and other microorganisms. AgNPs have been employed as bactericidal agents in hygiene and medical purposes, and increasingly as nanomedicine for treating and cleaning burns and diabetic wounds, as well as for purification of equipment. Studies show AgNPs have antibacterial properties and absorb in the UV range. The current studies revealed that AgNPs protect skin from UVB rays in both mouse models and cell cultures.[10] Green synthesized AgNPs were added to pluronic F-127 hydrogels at different concentrations to study their effect on the wounds of mice. The closure of the wound was confirmed on the 10th day with 95% healing, and the treatment did not observe any skin irritation and can be used as a topical ointment.[11]

Kim *et al.* (2016) studied the dermal and eye toxicity of AgNPs in rats and rabbits and did not observe any sign of erythema, oedema, or irritation effects on the eyes.[12] Recently, some studies have been published by Elatter *et al.* on Ag@TiO_2, and Ag@SeO_2 core/shell NCs synthesized by *Beta vulgaris* L. extract for *R. solani* and *S. sclerotiorum* fungi. The reported values found for Ag@TiO_2 NCs were 314, 624 μg mL^{-1} and 930, 1390 μg mL^{-1}.[13] The studies reveal that Ag@TiO_2 NCs were (one and a half times) more efficacious than Ag@SeO_2 NCs. Citrus lemon leaf extract-mediated AgNPs have been applied as an ointment to treat the wounds of Sprague Dawley rats, and it could be used for the treatment of human wounds.[14] Ag–TiO_2 NCs are composite materials that combine different phases of silver nanoparticles (AgNPs) and titanium dioxide nanoparticles (TiO_2 NPs). These nanocomposites have attracted significant attention due to their unique properties, which are different from those of their components. The AgNPs that occupy the surface of TiO_2 decrease the band gap and increase the photocatalytic properties, and are proposed to be effective nanocomposites in sunscreen.[15] The combination of AgNPs and TiO_2 NPs can lead to synergistic effects, such as enhanced photocatalytic activity, antibacterial properties, and improved electrical conductivity. Ag–TiO_2 NCs have potential applications in various fields, such as water purification, food packaging, and medical devices. They can be used to disinfect surfaces, prevent the growth of bacteria and other microorganisms, and reduce the risk of infections.[16] Ag–TiO_2 NCs@PF-127 HNRs were studied against 4T1 cancer cells, and showed 90% inhibition against the cells at a concentration of 120 mg mL^{-1} under UV radiation.[17]

Despite numerous studies, TiO_2 NPs face several challenges, including high recombination rates of electron–hole pairs, a small band gap energy, and low absorption capacity. To address these issues, researchers have doped TiO_2 NPs with various metal nanoparticles, such as silver (Ag), gold (Au), and iron (Fe). This doping helps to reduce the band gap without compromising the crystallinity of the TiO_2 NPs, and facilitates the diffusion of charge carriers on the surface of the nanoparticles.[18]

The photocatalytic activity of Ag–TiO_2 NPs is enhanced at the interface between Ag and TiO_2 NPs when exposed to UV light. When UV light is present, photoinduced electrons are transferred from the silver nanoparticles (Ag NPs) to the titanium dioxide nanoparticles TiO_2 NPs, increasing their photocatalytic activity. The electrons that accumulate on the surface of the TiO_2 NPs react with oxygen or other molecules to produce superoxide radicals. Additionally, the holes generated during this process can oxidize hydroxyl groups (–OH) or water (H_2O) to create powerful oxidizing radicals, such as hydroxyl radicals ($^{\bullet}OH$). Moreover, TiO_2 NPs can be oxidized by these strong oxidizing radicals ($^{\bullet}OH$), resulting in the production of carbon dioxide CO_2 and several other oxidized intermediates.[19] The integration of AgNPs into TiO_2 NPs significantly diminishes the recombination of e^-/h^+ pairs, ultimately extending the lifespan of Ag–TiO_2 NCs. This remarkable enhancement owes its effectiveness to the captivating surface plasmon effect of AgNPs, which facilitates the transport of electrons from TiO_2 NPs to AgNPs.[20] Inorganic NPs (sunblock) were measured with SPF to reduce the erythema after evaluating the rate of absorption, toxicology, and skin reactions on human skin.[21] Recent studies on Ag/TiO_2 NPs (400 μg mL^{-1}) reveal that, when combined with blue light, there is a 42.79% reduction in cell viability in the A431 skin cancer cell line after 45 minutes, compared to using nanoparticles and light separately.[22] Ag-doped TiO_2 NPs studied on human cancer (liver) cell lines showed better results with the addition of AgNPs, as it only focused on destroying cancer cells.[23] Several studies show that TiO_2 NPs reduce the release of AgNPs, enhancing antimicrobial activity and decreasing health problems and pollution. It is highly used as a hydrogel for wound dressing, cosmetics, antimicrobial agents and semiconductors. These well-known studies promote for skin repairment and wound healing which indicate harmless process to human usage proved *via in vivo* and *in vitro*.[24]

In this study, we synthesized green Ag–TiO_2 NCs using extracts from medicinal plants, specifically *Azadirachta indica* (*A. indica*, also known as neem) and *Mangifera indica* (*M. indica*,

Open Access Article. Published on 27 March 2025. Downloaded on 4/8/2025 6:43:27 AM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

as mango in India), which are commonly found in Asian countries. The synthesized Ag–TiO_2 NCs were assessed for their effects on the cytotoxicity assay (MTT) fibroblast cell line. They were determined to be nontoxic at concentrations up to twice the minimum inhibitory concentration (MIC). Additionally, the sub-acute toxicity of Ag–TiO_2 NCs—derived from *A. indica* (ATN = Ag–TiO_2 neem) and *M. indica* (ATM = Ag–TiO_2 mango) extracts—was safe at levels up to five times the MIC. This research paper presents novel findings on the dermal toxicity of Ag–TiO_2 NCs in CF rats, which have not been previously published.

2. Experimental methods and materials

2.1. Materials

A. indica and *M. indica* leaves were collected from the campus of NPL (National Physical Laboratory), New Delhi, India. The precursors used for synthesis include titanium isopropoxide [$Ti(OC_3H_7)_4$], silver nitrate ($AgNO_3$), iodo-nitrotetrazolium bromide and nutrient broth, which were purchased from Sigma-Aldrich chemicals. The microbial strains (*Staphylococcus aureus* MTCC 96, *Staphylococcus hominis* MTCC 657, *Escherichia coli* MTCC 8739, *Salmonella typhimurium* (MTCC 3232) and *Candida albicans* MTCC 183) were procured from Microbial Type Culture Collection (MTCC), Chandigarh, India.

2.2. Preparation of plant extract

The collected leaves of *A. indica* and *M. indica* were appropriately cleaned under tap water to eliminate dust residue. 30 g of each plant (*A. indica* and *M. indica*) leaves were used to prepare the leaf extract using 100 mL of deionized water, and the mixture was heated at 80 °C for 40 min. The reaction mixture was then filtered using Whatman no. 1 filter paper after cooling to room temperature, and stored at 4 °C.

2.3. Synthesis of Ag–TiO_2 NCs using *A. indica* and *M. indica* leaves extract

The green synthesis of Ag–TiO_2 NCs, followed by precursors (0.2 M) titanium(IV) isopropoxide (TTIP–Ti[$OCH(CH_3)_2]_4$) in 30 mL deionized water, was placed in a round-bottom flask stirred at 70 °C for 20 min. Thereafter, 3 mM silver nitrate ($AgNO_3$) solution with 30 mL of leaves extract of *A. indica* and *M. indica* was added slowly to the round-bottom flask with continuous stirring at 70 °C, following the protocol reported by Bhardwaj *et al.*[25] Slowly, the colour of the reaction mixture changed into a light grey color (could be due to the surface plasmon resonance), which indicates the formation of Ag–TiO_2 NCs. The reaction mixture of the NCs solution was centrifuged at 4000 rpm for 30 min to separate the suspended particles. Furthermore, the NCs particles were washed twice with deionized water to remove the extra product. The collected NCs were then dried overnight at 70 °C in the oven. The resulting Ag–TiO_2 NCs were annealed at 400 °C for 2 h to enhance their crystallinity. A schematic of the synthesis of Ag–TiO_2 NCs is shown in Fig. 1(a). Furthermore, the mechanism of Ag–TiO_2 NCs using plant extract is shown in Fig. 1(b). The TTIP reaction with D.I. water produces unstable $Ti(OH)_4$, which is further converted into TiO_2 through the condensation process. The hydroxyl group of phytochemicals (such as ascorbic acid, quercetin, nimbin, benzoic acid propyl ester) is attached to the surface of TiO_2 after the addition of the plant leaf extract. The addition of AgNPs on the surface of TiO_2 NCs is achieved through hydrogen bonding of the OH group of phytochemicals.

2.4. Structural and morphological characterization

Ag–TiO_2 NCs synthesized using plant leaf extract of *A. indica* and *M. indica* were characterized using Raman, SEM, XRD, FTIR, TEM, and UV-vis analysis. XRD (X-ray diffraction) patterns were recorded using a Bruker D-8 Advance X-ray Diffractometer instrument with Cu Kα with $\lambda = 1.54059$ Å, employing a scanning rate of 0.05° s^{-1} in the 2θ range from 10° to 80°. The

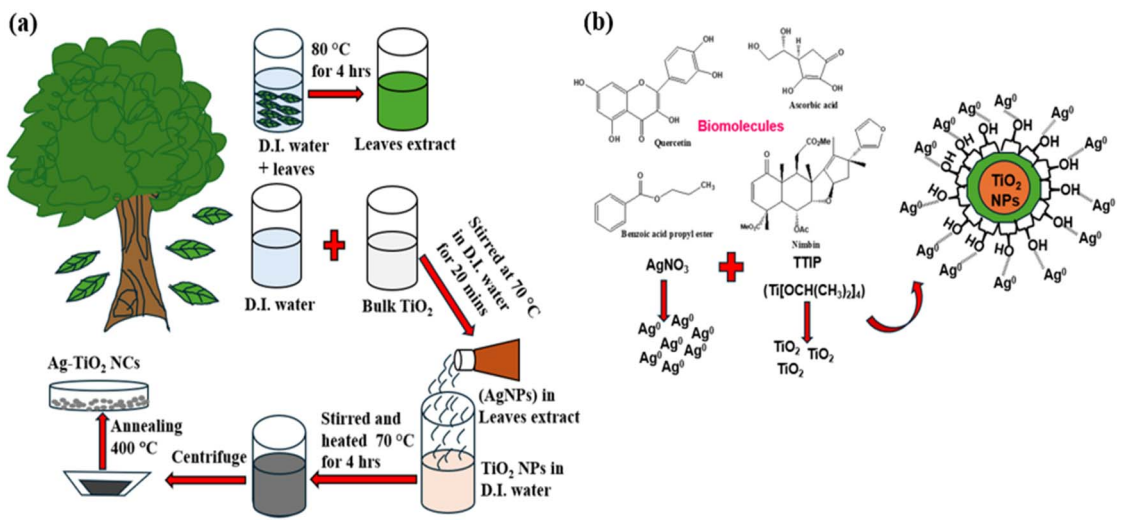


Fig. 1 Schematic of (a) the proposed green synthesis route of Ag–TiO_2 NCs using *A. indica* and *M. indica* leaf extracts and (b) mechanism of Ag–TiO_2 NCs.

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 27 March 2025. Downloaded on 4/8/2025 6:43:27 AM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

absorbance of NCs between 200–800 nm was measured using a UV-vis spectrophotometer (Shimadzu UV-3101). The bandgap energy (E_g) of the NCs obtained may be determined using the following equation:

$$E_g = 1240/\lambda \text{ eV} \quad (1)$$

λ = wavelength (nm) and E_g = band gap (eV).

The powder sample was characterized using Fourier-transform infrared (FTIR) spectra (PerkinElmer FTIR spectrum BX-II analyzer) to investigate the functional groups, which were recorded in the 4000–400 cm^{-1} region.

Raman spectroscopy of NCs was performed using laser excitation at 514.5 nm with an argon ion laser power of approximately 80–100 mW.

The Malvern Panalytical equipment was used to test the stability and hydrodynamic diameter of spherical NCs using dynamic light scattering (DLS). The particle size was estimated using a Malvern (1062761) Zetasizer at 25 °C with a cuvette of 2 mm in an aqueous solution. The biochemical and oxidative stress values were calculated through SpectraMax Plus (384) instrument.

The size and shape of the NPs were assessed using a TEM (Transmission Electron Microscope) instrument from Jeol Ltd [Tokyo, Japan] with an acceleration voltage of 100 kV. The NCs for the TEM investigation were ultrasonicated for 15 minutes before mounting onto a copper grid with carbon coating to dry in the air. The shape of the NCs were studied using a SEM (scanning electron transmission) and EDAX instrument (ZEISS instrument).

2.5. Antimicrobial activities

The Ag and TiO_2 NPs have been studied separately using similar plants and methods for antimicrobial activities on plant pathogens using the disc diffusion method. AgNPs were synthesized using *A. indica* and *M. indica* leaf extract with plant pathogens showing a zone of inhibition of 21 mm (*A. rolfsii* fungi) and 17 mm (*X. oryzae* bacteria) *via* the zone of inhibition disc diffusion method.[26] Similarly, TiO_2 NPs were also synthesized using the same plants and studied against plant pathogens, such as *A. alternata*, *C. gloeosporioides*, *T. harzianum* and *A. rolfsii* fungi and *X. oryzae* as bacteria using the disc diffusion method. Studies on the non-linear optical properties (absorption and refractive index) were also calculated through the Z-scan method.[27] In the current study of Ag–TiO_2 (*A. indica* and *M. indica*), the broth dilution method MIC was performed. Quantitative estimation of microbial growth inhibition by samples was observed by the micro broth dilution assay using 96 well 'U' bottom microtitre plates, as mentioned by Srivastava *et al.*[28] The method employs a two-fold serial dilution technique to assess any given antimicrobial's MIC (Minimum Inhibitory Concentration) under testing. The MIC of the test sample is considered as the concentration at which no colour change is observed, starting from the highest to lowest dilution after incubation of the microtiter plates.

Nutrient broth (100 μL) medium was added to all wells. 100 μL of the sample was added to the first well at a concentration of 5000 mg mL^{-1} diluted with broth medium. The solution was mixed by continuous pipetting (up to 5–6 times) and then transferred 100 μL from the 1st well into the 2nd well. It was serially diluted (2-fold dilution) up to the 10th well for testing as an antimicrobial agent. The schematic of the antimicrobial activity is shown in Fig. 2.

The working inoculum was prepared by diluting 1 mL of the overnight grown culture with 19 mL of sterile broth to achieve a (CFU) of 1×10^6 cfu mL^{-1}. To each well was added 100 μL of this working inoculum, except for the negative control. Column 11 was used for the negative (sterility) control, and column 12 was used for plate sterility with only nutrient broth. After 18 h incubation at 37 °C, plates were evaluated by adding 50 μL of iodo-nitrotetrazolium bromide (stock-1 mg mL^{-1}) dye to achieve a final concentration of 0.02 mg mL^{-1}. The well exhibiting no growth in each row corresponded to that antimicrobial agent's MIC (μg mL^{-1}). The MIC values (μg mL^{-1}) recorded in this study are the mean of three experiments performed in replicates.

3. Cytotoxicity assay (MTT) of ATN and ATM in the fibroblast cell line

3.1. Principle

In 1983, Mosmann introduced the MTT assay, a standard method for assessing cell viability and cytotoxicity. This assay is based on the ability of living cells to reduce yellow tetrazolium salt (MTT) into purple formazan crystals, primarily through the activity of mitochondrial dehydrogenases. Enzymes located in other organelles, such as the endoplasmic reticulum, also play a role in this reduction.[29] Once MTT is taken up by live cells *via* endocytosis, it is metabolized and converted into insoluble formazan crystals, which are transported back to the cell surface, forming needle-like structures. These insoluble crystals are then dissolved using a solubilizing agent, such as DMSO, acidified ethanol, or SDS in diluted hydrochloric acid. The resulting solution is measured using a spectrophotometer at 570 nm to determine the level of cell viability.

3.2. Method

Fibroblast cell lines were cultured in a growth medium (DMEM with 10% heat-inactivated fetal bovine serum), and seeded at a density of 0.5×10^4 cells per well in a 96-well plate. These cells were allowed to adhere overnight at 37 °C in a 5% CO_2 incubator. Following attachment, the cells were treated with ATN and ATM at concentrations of (312.5 μg mL^{-1} (1/2× of MIC), 625 μg mL^{-1} (1× of MIC), and 1250 μg mL^{-1} (2× of MIC)), along with the vehicle control (DMSO), for 12 hours in a culture medium supplemented with 1% serum. After the treatment, 20 μL of MTT solution (prepared from a 5 mg mL^{-1} stock in PBS, 12 mM MTT) was added to each well and incubated for 4 hours. The MTT-containing medium was then carefully removed, and the formed formazan crystals were solubilized in 100 μL of DMSO for 10 minutes, resulting in a blue-coloured solution. Cell viability was assessed by measuring the absorbance of the formazan solution at 570 nm using a spectrophotometer. The

Open Access Article. Published on 27 March 2025. Downloaded on 4/8/2025 6:43:27 AM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

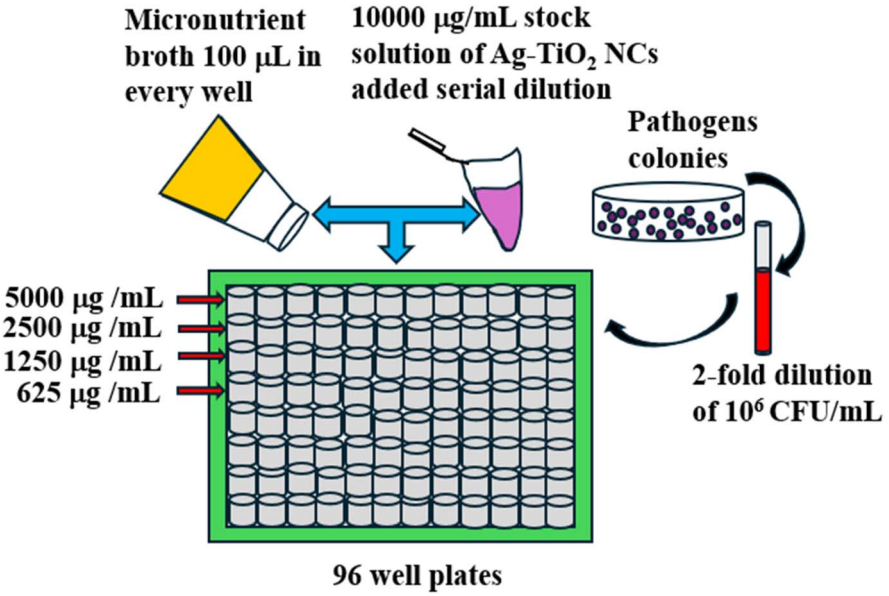


Fig. 2 Antimicrobial activity of Ag–TiO_2 NCs synthesized using *A. indica* and *M. indica* leaf extracts.

intensity of the color was directly proportional to the cell viability, and the viability was expressed as the percentage of survival. Each test was performed with ($n = 4$), and the data were presented as the mean ± SEM ($n = 4$).

4. Mixing of Ag–TiO_2 NCs (*A. indica* and *M. indica*) in sesame oil

Ag–TiO_2 NCs (*A. indica* and *M. indica*) having proven MIC were considered for evaluation of dermal toxicity and were accurately weighed. The MIC (μg mL^{-1}) was considered as × and in dermal toxicity, we have chosen 1×, 5×, and 10× concentrations for our study after suspending the NCs in sesame oil.

4.1. *In vivo* dermal toxicity

Animal studies were carried out according to protocols approved by the Institutional Animal Ethics Committee (IAEC) for Animal Facility, CSIR-CIMAP, Lucknow, UP, India (CIMAP/IAEC/2020-23/22), following the recommendations of the OECD test guidelines for dermal toxicity and FHSA regulations. Before commencing the study, the animals were familiarized with the experimental condition for one complete week (22 °C and a 12 hour dark/light cycle). The animals were given adequate rat feed and drinking water *ad libitum*. In acute dermal toxicity, 8 male and 8 female Charles Foster rats were taken. Following a week of acclimatisation, the animals were divided into 4 groups each with 2 male and 2 female rats (to obtain 8 spots of 1 square inch for the dermal application in each group). The lateral side of the abdominal region was clipped of hair and cleanly shaven with a marked area of a square inch on each side of the rats for applications of the NCs. Group 1 was considered as control, whereas groups 2 through 4 were considered the experimental groups, and were treated with the NCs at 1×, 5×, and 10× of MIC of the NCs. A similar grouping of animals and application procedure of NCs was followed for assessing the sub-acute dermal toxicity, wherein the test substance was applied once daily for 28 days.

4.2. Acute and subacute dermal toxicity

Ag–TiO_2 NCs were used at 1×, 5×, and 10× of MIC to evaluate the toxicity of the NPs. Ag–TiO_2 NCs were solubilized in sesame oil and applied on the marked area of the skin of the CF rats. We observed the animals for 24 hours, while archiving the images of the treated. We determined edema and erythema scores, and calculated the primary irritation index as reported previously.[30]

We performed the sub-acute dermal toxicity tests to determine the adverse effects of Ag–TiO_2 NCs upon repeated dermal exposure to rats. We gave 1×, 5×, and 10× of MIC to the animals for 28 days repeatedly in the marked area of the skin. We took a picture of the marked area before the application of the nanoparticle on day 0. After applying the NCs, the images of the marked area were collected once every 3 days. Animals were regularly monitored for any development of edema and erythema. The irritation index and primary irritation index were calculated for the experimental period of 28 days at 3 day intervals. We shaved the animal twice a week to remove the hairs for the application of the compound until the completion of the experiment. On the 28th day of the experiment, we recorded the body weight of the animals and collected blood from the retro-orbital plexus for biochemistry and hematology. The animals were then sacrificed under high-dose anesthesia (a combination of ketamine at 80 mg kg^{-1} and xylazine at 20 mg kg^{-1} given through the i.p. route), and vital organs like the lung, liver, heart, kidney, spleen, and brain were collected, inspected, and weights were recorded and expressed as absolute and relative organ weight.[31]

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 27 March 2025. Downloaded on 4/8/2025 6:43:27 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

4.3. Statistical analysis

Hematological, biochemical, and oxidative stress data were analyzed using statistical calculation in Excel. For these parameters, the data are presented as mean ± standard deviation. The relevant significance level was determined at ($P \leq 0.05$) as applicable. Two-way ANOVA tests were used to examine the differences in the data collected from the groups and evaluate the measurements. We performed all of the experiments in triplicates, and expressed the data as mean ± SEM. Two-way ANOVA was used to calculate the *p*-value.

5. Results and discussion

5.1. Crystallographic structure

XRD was used to determine the crystal structures and phases present in the NCs. In the XRD pattern of Ag–TiO_2 NCs, the peaks observed in the pattern correspond to the diffraction of X-rays by the atoms in the crystal lattice of the NPs. The XRD pattern of silver–TiO_2 NCs shows peaks that correspond to both the silver and TiO_2 phases. In the XRD pattern of the silver–TiO_2 nanocomposites, the peak positions and intensities can provide information about the crystallite size, crystal structure, and composition of the NPs. The XRD pattern of Ag–TiO_2 NC (*A. indica* and *M. indica*) samples is shown in Fig. 3(a). The diffraction peaks at 2θ angles (with corresponding planes) at 25.3° (101), 48.05° (200), 53.9° (105), 55.06° (211), 62.4° (204), 68.76° (220), 70.3° (116), and 75.06° (215), respectively, are attributed to the anatase phase of TiO_2 NPs (JCPDS card no. 21-1272). The XRD pattern of Ag–TiO_2 NCs shows the characteristic peak of Ag at 2θ of 32.3° (111), and 37.9° (004) (JCPDS no. 04-0784).[32] In the sample containing both Ag and TiO_2, the diffraction band of the Ti (004) lattice plane appears at an angle of 37.9°, which is high enough to overlap with the (111) characteristics peak of Ag.[33,34] The XRD pattern in the present study indicates that the dual phases of Ag–TiO_2 NCs have been confirmed by the Ag NPs that were deposited on the surface of the TiO_2 particles. The crystallite size of Ag–TiO_2 NCs (*A. indica* and *M. indica*) was calculated using the Scherrer equation, where β = full width at half maximum, θ = diffraction angle at peak position, and λ = wavelength of X-ray radiation.

$$D = \frac{0.94\lambda}{\beta \cos \theta} \quad (2)$$

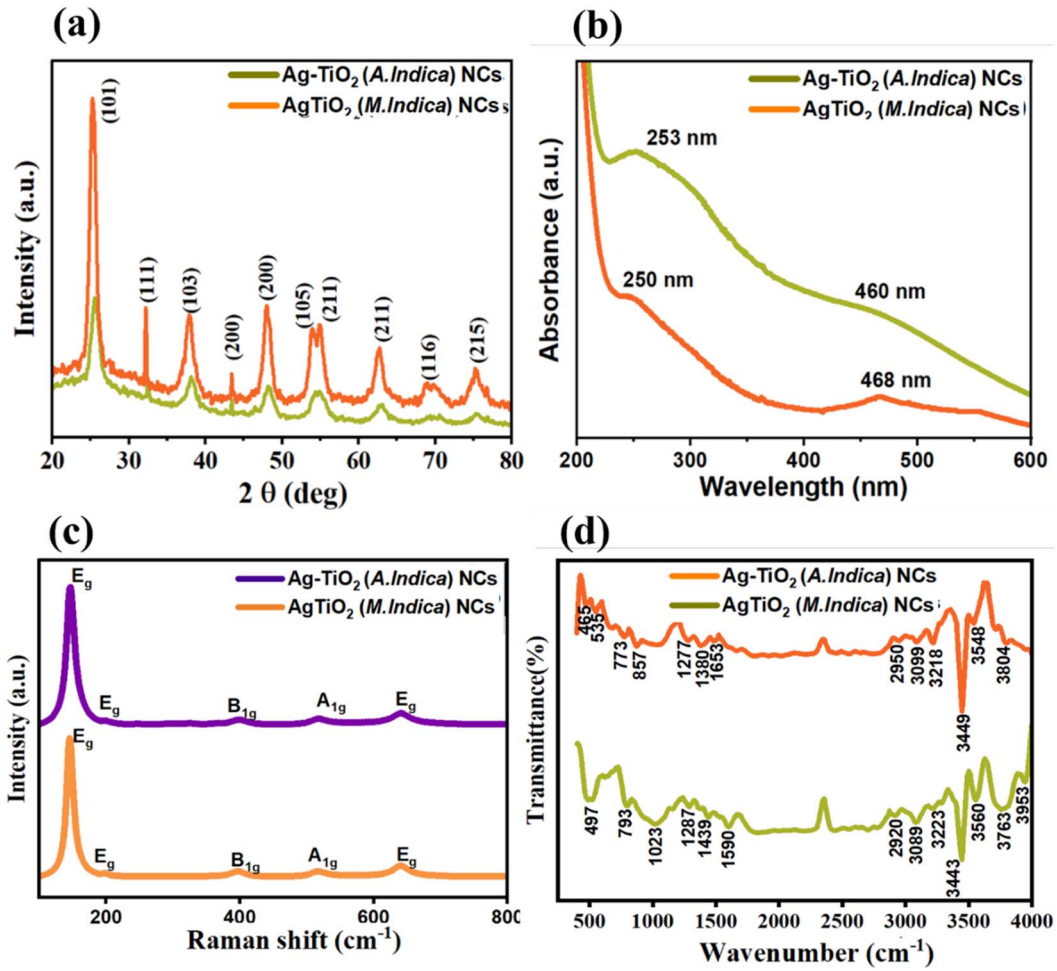


Fig. 3 (a) X-ray diffraction (XRD) patterns of Ag–TiO_2 NCs and (b) UV-vis spectra of Ag–TiO_2 NCs synthesized using *A. indica* and *M. indica* leaf extracts, (c) Raman spectra of Ag–TiO_2 NCs and (d) FTIR spectra Ag–TiO_2 NCs synthesized using leaf extracts of *A. indica* and *M. indica*.

Open Access Article. Published on 27 March 2025. Downloaded on 4/8/2025 6:43:27 AM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

The crystallite size for Ag–TiO_2 NCs was calculated as 7.23 nm and 7.9 nm at 400 °C annealing. Additionally, the XRD pattern exhibits the presence of AgNPs in NCs without coinciding with other peaks of TiO_2 NPs. The high crystalline nature of Ag–TiO_2 NCs is demonstrated by the presence of strong peaks in the data. The XRD pattern shows distinct diffraction peaks of Ag and TiO_2 phases for both samples, with the absence of any other peaks. The characterizations mentioned in the current paper have been studied only for Ag–TiO_2 NCs, whereas the characterization and results for Ag and TiO_2 NPs synthesized using the same plant extracts have been reported earlier for agricultural applications.[26,27]

5.2. Optical properties

The two main UV-vis absorption spectra of Ag–TiO_2 NCs (*A. indica* and *M. indica*) are shown in Fig. 3(b). The absorption bands are related to the electronic transitions in the NPs, which are affected by the size, morphology, and composition of the NPs. The first absorption band in the UV region is associated with the TiO_2 NPs' electronic transitions with a blue shift of 253 and 250 nm (*A. indica* and *M. indica*), which could be due to the AgNPs deposited on the surface of TiO_2 NPs.[35] The second absorption band in the visible region (redshift) is due to the SPR (Surface Plasmon Resonance) of the AgNPs observed at 460 and 468 nm (*A. indica* and *M. indica*).[36] The SPR band of AgNPs observed on the surface of TiO_2 significantly affected the absorption in the visible area of 450–550 nm, and the absorption edge also changes to the visible region for Ag–TiO_2 NCs.[37] The SPR band depends on the size, shape, and concentration of the AgNPs. Ag–TiO_2 NCs (*A. indica* and *M. indica*) was studied in the wavelength range of 200–600 nm. Furthermore, the significant shift in the dielectric property of the surrounding medium marked by the presence of AgNPs was also assigned to these improvements in the optical absorption of the Ag/TiO_2 NCs.[38]

5.3. Raman spectra

Raman spectroscopy is known as an incredibly effectual technique for detecting the vibrational modes of chemical bonds, and can precisely pinpoint the specific crystalline state of materials. In this study, we used this technique to study the Raman spectra of Ag–TiO_2 NCs in the frequency range of 150–800 cm^{-1}, shown in Fig. 3(c). The robust and intense peaks observed at 147 cm^{-1} (E_g (1)) and 199 cm^{-1} (E_g (2)) are optical modes of the anatase phase of TiO_2. The band located at 398 cm^{-1} (B_{1g}), 515 cm^{-1} (A_{1g}), and 641 cm^{-1} (A_{1g}/B_{1g}) assumed a low-frequency shift, corresponding to the LO-phonon in Ag–TiO_2 NCs. For Ag–TiO_2 NCs, all the Raman bands observed were for TiO_2, and the absence of an active band for Ag was due to the crystal symmetry of Ag.[39] However, the oscillations in the Raman spectra could be due to the oxygen atoms further changing the coordination environment of oxygen with the Ti^{4+} atom. Adding Ag atoms to the lattice of TiO_2 to maintain neutrality creates oxygen vacancies and forms Ag–O–Ti NCs.[40] The broadening of the peaks could be possibly attributed to the vibrational amplitudes and particle size.[41]

5.4. Evaluation of functional groups

The FTIR spectrum of Ag–TiO_2 NCs (*A. indica* and *M. indica*) shows the characteristic peaks of fundamental vibrations caused by N–H, O–H, and C–H stretching, measured in the 4000–2500 cm^{-1} range. The stretching vibrations for the absorbed –OH group of water molecules were observed at 3400 cm^{-1}. The C=C and C=O double bonds stretching vibrations are ascribed to the regions 2500–1500 cm^{-1}. The Ti–O–Ti bond stretching vibrations have been identified in the 700–500 cm^{-1} range. The peaks identified at 1287 and 1277 cm^{-1} represent the presence of Ag–O–Ti, which depicts the formation of NCs in Fig. 3(d).[42] The weak peaks observed for TiO_2 NPs were assigned to the lattice vibration range of 1380–1277 cm^{-1}. The Ti stretching vibrations were observed with a prominent peak at 1023 cm^{-1}.

5.5. DLS and zeta potential analysis

Dynamic light scattering is known for detecting the particle size distribution and zeta potential of aqueous nanocomposites of Ag–TiO_2 NCs (*A. indica* and *M. indica*) were determined using dynamic light scattering and a zeta sizer. Typically, the sample is dispersed in a suitable solvent and passed through a laser beam. The scattered light is then detected and analyzed to assess the size of the NPs in the sample. The size of Ag–TiO_2 NCs can vary depending on the synthesis method and the ratio of silver to titanium dioxide. It is important to note that DLS analysis can provide information about the hydrodynamic diameter of the NPs, which can be different from the actual size of the particles measured by TEM. It may be because the NPs are surrounded by a layer of solvent molecules, affecting their size and mobility. The average particle sizes of Ag–TiO_2 NCs (*A. indica* and *M. indica*) were 671 and 573 nm, respectively, as shown in Fig. 4(a) and (b). According to these findings, the presence of AgNPs on the surface of TiO_2 particles may prevent agglomeration.[43]

The zeta potential of these NCs refers to the electric charge that might be present on their surface, which affects their stability and dispersion in a liquid medium. The zeta potential of silver–TiO_2 NCs depends on various factors, such as the size and concentration of the NPs, the pH of the medium, and the presence of other ions. Typically, the zeta potential of these NCs is negative, indicating a repulsive force between them, preventing agglomeration or aggregation of the NPs. A high negative zeta potential is desirable for the stability of the nanocomposites in a liquid medium. The zeta potentials of Ag–TiO_2 NCs (*A. indica* and *M. indica*) were calculated as −27 mV and −23 mV, as shown in Fig. 4(c) and (d).

5.6. Morphological studies (TEM analysis)

The morphology and structure of Ag–TiO_2 NCs were analyzed using TEM, HRTEM and SAED, as shown in Fig. 5(a) and (b), to obtain more details of Ag–TiO_2 NCs (*A. indica* and *M. indica*). The TEM images demonstrate the mesoporous structure and explain the agglomeration of NCs. The TEM analysis of the TiO_2 NPs exhibited averaged particle sizes of 20–26 nm, which were made

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 27 March 2025. Downloaded on 4/8/2025 6:43:27 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

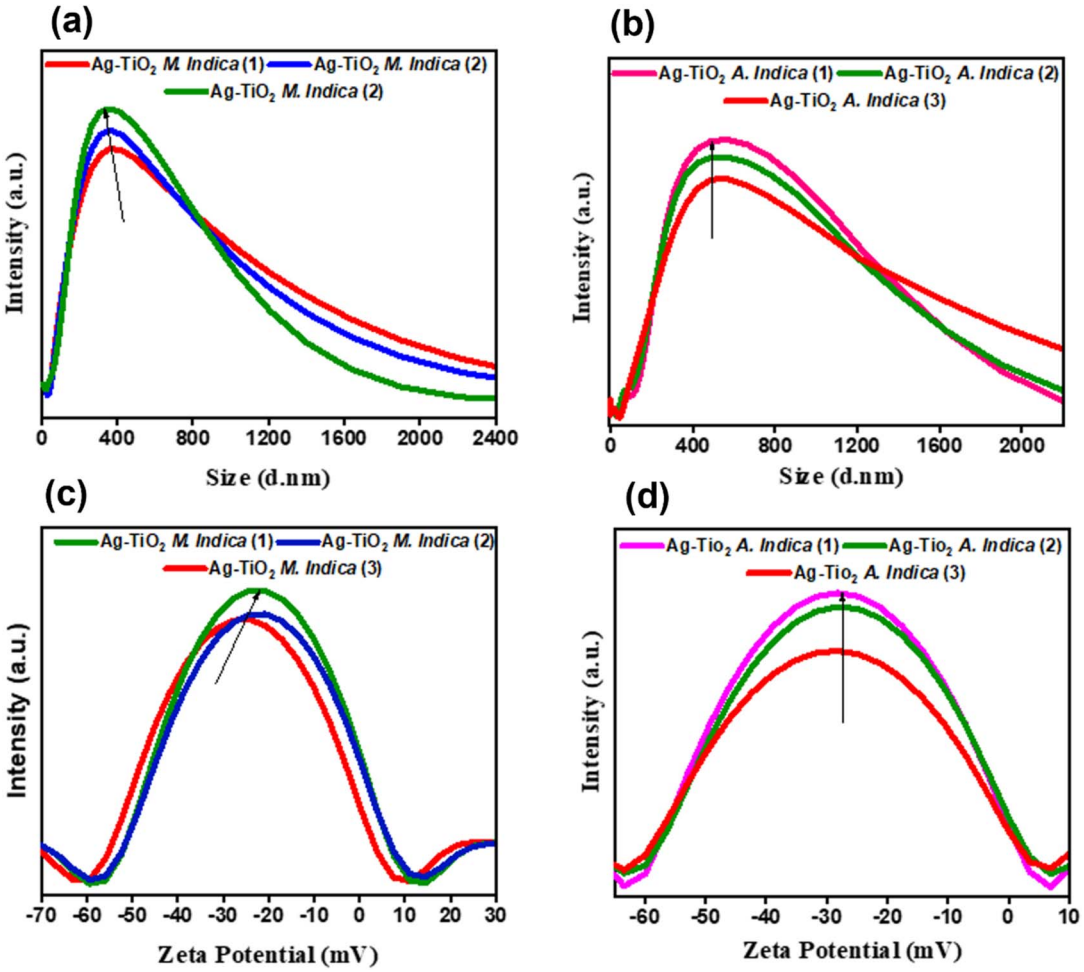


Fig. 4 (a and b) Dynamic light scattering and *Z*-average hydrodynamic diameter of Ag–TiO_2 NCs (*A. indica* and *M. indica*), (c and d) zeta potential measurements for Ag–TiO_2 NCs synthesized using *A. indica* and *M. indica* leaf extracts.

up of small-sized anatase crystallite-size NPs. Similar TiO_2 NPs have already been reported in our previous research paper.[28]

The small particles ranging from 5–8 nm for Ag NPs agree well with the XRD and UV results. The AgNPs were eventually distributed and arranged on the surface of the TiO_2 crystalline particles. The high-resolution TEM (HRTEM) images reveal that the lattice fringes of Ag have a plane spacing of 0.214 nm and 0.209 nm, which matches well with the (200) lattice distance, and the plane spacing of TiO_2 (anatase phase) is about 0.357 and 0.351 nm, which corresponds to the (101) plane, as shown in Fig. 5(c) and (d). These HRTEM parameters agreed with the X-ray diffraction (XRD) spectra. A SAED pattern of Ag–TiO_2 NCs (*A. indica* and *M. indica*) shows the polycrystalline nature of NCs. The observed rings in Fig. 5(e) and (f) are assigned to the (111), (004), (101), (200), and (211) planes of Ag and TiO_2 NPs.

5.7. SEM and EDAX analysis

The particle shape of green synthesized Ag–TiO_2 NCs were studied by SEM analysis which observed the spherical shape of both Ag and TiO_2 NCs. These images show agglomeration (minimum surface free energy) of the particles that exhibit a non-uniform particle distribution Fig. 6(a) and (b). These nanocomposites reported large surface areas, which makes them applicable for nano-medicine applications and also have more catalytic active sites. The elemental composition of Ag–TiO_2 NCs was determined by EDAX analysis and confirmed the presence of Ag (silver), Ti (titanium), and O (oxygen), as shown in Fig. 6(a) and (b).

5.8. Antimicrobial activity and determination of MIC of Ag–TiO_2 NCs

The minimum inhibitory concentration of Ag–TiO_2 NCs (*A. indica* and *M. indica*) against *Escherichia coli* and *Staphylococcus aureus* is tabulated below. Both Ag–TiO_2 NCs (*A. Indica* and *M. Indica*) with a MIC <2500 μg mL^{-1} were considered for testing against two Gram-positive bacteria (*S. aureus* & *S. hominins*), two Gram-negative bacteria (*E. coli* & *S. typhimurium*) and yeast (*C. albicans*). These NCs were active with the concentration (MIC <2500 μg mL^{-1}) against the mentioned microbes shown in Table 1. Based on the spectrum of the pathogens and the MIC, the activity is best in Ag–TiO_2 NCs (*A. indica*) compared to Ag–TiO_2 NCs (*A. indica*).

Open Access Article. Published on 27 March 2025. Downloaded on 4/8/2025 6:43:27 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

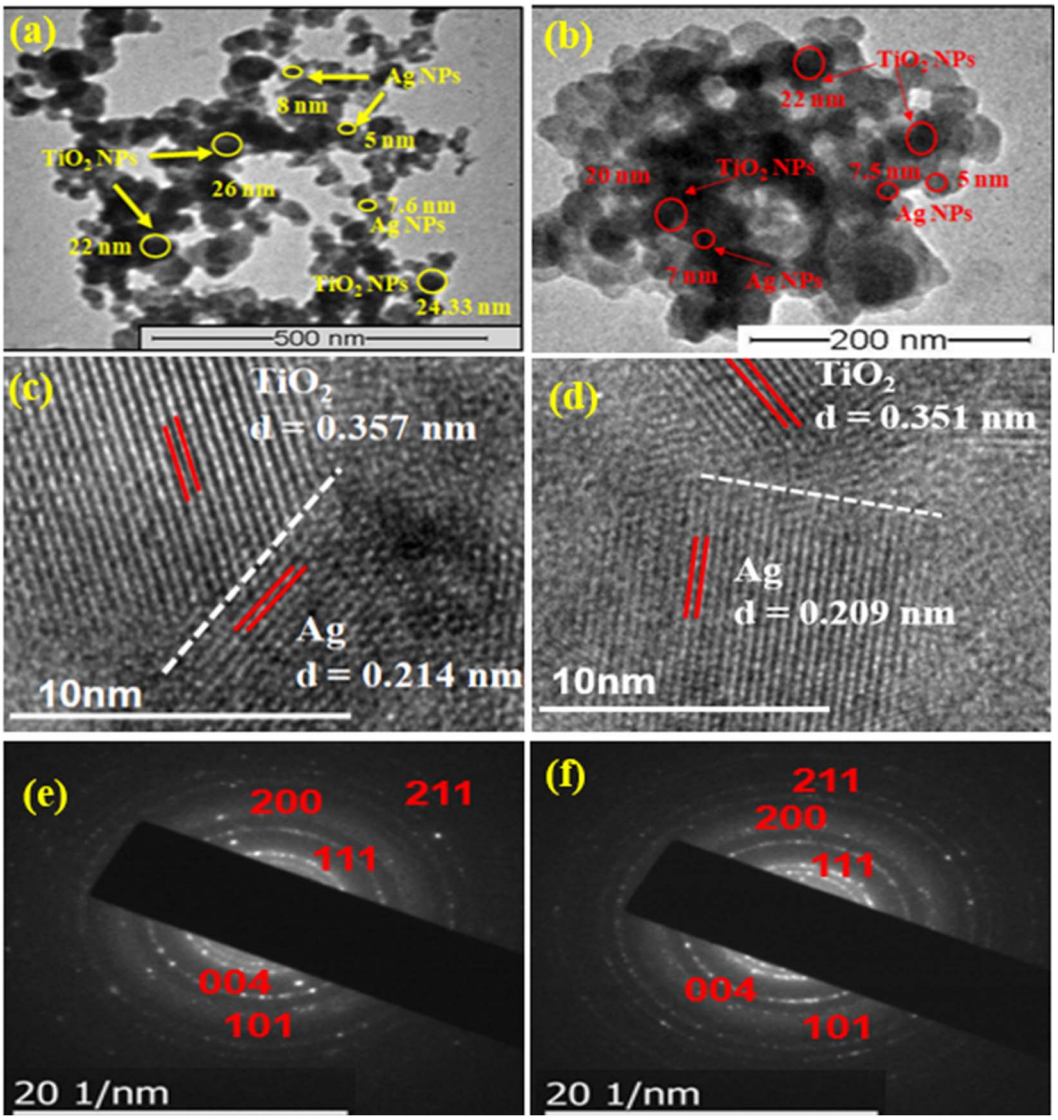


Fig. 5 (a and b) TEM images (c and d) HRTEM images (e and f) SAED images of Ag–TiO_2 NCs synthesized using *A. indica* and *M. indica* leaf extracts.

6. *In vitro* cytotoxicity of ATN and ATM in rat fibroblast cells

We evaluated the effect of ATN and ATM on the cytotoxicity and cell viability at three different concentrations (1/2×, 1×, and 2× of MIC) using the MTT assay in fibroblast cells. The results show that ATN and ATM did not affect the cell viability at any of the given concentrations, for up to 2× of MIC, as represented in Fig. 7(a) and (b). This suggests that ATN and ATM are non-toxic to fibroblast cells at the evaluated concentration of 2× of MIC.

7. Acute and sub-acute dermal toxicity

Based on the above performed antimicrobial tests, we chose 1×, 5× and 10× of MIC, *i.e.*, (625 μg mL^{-1}, 3125 μg mL^{-1} and 6250 μg mL^{-1}) for dermal toxicity. We observed that acute and sub-dermal toxicity testing of Ag–TiO_2 (*A. indica* and *M. indica*) NCs at up to 10× MIC in the CF rats did not cause any significant edema and erythema in the treated skin area in acute and sub-acute dermal toxicity (ESI: Fig. S1b.† Images of acute) (ESI: Fig. S1b.† Images of sub-acute). The erythema and edema in skin responses are each rated on a 0–4 scale individually. The irritation index was used to calculate the sum of the scores for erythema and edema. The primary irritation was calculated for each dose as the difference between the irritation index test site and control site, as reported by Kojic *et al.*[44] There was no sign of irritation and primary irritation index. These results show that Ag–TiO_2 is safe for short-term usage at up to 10× MIC. There was no significant skin damage or irritation, even with the compound used reputedly for 28 days. There was no significant behavioural change observed throughout the experiment.

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 27 March 2025. Downloaded on 4/8/2025 6:43:27 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

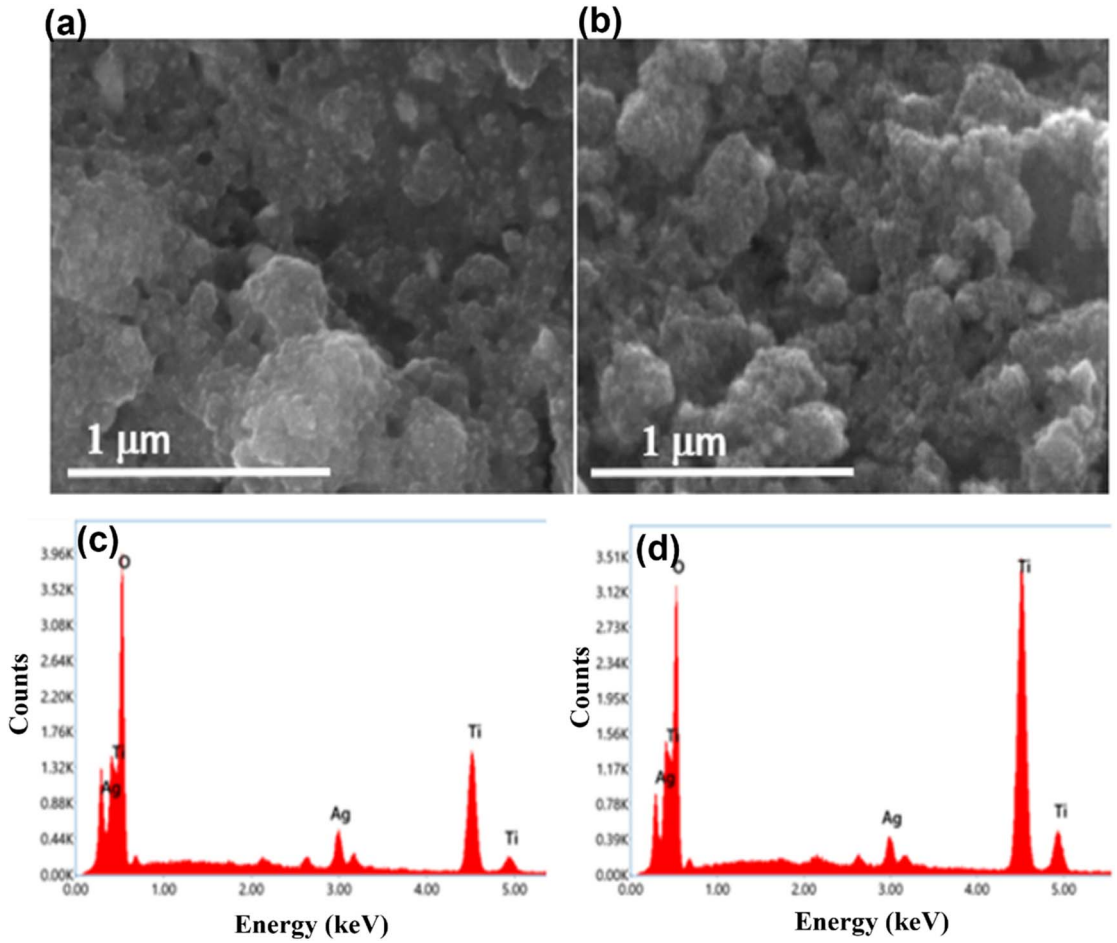


Fig. 6 (a and b) SEM images and (c and d) EDAX spectra of green-synthesized Ag–TiO_2 NCs (*A. indica* and *M. indica*).

Table 1 MICs of Ag–TiO_2 NCs (*A. indica* and *M. indica*)

Sample	*Staphylococcus aureus* (μg mL^{-1})	*Staphylococcus hominis* (μg mL^{-1})	*Escherichia coli* (μg mL^{-1})	*Salmonella typhimurium* (μg mL^{-1})	*Candida albicans* (μg mL^{-1})
MI	2500	2500	2500	2500	1250
AI	2500	2500	2500	2500	1250
ATM	**312.5**	2500	**625**	2500	**1250**
ATN	**625**	2500	**625**	2500	**312**
Vancomycin	0.0625	0.062	—	—	
Ofloxacin	—	—	≤0.0012	≤0.0017	—
Clotrimazole	—	—	—	—	≤0.0019

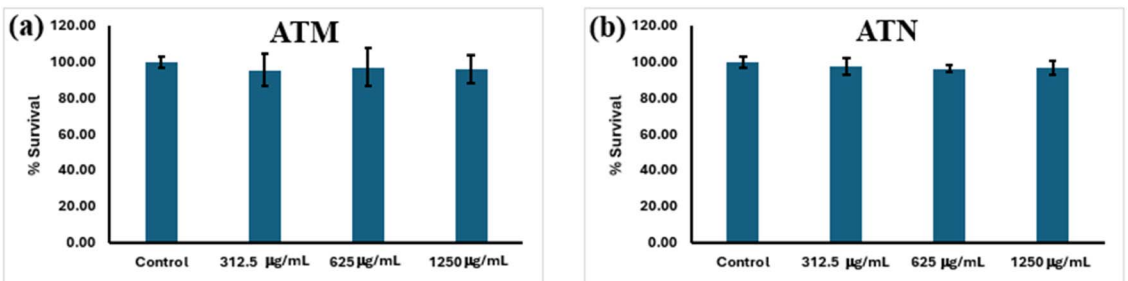


Fig. 7 (a and b) Percent cell viability of rat fibroblast cells treated with ATN and ATM (1/2×, 1×, and 2× of MIC) compared with the vehicle-treated cells using MTT [3-(4,5-dimethylthiazol-2-yl)-diphenyltetrazolium bromide] (mean ± SEM ($n = 4$)).

Open Access Article. Published on 27 March 2025. Downloaded on 4/8/2025 6:43:27 AM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Table 2 Irritation index (II) of Ag–TiO_2 NCs (*A. indica*) at 1×, 5×, and 10× of MIC (625 μg mL^{-1}, 3125 μg mL^{-1} and 6250 μg mL^{-1}) when applied once dermally for 28 days

Day	Control (sesame oil)	Ag–TiO_2 NCs (1× MIC)	Ag–TiO_2 NCs (5× MIC)	Ag–TiO_2 NCs (10× MIC)
0	0	0	0	0
3	0	0.75	0.5	0.5
7	0	0.75	0.25	0.25
11	0	0.75	0.75	0.5
14	0	0.75	0.5	0.5
17	0	1	0.5	0.75
21	0	0.25	0.75	1
24	0	1	0.75	0.5
28	0	0.75	1.75	1.5

7.1. Behavioural changes

The results from Ag–TiO_2 NCs (*A. indica* and *M. indica*) upon dermal application to CF rats are presented in Tables 2–5. The rats did not exhibit any significant behavioural abnormalities, skin effects, respiratory issues, difficulties in consuming food, or water, hair loss, or posture alignment.

7.2. Haematological, serum biochemical parameters and oxidative stress

The haematological profile was performed to understand whether the compound was causing any abnormalities. There were no significant changes observed in the total body weight, RBC, and WBC counts, as shown in Fig. S2(a) and (b).† Serum biochemistry also showed no substantial changes in the treated animals compared to the control. The cholesterol level, serum albumin, total serum protein, triglycerides, *etc.*, were comparable with the control group shown in S1, Tables 1 and 2.† No significant ($p > 0.05$) changes were observed in the biochemical parameters.

Nanomaterials are a developing technology with applications in science, and nanostructures are important tools within different areas of the field. However, there are various health hazards and environmental issues with chemically and physically synthesized NPs, and they are very expensive. A wide range of research has been carried out, and an alternative method has evolved, *i.e.*, green synthesis of the NCs. Medicinal plants are usually considered safe due to their natural occurrences. However, the toxicity and adverse effects of medicinal plant-based nanomaterials or composites need to be studied for regulatory clearances and their industrial application. Hence, we felt it was compulsory to perform *in vivo* dermal safety evaluation experiments for industrial uses to provide proven scientific evidence. Nowadays, dermal care products are gaining attention as individuals are showing greater interest in them. To present scientific evidence on environmental and health safety problems, this research incorporates both technical and societal difficulties.[45]

Nanomaterials have been studied *via* inhalation and oral and dermal routes based on biomedical needs.[46] Due to the unpredictability and ease of NCs exposure in typical applications, the dermal exposure route was selected amidst the possibilities.[47] AgNPs are being utilized in cosmetic consumer products, have the highest level of commercialization, and are employed in 30% of items that contain nanomaterials.[48]

The skin of rats was sensitized for 28 days, but no erythema or edema were observed (Tables 2 and 3). PII was also between 0–1, which means that the prepared $AgTiO_2$ NCs is safe to use in cosmetics for a short period (Tables 4 and 5). Acute and sub-acute toxicity help us to understand the toxic behaviour of compounds, and help to evaluate the morphological and physiological changes in the body.

For the acute toxicity test, Ag–TiO_2 NCs were left on rat's skin with different concentrations (625 mg mL^{-1} (1×), 3125 mg mL^{-1} (5×), and 6250 mg mL^{-1} (10×)) for 24 h, and there were

Table 3 Primary irritation index (PII) of Ag–TiO_2 NCs (*A. indica*) at 1×, 5×, and 10× of MIC when applied once dermally for 28 days

Day	Ag–TiO_2 NCs (1× MIC)	Ag–TiO_2 NCs (5× MIC)	Ag–TiO_2 NCs (10× MIC)
0	0	0	0
3	0.75	0.5	0.5
7	0.75	0.75	0.25
11	0.75	0.75	0.5
14	0.75	0.5	0.5
17	1	0.75	0.75
21	0.25	0.75	1
24	1	0.75	1.75
28	0.75	0.5	1.5

Table 4 Irritation index (II) of Ag–TiO_2 NCs (*M. indica*) at 1×, 5×, and 10× of MIC when applied once dermally for 28 days

Day	Control (sesame oil)	Ag–TiO_2 NCs (1× MIC)	Ag–TiO_2 NCs (5× MIC)	Ag–TiO_2 NCs (10× MIC)
0	0	0	0	0
3	0	0.75	0.5	0.5
7	0	0.75	0.25	0.25
11	0	0.75	0.75	0.5
14	0	0.75	0.5	0.5
17	0	1	0.5	0.75
21	0	0.25	0.75	1
24	0	1	0.75	0.5
28	0	0.75	1.75	1.5

Table 5 Primary irritation index (PII) of Ag–TiO_2NCs (*M. indica*) at 1×, 5×, and 10× of MIC when applied once dermally for 28 days

Day	Ag–TiO_2 NCs (1× MIC)	Ag–TiO_2 NCs (5× MIC)	Ag–TiO_2 NCs (10× MIC)
0	0	0	0
3	0.75	0.5	0.75
7	0.75	0.75	0.25
11	0.75	0.5	0.75
14	0.75	0.5	0.5
17	1	0.5	0.75
21	0.25	0.75	1
24	0.75	1	1
28	1.25	1.25	1.5

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 27 March 2025. Downloaded on 4/8/2025 6:43:27 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

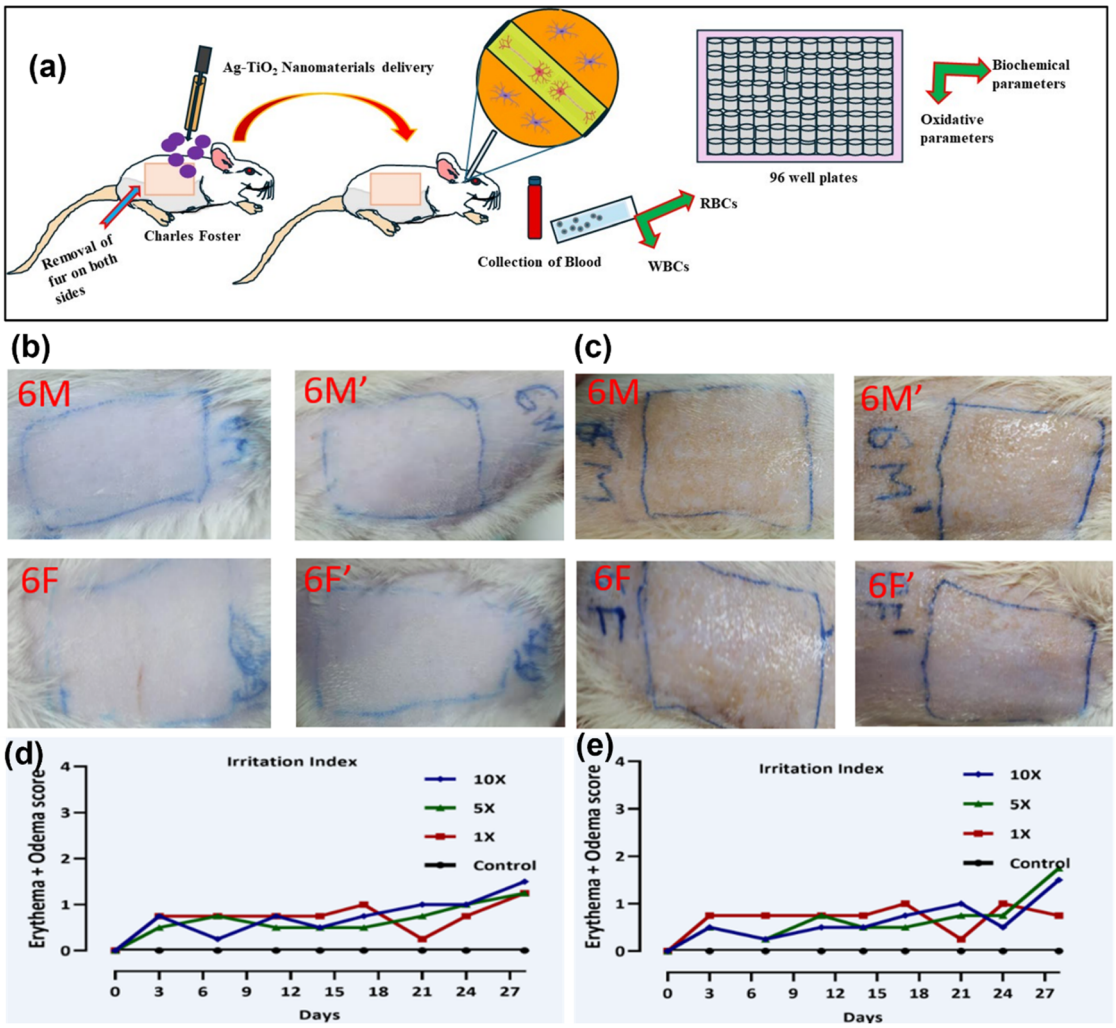


Fig. 8 (a) Illustration of the nanomaterial toxicity in Charles Foster species, (b and c) observation of high doses of Ag–TiO_2 NCs (derived from different plants), and (d and e) irritation index of both *A. indica* and *M. indica* Ag–TiO_2 NC samples.

no significant changes in the body weight and no skin irritation was observed in the rats. Furthermore, the sub-acute toxicity test after 28 days of Ag–TiO_2 NCs exposure on the skin with the above-mentioned concentration did not induce any skin irritation, as observed in Fig. S3(a) and (b).† The body weight, biochemical, and oxidative stress parameters, with repeated administration of Ag–TiO_2 NCs, were applied and compared with the control animals (applied only sesame oil) (S1 Tables 1 and 2†). The depiction of nanomaterials in Charles Foster rats has been studied using biochemicals, oxidative parameters, and irritation index, as illustrated in Fig. 8(a)–(e). The toxic traits of the current $AgTiO_2$ NCs are represented in Fig. 8(a), from removing fur of CF rats to testing for WBCs, RBCs, and other parameters. Fig. 8(b) and (c) shows the first and last day of $AgTiO_2$ NCs formulation on CF rats. Similarly, Fig. 8(d) and (e) depicts the irritation index behaviour of $AgTiO_2$ NCs after 28 days.

Based on the experimental observations, it is evident that the $AgTiO_2$ NCs can be used in cosmetics and topical formulations. However, further clinical studies are required to validate and ensure the safety and efficacy of $AgTiO_2$ NCs. NCs are essential to be used as a scientifically verified product. A thorough physicochemical characterization is crucial for establishing the link between NPs and toxicological reactions.[49] These statistics provide basic information about NPs that may be applied to risk evaluation in nanosafety. Modern techniques, including TEM, XRD, UV-vis, DLS, FTIR, and zeta-potential, were employed to evaluate the morphology, crystallinity, particle size, functional groups, and surface charge for Ag–TiO_2 NCs identification. They may have an impact on Ag–TiO_2 NCs toxicity.[50] Shaaban *et al.* mentioned that the AgNPs hydrogel formulation with a 2% addition showed wound healing of the skin without any side effects of erythema, edema, or irritation and hence can be well corroborated with the current observation.[51]

8. Conclusion

The results of this investigation demonstrate that Ag–TiO_2 NCs synthesized using *A. indica* and *M. indica* leaf extract have shown no harmful effects on animals during dermal

Open Access Article. Published on 27 March 2025. Downloaded on 4/8/2025 6:43:27 AM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

application. We evaluated the effect of ATN and ATM on the cytotoxicity and cell viability at three different concentrations (1/2×, 1×, and 2× of MIC) using the MTT assay in fibroblast cells showing no toxic behaviour up to 2× of MIC. Rats treated with different doses of Ag–TiO_2 NCs (1×, 5×, and 10×) exhibited no side effects or abnormal symptoms over 4 weeks, as confirmed by the irritation index. Moreover, haematological, biochemical, and oxidative stress assessments revealed no differences between the control and treatment groups. Based on these findings, it can be concluded that green-synthesized Ag–TiO_2 NCs in concentrations up to 10× of MIC (6250 μg mL^{-1}) can be deemed safe for use in rat models for cosmetics and drug delivery. However, further investigation is necessary to evaluate the long-term exposure effects of green-synthesized Ag–TiO_2 NCs in cosmetics and medical treatments.

Data availability

Raw data were generated at the National Physical Laboratory, Pusa Road, Delhi and Central Institute of Medicinal and Aromatic Plants, Lucknow, Uttar Pradesh.

Author contributions

Archana Rana: conceptualization, writing – original draft, methodology, investigation; Shweta Parashar: formal analysis, supervision, investigation; Diksha Singh and Kavita Singh: formal analysis, investigation; Debabrata Chanda: resources, validation, research, software; Anirban Pal: resources, validation, research, software; Ritu Srivastava: supervision, and validation; Shailesh Narain Sharma: supervision, validation, conceptualization and editing.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgements

The author Archana Rana would like to thank the University Grants Commission (UGC-90802) for granting a Research Fellowship, as well as AcSIR for providing the resources needed for conducting this experiment.

References

1 A. C. Santos, J. Marto, R. Cha-Cha and A. M. Martins, Nanotechnology-based sunscreens-a review, *Mater. Today Chem.*, 2022, **23**, 100709.
2 A. Rana, A. K. Chaudhary, *et al.*, An Investigation of antimicrobial activity for plant pathogens by green-synthesized silver nanoparticles using Azadirachata indica and Mangifera indica, *Physchem*, 2023, **3**, 125–146.
3 P. L. Sanches, L. R. O. Geaquinto, R. Cruz, D. C. Schuck, M. Lorencini, J. M. Granjeiro and A. R. L. Ribeiro, Toxicity Evaluation of TiO_2 Nanoparticles on the 3D Skin Model: A Systematic Review, *Front. Bioeng. Biotechnol.*, 2020, **8**, 575.
4 Y. Liang, A. Simaiti, M. Xu, S. Lv, H. Jiang, X. He, Y. Fan, S. Zhu, B. Du, W. Yang, X. Li and P. Yu, Antagonistic Skin Toxicity of Co-Exposure to Physical Sunscreen Ingredients Zinc Oxide and Titanium Dioxide Nanoparticles, *Nanomaterials*, 2022, **12**(16), 2769.
5 G. M. Nair, T. Sajini and B. Mathew, Advanced green approaches for metal and metal oxide nanoparticles synthesis and their environmental applications, *Talanta Open*, 2022, **5**, 1000800.
6 A. Rana, S. Pathak, D. K. Lim, *et al.*, Recent Advancements in Plant- and Microbe-Mediated Synthesis of Metal and Metal Oxide Nanomaterials and Their Emerging Antimicrobial Applications, *ACS Appl. Nano Mater.*, 2023, **6**(10), 8106–8134.
7 D. Kwek, D. Yu, M. I. Setyawati, A. Gautam, S. S. Adav, E. C. Cheong and K. W. Ng, Understanding the toxicological effects of TiO2 nanoparticles extracted from sunscreens on human keratinocytes and skin explants, *Part. Fibre Toxicol.*, 2024, **21**(1), 1–16.
8 M. Nilavukkarasi, S. Vijaykumar, M. Kalaskar, N. Gurav, S. Gurav, P. K. Praseetha and L. Cappariszeylanica, conjugated TiO_2 nanoparticles as bio-enhancers for antimicrobial and chronic wound repair, *Biochem. Biophys. Res. Commun.*, 2022, **623**, 127–132.
9 S. Li, J. Zeng, D. Yin, P. Liao and S. Ding, Synergic fabrication of titanium dioxide incorporation into heparin-polyvinyl alcohol nanocomposite: enhanced *in vitro* antibacterial activity and care of *in vivo* burn injury, *Mater. Res. Express*, 2021, **8**, 085012.
10 Y. Y. Ho, D. S. Sun and H. H. Chang, Silver Nanoparticles Protect Skin from Ultraviolet B-Induced Damage in Mice, *Int. J. Mol. Sci.*, 2020, **21**(19), 7082.
11 G. Chinnasamy, S. Chandrasekharan, T. W. Koh and S. Bhatnagar, Synthesis, Characterization, Antibacterial and Wound Healing Efficacy of Silver Nanoparticles From Azadirachta indica, *Front. Microbiol.*, 2021, **12**, 611560.
12 J. S. Kim, K. S. Song, J. H. Sung, H. R. Ryu, B. G. Choi, H. S. Cho and I. J. Yu, Genotoxicity, acute oral and dermal toxicity, eye and dermal irritation and corrosion and skin sensitization evaluation of silver nanoparticles, *Nanotoxicology*, 2012, **7**(5), 953–960.
13 K. M. Elattar, F. O. Al-Otibi, M. S. El-Hersh, A. A. Attia, N. M. Eldadamony, A. Elsayed, F. Menaa and W. I. A. Saber, Multifaceted chemical and bioactive features of Ag@ TiO_2 and Ag@ SeO_2 core/shell nanoparticles biosynthesized using Beta vulgaris L. extract, *Heliyon*, 2024, **10**(7), e28359.
14 N. Abbasi, H. Ghaneialvar and R. Moradi, Formulation and characterization of a novel cutaneous wound healing ointment by silver nanoparticles containing *Citrus lemon* leaf: A chemobiological study, *Arabian J. Chem.*, 2021, **14**(7), 103246.
15 A. I. Nicoara, V. L. Ene, B. B. Voicu, M. A. Bucur, I. A. Neacsu, B. S. Vasile and F. Iordache, Biocompatible Ag/Fe-Enhanced TiO_2 Nanoparticles as an Effective Compound in Sunscreens, *Nanomaterials*, 2020, **10**(3), 570.
16 M. Saraswati, R. L. Permadani, A. Slamet, *et al.*, The innovation of antimicrobial and self-cleaning using Ag/

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 27 March 2025. Downloaded on 4/8/2025 6:43:27 AM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

TiO_2 nanocomposite coated on cotton fabric for footwear application, *IOP Conf. Ser.:Mater. Sci. Eng.*, 2021, **509**, 012091.

17 Y. Hou, A. Mushtaq, Z. Tang, E. Dempsey, Y. Wu, *et al.*, ROS-responsive Ag-TiO_2 hybrid nanorods for enhanced photodynamic therapy of breast cancer and antimicrobial applications, *J. Sci.:Adv. Mater. Devices*, 2022, **7**(2), 100417.

18 S. M. Gupta and M. Tripathi, A review of TiO_2 nanoparticles, *Chin. Sci. Bull.*, 2011, **56**, 1639–1657.

19 Y. Zhang, F. Fu, Y. Li, D. Zhang and Y. Chen, One-Step Synthesis of Ag@TiO_2 Nanoparticles for Enhanced Photocatalytic Performance, *Nanomaterials*, 2018, **8**(12), 1032.

20 F. Khojasteh, M. R. Mersagh and H. Hashemipour, The influences of Ni, Ag-doped TiO_2 and SnO_2, Ag-doped SnO_2/TiO_2 nanocomposites on recombination reduction in dye synthesized solar cells, *J. Alloys Compd.*, 2022, **890**, 161709.

21 G. J. Nohynek, E. Antignac, T. Re and H. Toutain, Safety assessment of personal care products/cosmetics and their ingredients, *Toxicol. Appl. Pharmacol.*, 2010, **243**, 239–259.

22 D. J. Dheaa and S. A. Mahdi, Photodynamic therapy and TiO_2-decorated Ag nanoparticles: Implications for skin cancer treatment, *Rev. Compos. Mater. Av.*, 2023, **33**(6), 393–398.

23 M. Ahamed, M. A. M. Khan, M. J. Akhtar, *et al.*, Ag-doping regulates the cytotoxicity of TiO_2 nanoparticles *via* oxidative stress in human cancer cells, *Sci. Rep.*, 2017, **7**, 17662.

24 G. Li, X. Tan, W. Zhao, M. A. I. A'srai and M. H. Razali, In-vitro and in-vivo wound healing studies of Ag@TiO2NRs/GG hydrogel film for skin tissue regeneration, *Mater. Res. Express*, 2023, **10**, 045401.

25 D. Bhardwaj and R. Singh, Green biomimetic synthesis of Ag–TiO_2 nanocomposite using *Origanum majorana* leaf extract under sonication and their biological activities, *Bioresources and Bioprocessing*, 2021, **8**, 1.

26 A. Rana, A. Kumari and A. K. Chaudhary, An investigation of Antimicrobial activity for plant pathogens by Green-Synthesized Silver Nanoparticles using Azadirachata indica and Mangifera indica, *Physchem*, 2023, **3**, 125–146.

27 A. Rana, S. Pathak and K. Kumar, Multifaceted Properties of TiO2 Nanoparticles Synthesized via Mangifera indica and Azadirachta indica Plant Extracts: Antimicrobial, Antioxidant, and Non-linear Optical Investigations for Sustainable Agriculture Applications, *Mater. Adv.*, 2024, **5**, 2767.

28 S. Srivastava, R. K. Lal, K. Yadav, Y. Pant, L. Bawitlung, P. Kumar and C. S. Chanotiya, Chemical composition of phenylpropanoid rich chemotypes of Ocimumbasilicum L. and their antimicrobial activities, *Ind. Crops Prod.*, 2022, **183**, 114978.

29 J. C. Stockert, A. Blázquez-Castro, M. Cañete, R. W. Horobin and Á. Villanueva, MTT assay for cell viability: Intracellular localization of the formazan product is in lipid droplets, *Acta Histochem.*, 2012, **114**(8), 785–796.

30 A. D. Ankomah, Y. D. Boakye, T. A. Agana, V. E. Boamah, P. P. S. Ossei, F. Adu and C. Agyare, Evaluation of Dermal Toxicity and Wound Healing Activity of *Cnestis ferruginea* Vahl ex DC, *Adv. Pharmacol. Pharm. Sci.*, 2022, 5268613.

31 A. Rana, A. K. Chaudhary, S. Saini, R. Srivastava, M. Kumar and S. N. Sharma, Ultrafast transient absorption spectroscopic (UFTAS) and antibacterial efficacy studies of phytofabricated silver nanoparticles using Ocimum Sanctum leaf extract, *Inorg. Chem. Commun.*, 2023, **147**, 110233.

32 Y. Zhang, F. Fu, Y. Li, D. Zhang and Y. Chen, One-Step Synthesis of Ag@TiO2 Nanoparticles for Enhanced Photocatalytic Performance, *Nanomaterials*, 2018, **8**(12), 1032.

33 T. N. Rao, Riyazuddin, P. Babji, N. Ahmad, *et al.*, Green synthesis and structural classification of *Acacia nilotica* mediated-silver doped titanium oxide (Ag/TiO_2) spherical nanoparticles: Assessment of its antimicrobial and anticancer activity, *Saudi J. Biol. Sci.*, 2019, **26**, 1385–1391.

34 F. Caputo, J. Clogston, L. Calzolai, M. Rösslein and A. Prina-Mello, Measuring particle size distribution of nanoparticle enabled medicinal products, the joint view of EUNCL and NCI-NCL. A step-by-step approach combining orthogonal measurements with increasing complexity, *J. Controlled Release*, 2019, **299**, 31–43.

35 D. Bhardwaj, R. Singh, *et al.*, Green biomimetic synthesis of Ag–TiO_2 nanocomposite using Origanummajorana leaf extract under sonication and their biological activities, *Bioresources and Bioprocessing*, 2021, **8**, 1.

36 H. Iqbal, A. Razzaq, B. Uzair, N. Ul Ain, S. Sajjad, N. A. Althobaiti, A. E. Albalawi, B. Menaa, M. Haroon, M. Khan, N. U. Khan and F. Menaa, Breast Cancer Inhibition by Biosynthesized Titanium Dioxide Nanoparticles Is Comparable to Free Doxorubicin but Appeared Safer in BALB/c Mice, *Materials*, 2021, **14**(12), 3155.

37 X. H. Yang, H. T. Fu, X. C. Wang, J. L. Yang, X. C. Jiang and A. B. Yu, Synthesis of silver-titanium dioxide nanocomposites for antimicrobial applications, *J. Nanopart. Res.*, 2014, **16**(8), 2526.

38 S. P. Lim, S. Mehmood and P. Rameshkumar, Amperometric detection of hydrogen peroxide and its density functional theory for adsorption on Ag/TiO2 nanohybrid, *J. Mater. Sci.:Mater. Electron.*, 2020, **31**(8), 6017–6026.

39 B. Govindan and A. R. Soliman, Sol-Gel-Assisted Microwave-Derived Synthesis of Anatase Ag/TiO2/GO Nanohybrids toward Efficient Visible Light Phenol Degradation, *Catalysts*, 2017, **7**(5), 133.

40 M. M. Zahornyi, N. I. Tyschenko, T. F. Lobunets, O. F. Kolomys, V. V. Strelchuk and K. S. Naumenko, The Ag influence on the surface states of TiO_2, Optical activity and its cytotoxicity, *J. Nano- Electron. Phys.*, 2021, **13**(6), 06009.

41 T. Ali, A. Ahmed, U. Alam, I. Uddin, P. Tripathi and M. &Muneer, Enhanced photocatalytic and antibacterial activities of Ag-doped TiO_2 nanoparticles under visible light, *Mater. Chem. Phys.*, 2018, **212**, 325–335.

42 T. N. Rao, Riyazuddin, P. Babji, N. Ahmad, *et al.*, Green synthesis and structural classification of *Acacia nilotica* mediated-silver doped titanium oxide (Ag/TiO_2) spherical

Open Access Article. Published on 27 March 2025. Downloaded on 4/8/2025 6:43:27 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

nanoparticles: Assessment of its antimicrobial and anticancer activity, *Saudi J. Biol. Sci.*, 2019, **26**, 1385–1391.

43 S. Bhutani, T. Kalaivani and P. Anilkumar, Synthesis of Ag decorated TiO 2 nanoparticles for ammonia gas sensor application, *J. Environ. Nanotechnol.*, 2021, **10**, 25–34.

44 Z. Kojic, D. Stojanovic, S. Popadic, M. Jokanovic and D. Janackovic, The irritative property of a-tricalcium phosphate to the rabbit skin, *Gen. Physiol. Biophys.*, 2009, **28**, 168–173.

45 S. A. S. Nurul, H. Hazilawati, R. S. Mohd, *et al.*, Subacute Oral Toxicity Assesment of Ethanol Extract of Mariposa christiavespertilionis Leaves in Male Sprague Dawley Rats, *Toxicol. Res.*, 2018, **34**, 85–95.

46 Y. Li and E. Cummins, Hazard characterization of silver nanoparticles for human exposure routes, *J. Environ. Sci. Health, Part A: Toxic/Hazard. Subst. Environ. Eng.*, 2020, **55**(6), 704–725.

47 D. B. Warheit and E. M. Donner, Risk assessment strategies for nanoscale and fine-sized titanium dioxide particles: Recognizing hazard and exposure issues, *Food Chem. Toxicol.*, 2015, **1**(85), 138–147.

48 R. Gautam, S. Yang, A. Maharjan, J. Jo, M. Acharya, Y. Heo and C. Kim, Prediction of Skin Sensitization Potential of Silver and Zinc Oxide Nanoparticles Through the Human Cell Line Activation Test, *Frontiers in Toxicology*, 2021, **3**, 649666.

49 K. Rasmussen, H. Rauscher, A. Mech, J. R. Sintes, D. Gilliland, M. González, P. Kearns, K. Moss, M. Visser, M. Groenewold and E. A. Bleeker, Physico-chemical properties of manufactured nanomaterials-Characterisation and relevant methods. An outlook based on the OECD Testing Programme, *Regul. Toxicol. Pharmacol.*, 2018, **92**, 8–28.

50 J. McClements and D. J. McClements, Standardization of nanoparticle characterization: methods for testing properties, stability, and functionality of edible nanoparticles, *Crit. Rev. Food Sci. Nutr.*, 2016, **56**(8), 1334–1362.

51 S. Shaaban, M. S. S. Adam and N. M. EI-Metwaly, Novel organoselenium-based N-mealanilic acid and its zinc (II) chelate: Catalytic, anticancer, antimicrobial, antioxidant, and computational assessments, *J. Mol. Liq.*, 2022, **363**, 119907.

© 2025 The Author(s). Published by the Royal Society of Chemistry

Current Psychology
https://doi.org/10.1007/s12144-025-07782-3

Exploring the psychometric properties of the adult hope scale in an Indian university student sample

Suhans Bansal[1] · **Naval Garg**[2] · **Jagvinder Singh**[3]

Accepted: 31 March 2025
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2025

Abstract
The present study explores the psychometric properties of Synder et al.'s (1991) Adult Hope Scale in the Indian milieu and is based on three independent studies. Study 1 examines the factor structure of the scale using Exploratory Factor Analysis and Confirmatory Factor Analysis. Study 2 tests the internal consistency reliability using Cronbach's alpha and Composite Reliability. It aslo assesses convergent, discriminant, and criterion validities using the Average Variance Extracted, Fornell and Larcker testing system, and correlation coefficients between hope, gratitude, spiritual well-being, and distress scores. The final study examines the stability of the scale using test-retest reliability. The results of Study 1 suggested a two-factored (agency and pathways thinking) structure alike the original study. The second study asserted acceptable reliability and validity estimates and convergence with gratitude questionnaire- 6 (GQ- 6) and divergence with psychological distress scale (K10) scores. Furthermore, the scale successfully predicted the spiritual well-being of Indian students, thereby suggesting the scale's predictive validity in the Indian context. Finally, the test-retest reliability confirmed the stability of scale among Indian students. The study is one of the first to test the reliability and validity of the Adult Hope Scale and its associations with spiritual well-being, gratitude, and psychological distress, specifically in India.

Keywords Hope · India · Spiritual well-being · Gratitude · Psychological distress

Introduction

Hope is among the most revered concepts explored in various fields, including psychology, medical, organizational behavior, and theology (Bowers & Bowers, 2023). However, it has been defined differently by different researchers. Fromm (1968) suggested hope as a state of being. Lazarus (1999) expressed hope as an emotion. Waterworth (2004) labelled it as a cognitive process, and Pettit (2004) understood hope as a state of mind. Various researchers explain hope as a profound and all-prevailing human emotion, embodying optimism and a strong belief in the probability of the occurrence of positive events (Doe, 2020). It is one of the central facets of human nature that fuels resilience and encourages people to strive for a better future. It helps people realize a sense of purpose and gives them the courage to tackle challenges (Cheavens et al., 2019). Basically, it is an instilled longing for something better or a change from the present conditions and arises when people imagine a promising future that aligns with their ambitions (Doe, 2020).

Researchers like Garg (2023) compared hope with a torch that guides people through their dark and low times by reminding them that a new, bright, cherished, and improved future lies ahead. Researchers like Doe (2020) and Colla et al. (2022) highlighted that hope, accompanied by a sense of expectation, trust in positive future events, and faith in the ability and resources, encourages people to set and execute goals. At the same time, researchers also explain that hope recognizes the realism of the current situations and hindrances, thereby not restricting it to blind sanguinity and aspirant thinking (Garg, 2023). Instead, it acts as a dynamic and flexible mindset allowing people to reframe

✉ Naval Garg
naval.garg@dtu.ac.in

Suhans Bansal
suhansbansal@gmail.com

Jagvinder Singh
jagvinder.singh@gmail.com

[1] Delhi Technological University, New Delhi, India

[2] Netaji Subhas University of Technology, Delhi, India

[3] University of Delhi, New Delhi, India

Published online: 14 April 2025

disappointments and hindrances and motivate them to learn from their mistakes and change strategies accordingly (Doe, 2020). The nature of hope has attracted the world's attention, contributing to conceptualizing various theories and models involving hope.

Various researchers and social scientists have contributed to the literature and have presented seminal works on hope. Authors have conceptualized the psychodynamic perspective of hope as a fundamental basis for transformation. Further, the cognitive-behavioral perception of hope visualizes hope as "a desire to formulate goals or an individual's attempt to receive positive reinforcement by learning new behaviors" (Doe, 2020, p. 56). Synder et al.'s (1991) work on hope and the subsequent development of the *Adult Hope Scale* (AHS) has been regarded as one of the most pivotal contributions to the knowledge of hope. Snyder (2002) defined hope as "the perceived capability to derive pathways to desired goals and to motivate oneself via agency thinking to use those pathways" (p. 249). As per Synder et al.'s (1991) hope theory, hopeful thinking necessitates the effective amalgamation of agency and pathway thinking to account for future endeavours (Snyder, 2002). The development of pathways indicates an 'activity-oriented approach' when one has to achieve a major goal that has smaller objectives to achieve as well. These pathways fuel positive emotions in people that help them achieve the overarching goal (Yaghoobzadeh et al., 2018). In its dual facets, agency and pathways, hope is explained as one of the major positive sentiments that help individuals meet positive future events (Doe, 2020). The two factors of hope present goal-focused fortitude and a plan to meet those goals (Snyder et al., 1991).

Several studies have examined the antecedents and benefits of hope in different fields of study. Different scholars have asserted that hope leads to subjective well-being (Doe, 2020), greater purpose in life (Aghababaei et al., 2016), optimism (Hutz et al., 2014), gratitude (Witvliet et al., 2018), academic achievement, social change (Greenaway et al., 2016), resilience (Doe, 2020), and harmony in life and vitality (Doe, 2020). Other researchers reported a negative linkage of hope with depression and anxiety (Doe, 2020), posttraumatic disorder (Cheng et al., 2023), probability of developing a disease (Doe, 2020), and stress (Warber et al., 2011). Further, studies in the healthcare sector suggested that hope acts like an inner voice, the voice of others, or a motivator that helps patients maintain faith in a positive future and draws upon their spiritual well-being to respond to their current situation positively (Espedal, 2021).

The importance of hope further increased during the Covid- 19 period. During the period of covid 19, when the world halted its physical activities due to lockdown, people around the globe wished for faster control of the outbreak, and the 'hope' kept our expectations for a better future alive (Garg, 2023). The horrors of hospitals during the pandemic severely impacted medical personnel worldwide. Various studies on the mental health of medical practitioners suggested that hope and optimism positively moderated the mental health of medical practitioners (Doe, 2020). Studies on mental health further suggest that hope acts as an agent to reduce the negative impact of stress and anxiety and helps in faster recovery of mental health (Doe, 2020).

Motivation for the study

Positive psychology (1.0) asserts the role of positive emotions like hope and gratitude in ensuring well-being in normal times; however, positive psychology (2.0) looks beyond normal times. It tends to observe the efficacy of any positive emotion-based interventions in challenging times (Susarla et al., 2021). With the world facing tremendous challenges in the form of the Ukraine-Russia war, the rise of the Taliban in Afghanistan, impending global recession, the new virtual cold war, climate crisis, throat-cut competition, materialism, and greed, there is a need to reassess the implications of constructs like hope and gratitude amid these difficult times (Garg, 2023). Also, studies have shown strong linkages between gratitude and hope (Witvliet et al., 2018) during harsh times, suggesting a strong connection between the two positive emotions and the necessity to practice them in daily life to bounce back from such times (Susarla et al., 2021). Such studies have suggested that people should know to whom they are grateful, like their soldiers, doctors and others, and always hope to overcome difficult times (Susarla et al., 2021). Studies have also reported that people with higher levels of hope during difficult times, show a presence of spiritual well-being compared to those with lower levels or no hope of surviving such times (Cheng et al., 2023). Therefore, the present study tries to reignite the interest of scientists, academicians, and researchers in analysing hope, especially in a country like India, which is struggling with its own set of problems like population explosion, unemployment, price rise, communalism, casteism, poverty, wealth and digital divide, the disintegration of the robust joint family system, erosion of cultural values and health-related issues. The present study was conducted in the Indian context for the following plausible reasons.

First, like the rest of the world, India is also witnessing a sharp increase in cases related to mental health like psychological distress, anxiety, depression, loneliness, and suicidal tendencies. However, what makes the situation worse is the fact that mental health remains a taboo in Indian society. Talks and discussions over mental issues are still uncommon, especially among Indian teens and collegiates (Gaiha et al., 2020). Psychiatrists/psychologists tend to develop small interventions based on different positive constructs

like gratitude, hope, and optimism for both preventive and curative purposes. Thus, it is essential to understand the Indian notion of hope.

Second, hope has special religious and philosophical connotations in India. The Indian philosophy (*darshan shastra)* highlights the need to understand the context and meaning of hope. *Swami Vivekananda* has highlighted the importance of hope in human life. He explains that in a world governed by sufferings and the cycle of birth and death, people hope they never get ill or die (Vivekananda, 2017). Further, *Advait Vedant*, one of the most prominent Indian schools of *darshan shastra*, suggests hope is the essence of life that enables and motivates an individual to quench the thirst for answers to questions such as self, the god, creation, or our objectives of wondering in this world (Namastu, 2018). This school explains that hope keeps the light of pursuit lit even in the lowest periods of one's life and connects them with God ("Hope for a Good Life," 2018). Shree Aurobindo explains that the *rajasic* actions (*karmas)* are performed under the influence of desire and hope to achieve the desired fruits (Namastu, 2018). Lord Krishna explains that every action has its karmic bondage, which awards the appropriate results of the actions done. Hence, for every righteous action, people hope for favorable rewards. The Lord further explains that liberated people are free from the karmic bondages, the cycle of birth and death, and the hope and sorrows of life (Namastu, 2018). India, a deeply religious country, is poised to accept any religiously ingrained intervention more easily and thoroughly than any other intervention. Therefore, it makes it more worthy of exploring hope in the Indian context.

Hope scales and adult hope scale

However, researchers have emphasized the crucial requirement of having a scale to precisely measure a phenomenon while conducting research (Garg, 2023). Regrettably, the existing Indian literature almost lacks studies that have either established a new scale or validated an existing one to assess hope in Indian collegiate. Previous researchers have developed and validated a number of hope scales like the AHS (Snyder et al., 1991), Herth Hope Index (HHI) (Herth, 1992), Integrative Hope Scale (IHS) (Schrank et al., 2011), Comprehensive Hope Scale (CHS) (Scioli et al., 2011), etcetera. Snyder et al.'s (1991) AHS is a twelve-item self-reported. It captures the trait hope in adults and college students. It is a two-factor scale: agency and pathways. Four items capture agency, the other four pathways, and the remaining four are filler items (Snyder et al., 1991). The items are rated on an eight-point Likert scale ranging from 1 (definitely false) to 8 (definitely true). The agency dimension denotes a sense of willpower to meet goals, whereas pathways are the ability to develop successful plans to realize these goals. Initially, Snyder et al. (1991) administered 45 items of the AHS to 384 students at the University of Kansas to explore the psychometric properties of the originally developed AHS. The reliability estimate-Cronbach's alpha of agency ranged between 0.71 and 0.76, and pathways ranged between 0.63 and 0.80. The authors also conducted temporal stability using test-retest reliability. The results of temporal stability reported significant coefficients between the different periods of tests. Finally, the convergent validity was tested on the data collected from 241 students. AHS positively correlated with the life orientation test (LOT), control perceptions, and esteem.

The present study selected AHS for psychometric evaluation for three reasons. First, the AHS scale measures trait hope, which is a stable and enduring characteristic than emotion and mood-based conceptualization of hope. Trait hope is one's general tendency to approach the goal with optimism and positivity. Second, AHS is based on the more comprehensive conceptualization of hope with two distinct components: agency thinking and pathways thinking. 'Agency' thinking represents one's capability to recognize, pursue, and achieve goals. 'Pathways' thinking reflects one's perceived ability to frame paths, strategies, and plans to achieve expressed goals. Thus, AHS offers a more comprehensive and holistic assessment of dispositional hope.

Third, AHS is one of the most widely used hope scales with psychometric properties examined in different settings. Redlich-Amirav et al. (2018) reported that thirty-one studies had explored the psychometric properties of AHS in different cultural contexts. And the scale has been adopted in these settings with/without modifications. Compared to AHS, the psychometric properties of the HHI, HIS and CHS have been evaluated only one to three times (Redlich-Amirav et al., 2018). It reflected the wide acceptability and adaptability of the AHS scale. This meant that AHS was already being used in India. However, upon conducting a detailed literature review, it was found that although AHS was used in the Indian context, a formal investigation of its psychometric properties was required, more specifically following the guidelines provided in STANDARDS (American Educational Research Association et al., 2014) and that too in the collegiates. Further, previous studies have not examined its second-order CFA and test-retest reliability. Also, earlier studies evaluated only Cronbach's Alpha to validate the scale's internal consistency reliability. Additionally, India is a diverse nation with varying cultural and linguistic backgrounds, so the psychometric properties of a psychometric instrument may vary accordingly. Therefore, based on the recommendation from authors like Deb et al. (2023), Garg (2023) and Kapur and Teotia (2018) to conduct a formal investigation of psychometric properties of AHS in

India, the current study not only examines the psychometric properties of AHS in India, it also bridges the gap in the literature by conducting a test-retest reliability study and uses more robust statistical techniques like McDonalds' Omega, average variance explained, Fornell and Larcker testing system among others to test the reliability and validity of AHS in India, that too among Indian university students.

Lastly, researchers like Garg (2023) have argued that when considering the adult population, researchers may tend to ignore university students or may have to ignore certain groups of university students based on the definition of adulthood specified in their respective national laws. Also, university life is considered a transition stage in which students face various challenges related to their academic performance, personal identity, and career planning issues, such as unemployment owing to the growing Indian population and relationship issues (Garg, 2023). The significance of such challenges has taken a toll on university students. Studies have shown that Indian university students suffer from various mental health issues like increased stress, depression, anxiety and identity issues due to challenges discussed earlier (Gaiha et al., 2020). Also, university students represent a new generation, which has moral and social values, beliefs, and attitudes developed based on various environmental factors affecting them. Hence, the challenges and generational differences that were previously discussed help create a unique perspective of hope among them. Also, these perspectives not only differ based on Western and Eastern cultural differences but also vary from state to state in a country like India, which calls for unity in its diversity. Therefore, it has become imperative to formally test the reliability and validity of the AHS in this context to ensure that the scale reflects how Indian university students experience hope.

Hence, the discussed motives mandate an examination of the psychometric properties of AHS in India, specifically, among Indian university students. The study's methodology is based on the guidelines provided by the American Educational Research Association et al. (2014). The first study validates the factor structure of AHS based on exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). The next study examines the reliability and validity of the scale. Finally, the last study assesses the test-retest reliability of the scale.

Table 1 Demographic description

Variables	Category	Study 1		Study 2	Study 3
		Ss1	Ss2		
Gender	**Female**	163	162	269	112
	Male	208	208	231	140
Age	**Below 20**	162	152	113	65
	21–25	117	127	211	66
	26–30	41	71	91	72
	Above 31	51	20	85	49
Course	**UG**	166	227	191	125
	PG	173	123	246	83
	PhD	32	20	63	44
Steam	**Science**	185	100	175	84
	Management	120	111	146	118
	Humanities	66	159	179	50

Source: Primary Data, Ss1 – Subsample 1, Ss2 – Subsample 2

Study 1: factorial validity

Study 1 examined the factorial structure of the AHS in the Indian context with the help of exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) conducted on two independent samples.

Data collection and participants

Data was collected from 741 students enrolled in Delhi Technological University, India, using an online survey. A Google form was randomly circulated among students' email addresses available on the university's website. It was taken care that no student was already suffering from major mental health issues like depression, and no student had more than five back papers as this pressure might cause the development of major mental health issues like depression. There were 325 (43.9%) female and 416 (56.1%) male participants. Students below 20 years were 314 (42.4%), 21–25 years were 244 (32.9%), 26–30 years were 112 (15.1%), and above 31 years were 71 (9.6%). Students from undergraduate (UG) courses numbered 393 (53.0%), postgraduate (PG) courses numbered 296 (39.9%), and doctor of philosophy (PhD) courses numbered 52 (7.0%). Among these 741 students, 285 (38.5%) belonged to science programs, 231 (31.2%) to management programs, and 225 (30.4%) to humanities programs. The sample size was deemed appropriate as Cochran (1977) has suggested a sample size of 384 at $p < 0.05$. Table 1 presents the demographic information of the participants.

Instrument

A structured survey form divided into three sections was used to collect data. The first section highlighted data privacy and ethical statements. Participants were promised anonymity and strict academic use of data. In addition, they were told that the data would not be shared with any third party. This section also collected informed consent from the participants. The second section collected demographic data like age, gender, education level, etcetera. The last section measured participants' hope using Snyder et al.'s (1991) AHS. The form was designed to end the survey at the end

of Sect. 1 if the participant clicks 'NO'/does not wish to participate. The survey form was designed in the English language. Also, the original scale was developed in the English language by Snyder et al. (1991). Further, English is promoted as one of the essential and official languages in India, especially by the Indian education system. It is taught from the initial academic years of the students and is emphasized as a language for educational and professional purposes. Hence Indian students develop a strong grasp of the language till they reach college to pursue their UG and PG degrees (Garg, 2023).

Scoring and interpretation of the scale

The scale is an eight-item two-dimensional scale, where one dimension is Pathway and the other is Agency. This is a positively rated scale, meaning the score for both dimensions and the overall scale are calculated based on an eight-point Likert scale ranging between 1 (definitely false) and 8 (definitely true). The scores are calculated by averaging the scores of items 2, 9, 10 and 12 for agency and 1, 4, 6 and 4 for pathways. The total score of the overall scale is calculated by averaging the scores of four statements from each dimension, i.e., by adding the scores of each of the four statements of agency and pathways. Further, lower average scores, i.e., less than 4, indicate lower levels of agency, pathways and overall AHS, and average scores above 5 indicate higher levels of hope in an individual (Snyder et al., 1991).

Analysis and results

The sample of 741 students was divided into two subsamples using the SPSS split sample function. Odd serial-numbered responses constituted subsample one (S1 with $n_1 = 371$) on which EFA was applied. Even serial-numbered responses formed subsample two (S2 with $n_2 = 370$) on which CFA was applied to validate the model fit of the factor structure resulting from EFA. The EFA was applied on the eight items forming agency and pathways subscales of the THS, and four filler items were neglected as directed by Snyder et al. (1991). Before EFA, Kaiser-Meyer-Olkin (KMO), and Bartlett's test of sphericity were carried out to examine the suitability of the collected data for factor analysis. The value of the KMO (0.837) was above the minimum required value of 0.70. Additionally, Bartlett's Test of sphericity proved to be significant at $p < 0.05$ level of significance. The results confirmed the appropriateness of data for conducting EFA. Consequently, EFA was applied using the principal axis factoring with *promax rotation* (oblique rotation, kappa =4) and Kaiser normalization. It resulted in the same two factors as the original AHS with nomenclature as pathways and agency. They could explain that 60.12% of the total variance resulted from EFA. Both factors comprised four items each. Table 2 reports factor loadings, commonalities (h^2), and total variance of eight items of the AHS in the Indian context.

CFA was conducted on subsample two of 370 college students to validate the two-factor solution derived from EFA. Bentler and Bonett (1980) suggest values of RMSEA < 0.08, SRMR ≤0.05, GFI ≥0.95, CFI ≥0.97, and TLI >0.90 as model fit indices. The two-factor THS reported satisfactory values of RMSEA =0.055, SRMR =0.029, GFI =0.963, CFI =0.982, and TLI =0.943. Additionally, the value of $\chi^2 = 53.906$, $df = 19$ at $p < 0.05$ ($p = 0.02$). Next, we conducted a second-order CFA. It, too, gave a satisfactory model fit based on the values of RMSEA (0.051), SRMR (0.027), GFI (0.965), CFI (0.986), TLI (0.947), $\chi^2 = 53.871$, $df = 18$ at $p < 0.05$ ($p = 0.01$). We compared both models based on χ^2 difference test. We found that the $\chi^2_{difference}$ was 0.035 for $df_{difference} = 1$, which was less than the recommended p-value at *significance level 0.05* and $df_{difference} = 1$, i.e. the obtained value was less than 1.84. Therefore, it was safe to assume that the model reported in the second-order CFA was better than the model reported in the first-order CFA

Table 2 Factor structure

No.	Statements	Loadings	h^2	Eigenvalues
Agency				
1	I energetically pursue my goals.	0.587*	0.345	1.701
2	My past experiences have prepared me well for my future.	0.642*	0.412	
3	I've been pretty successful in life.	0.712*	0.507	
4	I meet the goals that I set for myself.	0.661*	0.437	
Pathways				
1	I can think of many ways to get out of a jam.	0.897*	0.805	2.655
2	There are lots of ways around any problem.	0.716*	0.513	
3	I can think of many ways to get the things in life that are important to me.	0.781*	0.610	
4	Even when others get discouraged, I know I can find a way to solve the problem.	0.853*	0.728	
	Number of items	8		
	Eigenvalues	4.356		
	Explained Variance	60.12%		

Source: Primary Data. *Satisfactory Factor Loadings (i.e., greater than |0.5|), h^2: Communalities

and that AHS is a multidimensional model. Thus, the results of CFA suggested satisfactory factorial validity of the two-factored eight-item AHS among Indian college students. These results follow the suggestions of the study conducted by Snyder et al. (1991). After factorial validity, the scale is subjected to further examination of the reliability and validity in Study 2. Figures 1 and 2 present the first-order CFA and second-order CFA models of AHS in the Indian context.

Study 2: reliability and validity testing

The study 2 examined the reliability of the two-factored AHS in the Indian context using Cronbach's Alpha (CA) and composite reliability (CR), whose acceptable values are 0.70 and above (Cronbach & Shavelson, 2004). It evaluated the convergent validated using average variance explained (AVE) values, which should be more than 0.50 to establish convergent validity. Discriminant validity was ensured with the help of Fornell and Larcker's (1981) criterion, which states that the square root of AVE values should be less than correlation values to establish the discriminant validity of a scale. Construct validity was adjudged through correlation analysis of AHS with gratitude and psychological distress. Finally, predictive validity was examined through correlation and regression analysis of AHS with spiritual well-being.

Data collection and participants

An online survey form was randomly circulated on email IDs of students of DTU; the only difference was that this time, we did not share the questionnaire with the students who were already approached during Study 1. Initially, the survey form was shared with 721 students, of whom 500 showed interest in the study and participated in the survey. Again, the exclusion criteria so formed in Study 1 were used, and the sample size was appropriate based on the sample size criteria suggested by Cochran (1977). There were 269 (53.8%) female participants and 231 (46.2%) male participants. Students from UG courses were 191 (38.2%), PG courses 246 (49.2%), and PhD courses 63 (12.6%). Of the surveyed participants, 175 (35.0%) were from science programs, 146 (29.2%) were from management programs, and 179 (35.8) were from humanities programs. There were 113 (42.2%) students below 20 years, 211 (42.2%) aged between 21 and 25 years, 91 (18.2%) aged between 26 and 30 years, and 85 (17.0%) were above 31 years. Table 1 presents the demographic information of the sample.

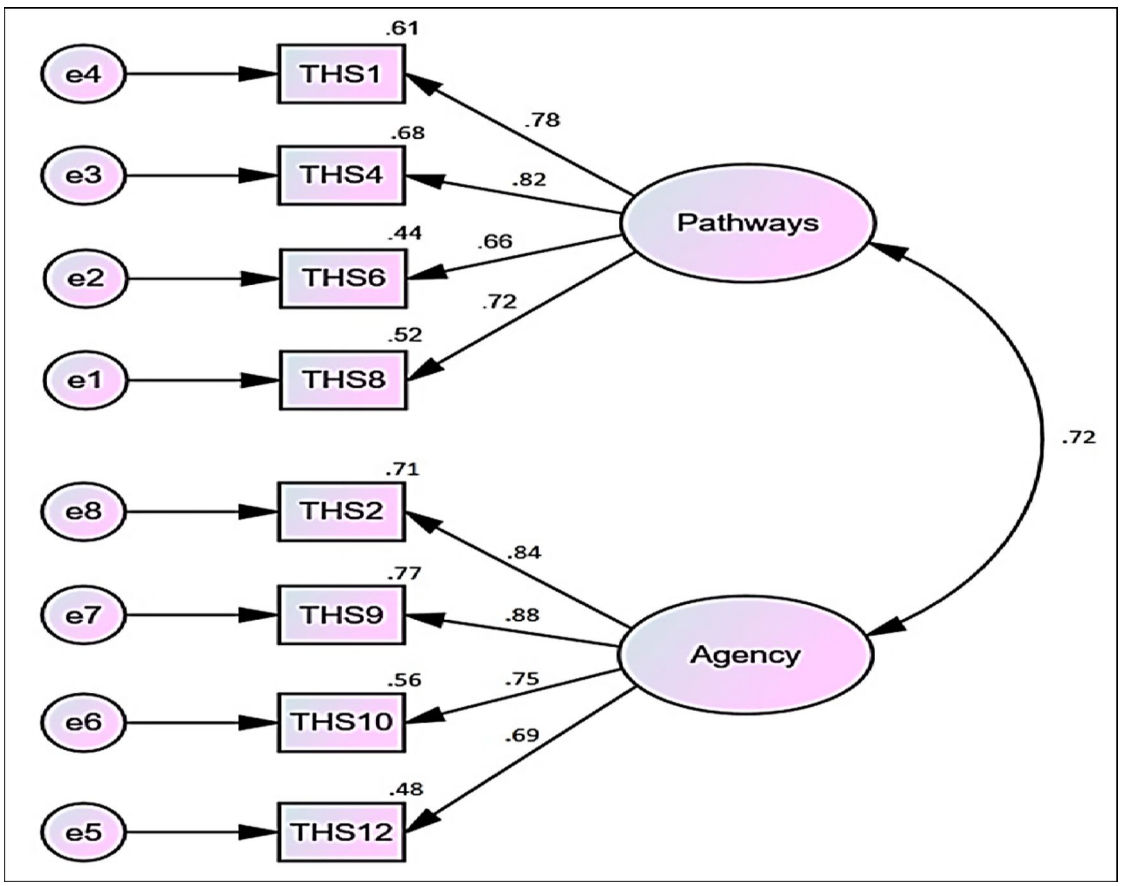


Fig. 1 Results of CFA model. Source: Primary Data

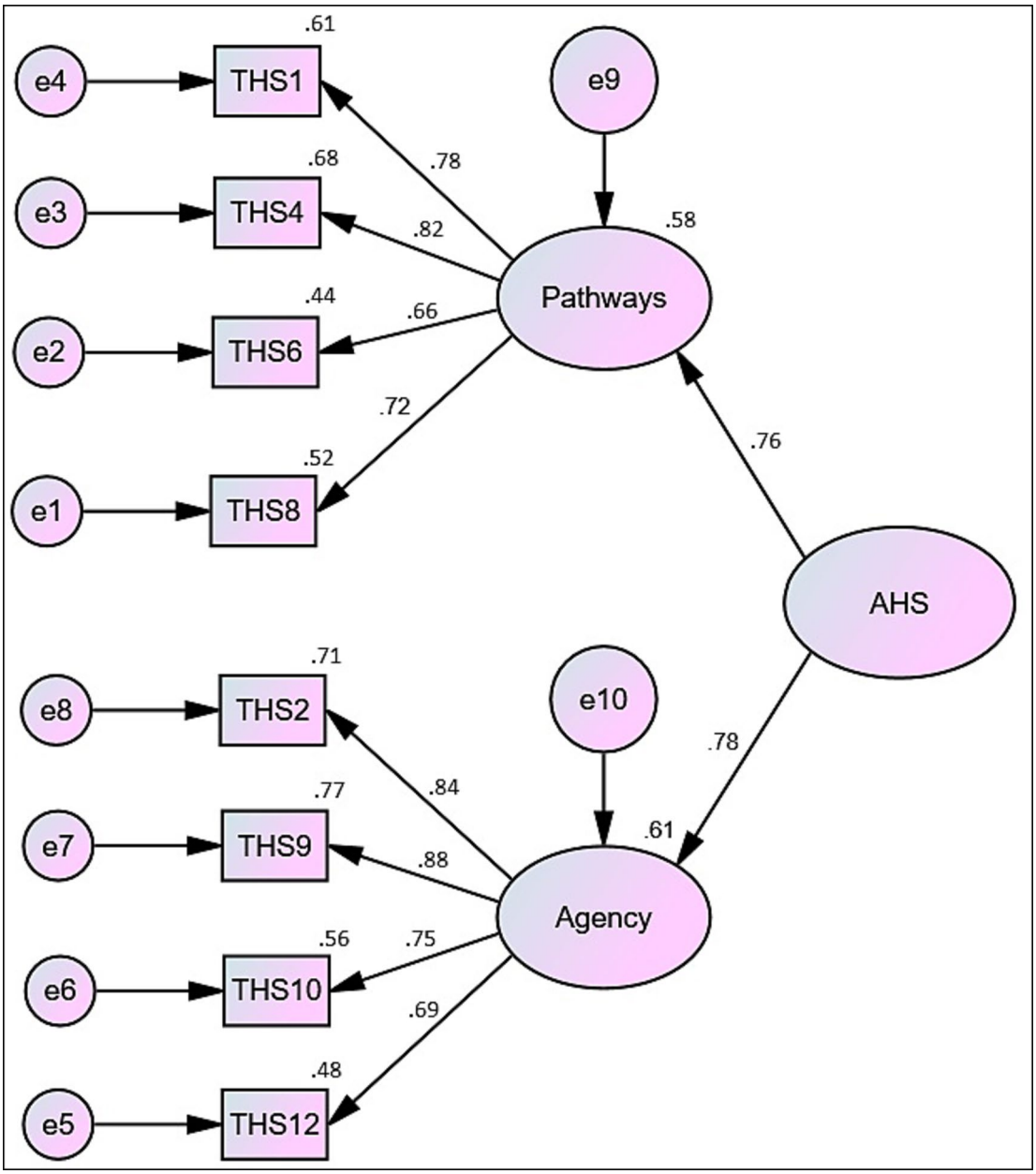


Fig. 2 Results of second order CFA model. Source: Primary Data

Instrument

An online structured questionnaire was used to collect the data. It was divided into six sections. The first section highlighted data privacy and ethical statements. Participants were promised anonymity and strict academic use of data. In addition, they were told that the data would not be shared with any third party. This section also collected informed consent from the participants. Section 2 gathered demographic information. Section 3 to Sect. 6 captured hope, spiritual well-being, gratitude, and distress, using respective instruments.

Hope

Since EFA and CFA did not suggest any deviation in the original AHS in the Indian context, the original AHS was used to measure college students' hope level on an eight-point rating scale ranging between 1 (definitely false) to 8 (definitely true).

Spiritual well-being

SWBS, established by Paloutzian and Ellison (1982), is a twenty-item six-point rating scale ranging between 1

(strongly disagree) to 6 (strongly agree). It examines an individual's relationship with the almighty and sense of life purpose and satisfaction. A higher score presents higher levels of relationship with God, a heightened sense of purpose, and satisfaction with life (Paloutzian & Ellison, 1982). A few items are "I do not find much satisfaction in private prayer with God" and "I feel that life is a positive experience."

Gratitude

GQ6, developed by McCullough et al. (2002), is a six-item seven-point rating scale ranging between 1 (strongly disagree) to 7 (strongly agree). It is used to examine an individual's experience of gratitude (Byrne, 2011). A few of its statements are "I have so much in life to be thankful for" and "I am grateful to a wide variety of people."

Psychological distress

Kessler Psychological Distress Scale (K10), developed by Kessler et al. (2003), is a ten-item five-point rating scale ranging between 1 (none of the time) to 5 (all of the time). It is used to measure one's psychological distress level. A few of its statements are "In the past four weeks, about how often did you feel hopeless?" and "In the past four weeks, about how often did you feel worthless?"

Analysis and results

Table 3 suggested that the CA values for overall AHS, pathways, and agency were 0.885, 0.879, and 0.862, respectively. Also, CR estimates for overall AHS, pathways, and agency were 0.880, 0.831, and 0.858, respectively. These values of CA and CR are greater than the threshold value of 0.7, thereby recommending AHS's reliability in the Indian context. The values of AVE for overall AHS, pathways, and agency were 0.57, 0.59, and 0.56, respectively, and were greater than the threshold value of 0.5. These results confirmed the convergent validity of AHS among Indian adults. Further, the table also reported that the square root AVE values were greater than the correlation between pathways and agency, confirming AHS's discriminant validity in the Indian sample.

The external convergent and discriminant validity were examined through the correlation coefficient between AHS, a similar (GQ- 6), and a dissimilar (K10) scale. Previous researchers have suggested that people with higher hope scores have higher levels of gratitude scores (Cheng et al., 2023). Also, people with higher hope scores have broadened their thought-action repertoire (Fredrickson, 2004) and, therefore, have lower psychological distress. Thus, it was hypothesized that AHS positively correlates with GQ6 and negatively correlates with K10. Table 3 recommended significant positive correlations between GQ6 and AHS ($r = 0.445$, $p < 0.05$), agency ($r = 0.419$, $p < 0.05$), and pathways ($r = 0.422$, $p < 0.05$). Also, the results suggested significant negative correlations between K10 and AHS ($r = -0.431$, $p < 0.05$), agency ($r = -0.462$, $p < 0.05$), and pathways ($r = -0.528$, $p < 0.05$). These results construct (convergent and divergent) validity of AHS in an Indian sample.

The broaden-and-build theory of positive emotions states that positive emotions like hope builds and broadens psychological and social resources for a person, which is instrumental to one's emotional and spiritual well-being (Fredrickson, 2004). Accordingly, it was hypothesized that hope significantly predicts the spiritual well-being of Indian adults. The acceptance of such a hypothesis would confirm the criteria validity of the scale. Table 3 reported significant

Table 3 Descriptive statistics, reliability, and validity measures and Correlation Matrix

Descriptive statistics, reliability, and validity measures								
Variable	Mean	S.D	CA	CR	AVE	1		2
Agency	7.128	0.815	0.879	0.858	0.59	1 **(0.768)**		
Pathways	6.992	0.753	0.831	0.831	0.56	0.598†		1 **(0.56)**
Correlation Matrix								
Variables	AHS	Agency	Pathways	SWBS	RWB	EWB	GQ6	K10
AHS	1							
Agency	0.652*	1						
Pathways	0.598*	0.525*	1					
SWBS	0.657*	0.601*	0.623*	1				
RWB	0.512*	0.481*	0.490*	0.692*	1			
EWB	0.498*	0.438*	0.461*	0.703*	0.688*	1		
GQ6	0.445*	0.419*	0.422*	0.508*	0.455*	0.449*	1	
K10	− 0.431*	− 0.462*	− 0.528*	− 0.493*	− 0.611*	− 0.571*	− 0.328*	1

Source: Primary Data, S.D: Standard Deviation, CA: Cronbach's Alpha, CR: Composite Reliability, Bold values indicate the square root of AVE (diagonally), † indicates the correlation between agency and pathways, *Sig. at 0.05, AHS– Adult Hope Scale, SWBS– Spiritual Well-Being Scale, RWR– Religious Well-Being, EWB– Existential Well-Being, GQ6– Gratitude Questionnaire 6, K10– Kessler Psychological Distress Scale

Table 4 Regression analysis

Model	Variable	Unstandardized Coefficients		Standardized Coefficient Beta	*t-value*	Sig.
		B	Std. error			
Model 1 (DV– Spiritual Well-Being)						
1	***Constant***	0.325	0.101		5.758	0.028*
	Pathways	0.651	0.132	0.601	2.981	0.031*
	Agency	0.511	0.210	0.498	4.021	0.012*
	R	***R Square***	***Adjusted R Square***	**Std. error**	***F***	**Sig**
	0.728	0.529	0.408	0.112	7.183	0.014*

Source: Primary Data, *Sig. at. 0.05

positive correlations between SWBS and AHS ($r = 0.657$, $p < 0.05$), agency ($r = 0.601$, $p < 0.05$), and pathways ($r = 0.623$, $p < 0.05$). SWBS comprises two dimensions of spiritual well-being: religious and essential well-being. Table 3 also reported significant correlations between religious well-being and AHS ($r = 0.512$, $p < 0.05$), agency ($r = 0.481$, $p < 0.05$), and pathways ($r = 0.490$, $p < 0.05$). Also, there are positive associations between existential well-being and AHS ($r = 0.498$, $p < 0.05$), agency ($r = 0.438$, $p < 0.05$), and pathways ($r = 0.461$, $p < 0.05$). Table 4 reported that agency ($\beta = 0.498$, $p < 0.05$) and pathways ($\beta = 0.651$, $p < 0.05$) are significant predictors of spiritual well-being. These two variables of hop could explain 52.9% variations in adults' hope levels. Therefore, it recommended predictive validity of the scale in the Indian context.

Study 3: Test-retest reliability

In this study, the test-retest reliability of the AHS was explored by assessing the correlation coefficient between hope scores calculated by collecting the data twice.

Data collection and participants

An online survey form was shared with 400 randomly selected students from DTU, out of which 324 responded (at time T1). Again, we did not share the questionnaire with the students who were already approached during studies 1 and 2. After a gap of one month (at time T2), these respondents were again approached. The sequences of questions were changed to avoid familiarity bias. At time T2, 252 students answered the surveyed questions. There were 112 male (44.4%) and 140 (55.6%) participants. Participants aged below 20 were 65 (25.8%), between 21 and 25 years were 66 (26.2%), between 26 and 30 years were 72 (28.6%), and above 31 were 49 (19.4%). There were 125 (49.6%) undergraduate, 83 (32.9%) postgraduate, and 44 (17.5%) PhD students. Respondents from science courses were 84 (33.3%), management courses 118 (46.8%), and humanities courses 50 (19.8%). Again, the same participant selection criteria were followed, as the one followed in the study 1. The sample description is presented in Table 1.

Instrument

A structured questionnaire was shared with the students of DTU. It was divided into three sections. The first section provided data privacy, ethical, and informed consent statements as a standard practice. Next, the section captured the demographic information. Finally, the last section recorded the responses based on the AHS scale.

Analysis and results

Test-retest reliability was examined to assess AHS's stability and reliability over time. The correlations between scores of AHS and its two dimensions obtained in two different periods (time T1 and T2) were explored. Results recommended positive correlations in two scores of pathways ($r = 0.621$, $p < 0.05$), agency ($r = 0.695$, $p < 0.05$), and overall AHS ($r = 0.658$, $p < 0.05$), which confirmed the test-retest reliability of AHS in the Indian context.

Discussion

Hope is an essential trait that helps humans formulate goals and take action to achieve those goals (Snyder et al., 1991). It motivates people in their low times by letting them wish for a positive future (Doe, 2020). Hope is vital for physical, mental, emotional, and spiritual well-being (Doe, 2020). Researchers suggest hope provides comfort and decreases psychological distress (Doe, 2020). The presence of hope can diminish feelings of helplessness and give people a sense of control (Snyder, 2002). The Indian philosophy provides contexts and reasons, stating that hope connects people with the Almighty ("Hope for a Good Life," 2018).

Additionally, in the present world, when regional events have global impacts and life has become dull, positive emotions like hope and gratitude provide some relief to people and cushion the negative impact of unwishful incidences (Doe, 2020). Hence, the second wave of positive

psychology (PP 2.0) provides pieces of evidence to study and understand positive emotions. According to Fredrickson's (2004) broaden-and-built theory, positive emotions like hope and gratitude broaden a person's awareness and build psychological resources like spiritual quotient and resilience that cushion an individual against psychological distress and other negativities (Garg, 2023). Based on PP 2.0, Synder et al.'s (1991) work on dispositional hope has received worldwide acceptance. The authors provide a bi-dimensional model and scale for measuring dispositional hope. The first dimension is agency, which is the motivation to formulate goals, and the second is pathways which are an individual's motivation to take action to achieve planned goals. Also, Synder et al. (1991) have suggested validating the scale in cross-cultural contexts. Therefore, based on the recommendation of Snyder et al. (1991 and 2022), the current study's authors conducted a psychometric analysis of the AHS in the Indian milieu. The analysis is based on three independent studies.

Study 1 explored the factorial structure of AHS. EFA and CFA were used to analyze the factor model of AHS in India. The results of Study 1 confirmed the two-factor structure of AHS. For study 1, the sample was divided into two subsamples (SS1 and SS2). EFA was applied in SS1, and CFA was applied in SS2. Prior to applying EFA, the SS1 was subjected to the KMO test of sampling adequacy and Bartlett's test of sphericity. The value of the KMO test and significant Bartlett's test of sphericity confirmed the suitability of the sample to apply EFA. No statement was dropped during EFA, and it reported a two-factor model. This eight-item scale was further subjected to first-order and second-order CFAs to validate the factorial structure of the scale. The model fit indicators confirmed the factorial validity of the scale in both models. Further, the second-order CFA confirm that the AHS is a multidimensional scale. The model derived from this study is in line with the actual model presented by Snyder et al. (1991).

The second study assessed the reliability and validity of the AHS in India. The values of CA and CR suggested the internal consistency reliability of the AHS. Additionally, the AVE values of AHS and its two dimensions were greater than the threshold value (0.05), and the squared correlation of the factors was less than their AVE values, thus confirming the convergent validity of the scale. The criterion validity of the scale was explored by analyzing the correlation coefficients between AHS, GQ6, and K10. Significant positive correlations ($p < 0.05$) between AHS and GQ6 were reported, whereas a significant negative correlation was reported between AHS and K10, thus confirming the criterion validity of the scale. The two dimensions of the AHS, agency and pathway, contribute to the relationship, i.e., they indicate that both the ability to find routes to goals and the motivation to pursue them are associated with higher levels of gratitude. Hence, these results suggest that hope and gratitude are positively related, indicating that people with higher levels of hope also tend to experience higher levels of gratitude. Also, the Broaden-and-Built Theory suggest that positive emotions like hope broaden an individual's thought-action repository that leads to the development of psychological and social resources (Fredrickson, 2004). This broadening effect help people to manage and reduce psychological distress and provide a brighter outlook to life. On this ground, the results from our study suggest that pathways provide cognitive flexibility to the individuals, enabling them to navigate obstacles more effectively. Similarly, agency thinking empowers people to initiate and sustain actions towards their goals. Therefore, people with higher pathways and agency scores tend to reduce psychological distress. Further, to test the predictive validity of AHS, a regression analysis was conducted between AHS and SWBS, which confirmed the predictive validity of the scale. The results of regression analysis suggest that agency and pathways are strong predictors of spiritual well-being, suggesting that people with higher scores of agency and pathways can have higher sense of spiritual well-being. In other words, people with high agency believe in their abilities to achieve their goals, which gives them a higher sense of control which adds to their spiritual well-being as these types of people are likely to sense that they have a purpose and direction in their lives. Similarly, higher pathways scores suggest that people have the ability to generate multiple solutions to a problem, which can enhance their spiritual experience by fostering a proactive approach to life and its problems. Further, the results of this study align with those conducted by Cheng et al. (2023) and Witvliet et al. (2018). These authors have suggested that people with higher hope levels are more likely to develop gratuitous behavior and spiritual well-being than those with lower hope levels. Moreover, the results of our study align with other reliability and validity studies on the AHS, like those conducted by Kraft et al. (2021) in Pakistan, Vinueza-Solórzano et al. (2023) in Ecuador, and Nooripour et al. (2022) in Iran. These studies also reported a two-factor solution of the scale, the positive relationship of AHS with work engagement and AHS's negative relationships with burnout, mental distance, depression, stress, and anxiety (Nooripour et al., 2022; Vinueza-Solórzano et al., 2023).

The final study was conducted to explore the test-retest reliability of AHS. For this, the scale was administered to randomly select Indian college students. After one month, the scale was again administered to the students with a changed orientation/flow of the items. A significant correlation between the two scores confirmed the test-retest reliability of the scale. The results of the test-retest align with

the results of a similar study conducted in Pakistan by Kraft et al. (2021). Similar to the results of our study, Kraft et al. (2021) also reported a positive correlation between the results of test-retest reliability. From this standpoint, this study validates the AHS in the Indian context.

Implementations

The findings of the present study offer significant theoretical and practical implications. First, there is a dearth of a reliable and valid scale to measure dispositional hope in Indian college students. Validation of AHS in India can be considered a step toward addressing this gap. Secondly, the findings regarding dispositional hope are likely to motivate academicians, researchers, and practitioners to broaden their understanding of hope. While research on hope has primarily focused on its state and domain-specific nature, there is room in the literature to explore a more generalized nature of hope in psychological, philosophical, and other contexts ("Hope for a Good Life," 2018). Thirdly, researchers have suggested that there exists a gap in the use of hope in Indian psychology, and at the same time, philosophical thinkers around the globe also suggest that understanding hope merely limiting to a theoretical concept is not possible; instead, a generalized form of hope has to be talked about (Krafft & Choubisa, 2018). Hence, the results generated after utilizing the robustly validated AHS in India will contribute to the Indian psychology literature and help philosophical thinkers better explain their ideas on hope. Fourth, the analysis of the psychometric properties of AHS adds to the literature on positive psychology and the broad-and-built theory by Fredrickson (2004). Fourthly, validating the scale in the cross-cultural context allows researchers, academicians, and policymakers to comprehensively understand hope in different cultures. Also, the cross-cultural validation of the AHS will help the research community to certify that the scale is reliable and valid across the globe by providing evidence from India. Even the developers of AHS recommend a cross-cultural validation of the scale (Snyder et al., 1991). Lastly, researchers, academicians, and managers can use the scale to examine the interventions' efficacy to increase hope in people.

Limitations, the scope for future studies, and the conclusion

While the study presents insightful findings, it is essential to consider the study's limitations when understanding these findings. Firstly, the study is based on a sample of university students only, which might only present the perspective of hope from their side. In the future, researchers could gather data from a diverse sample of people of all ages. Secondly, apart from the test-retest reliability analysis (which makes the final study a longitudinal study), other studies are based on a cross-sectional survey, which presents its own inherent limitations. To overcome these limitations, Garg (2023) suggests that future researchers may use a mixed approach of qualitative and quantitative surveys like interviews and observations. Thirdly, the reliability testing was limited to internal consistency and test-retest reliability. Fourthly, the original scale was developed in English language by Snyder et al. (1991). Although we have discussed our reasons as to why we did not back translate the scale in Hindi language, future researchers are recommended to back translate the scale in Hindi, or any other language as deemed necessary. Future researchers may test inter-rater or alternate-form reliabilities for better understanding. Lastly, in accordance with the suggestions of previous psychometric property analysis studies, the scale needs to be validated in a cross-cultural setting.

To summarize, hope is one of the most powerful emotions and feelings in human life that can motivate people to formulate goals and act upon them. Therefore, to analyze the level of hope in people, a reliable and valid scale that measures hope is required. Hence, the present study tested the psychometric properties of AHS in India. The results from the factorial structure validity study confirmed the two factored solutions of the scale. The study on reliability and validity confirmed to scale to be reliable and valid in the Indian context.

Funding Not applicable.

Data availability As promised to participants, data will be kept confidential and will not be shared.

Declarations

We hereby declare that the present article is compiled with Ethical Standards as below.*

Conflict of interest Authors declare No conflict of interest.

Ethical approval All methods in studies involving human subjects were carried out in line with the institutional and national research committee's ethical standards, as well as the 1964 Helsinki Declaration and its subsequent amendments or similar ethical standards. Informed consent of participants was taken.

References

Aghababaei, N., Sohrabi, F., Eskandari, H., Borjali, A., Farrokhi, N., & Chen, Z. (2016). Predicting subjective well-being by religious and scientific attitudes with hope, purpose in life, and death anxiety as mediators. *Personality and Individual Differences*, *90*, 93–98. https://doi.org/10.1016/j.paid.2015.10.046

American Educational Research Association, American Psychological Association, National Council on Measurement in Education,

Testing, J. C. O. S. F. E. A. P., American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing (U.S.). (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*(3), 588–606. https://doi.org/10.1037/0033-2909.88.3.588

Bowers, R. M., & Bowers, E. P. (2023). A literature review on the role of hope in promoting positive youth development across Non-WEIRD contexts. *Children (Basel)*, *10*(2), 346. https://doi.org/10.3390/children10020346

Byrne, B. M. (2011). *Structural equation modeling with EQS*. In Routledge eBooks. https://doi.org/10.4324/9780203726532

Cheavens, J. S., Heiy, J. E., Feldman, D., Benitez, C., & Rand, K. L. (2019). Hope, goals, and pathways: Further validating the hope scale with observer ratings. *The Journal of Positive Psychology*, *14*(4), 452–462. https://doi.org/10.1080/17439760.2018.1484937

Cheng, Q., Lu, W., Duan, Y., Li, J., Xie, J., & Chen, Y. (2023). Spiritual well-being and its association with hope and meaning in life among gastrointestinal cancer patients: A cross-sectional study. *Supportive Care in Cancer*, *31*(4). https://doi.org/10.1007/s00520-023-07696-5

Cochran, W. G. (1977). *Sampling techniques*. Wiley.

Colla, R., Williams, P. L., Oades, L. G., & Camacho-Morles, J. (2022). A new hope for positive psychology: A dynamic systems reconceptualization of hope theory. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.809053

Cronbach, L. J., & Shavelson, R. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement*, *64*(3), 391–418. https://doi.org/10.1177/0013164404266386

Deb, S., Dorji, N., Kar, S., Sunny, A. M., Deb, S., Ghosh, S., & Chakraborty, S. (2023). COVID-19 and stress of Indian youth: An association with background, on-line mode of teaching, resilience and hope. *Journal of Affective Disorders Reports*, *12*, 100502. https://doi.org/10.1016/j.jadr.2023.100502

Doe, M. J. (2020). Conceptual foreknowings: An integrative review of hope. *Nursing Science Quarterly*, *33*(1), 55–64. https://doi.org/10.1177/0894318419881805

Espedal, G. (2021). Hope to see the soul: The relationship between spirituality and hope. *Journal of Religion & Health*, *60*(4), 2770–2783. https://doi.org/10.1007/s10943-021-01245-2

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50. https://doi.org/10.1177/002224378101800104

Fredrickson, B. L. (2004). The broaden–and–build theory of positive emotions. *Philosophical Transactions of the Royal Society B*, *359*(1449), 1367–1377. https://doi.org/10.1098/rstb.2004.1512

Fromm, E. (1968). *The revolution of hope, toward a humanized technology*. https://ci.nii.ac.jp/ncid/BA04814834

Gaiha, S. M., Salisbury, T. T., Koschorke, M., Raman, U., & Petticrew, M. (2020). Stigma associated with mental health problems among young people in India: A systematic review of magnitude, manifestations and recommendations. *BMC Psychiatry*, *20*(1). https://doi.org/10.1186/s12888-020-02937-x

Garg, N. (2023). Development and validation of Hindu gratitude scale (HGS-15): A Rnas perspective. *Journal of Religion and Health*. https://doi.org/10.1007/s10943-023-01831-6

Greenaway, K. H., Cichocka, A., Van Veelen, R., Likki, T., & Branscombe, N. R. (2016). Feeling hopeful inspires support for social change. *Political Psychology*, *37*(1), 89–107. https://doi.org/10.1111/pops.12225

Herth, K. (1992). Abbreviated instrument to measure hope: Development and psychometric evaluation. *Journal of Advanced Nursing*, *17*(10), 1251–1259. https://doi.org/10.1111/j.1365-2648.1992.tb01843.x

Hutz, C. S., Midgett, A., Pacico, J. C., Bastianello, M. R., & Zanon, C. (2014). The relationship of hope, optimism, self-esteem, subjective well-being, and personality in Brazilians and Americans. *Psychology*, *05*(06), 514–522. https://doi.org/10.4236/psych.2014.56061

Kapur, N. M. S., & Teotia, N. D. A. (2018). Vital young adults: Hopeful and gritty. *International Journal of Indian Psychology*, *6*(4). https://doi.org/10.25215/0604.020

Kessler, R. C., Barker, P., Colpe, L. J., Epstein, J., Gfroerer, J. C., Hiripi, E., Howes, M., Normand, S. T., Manderscheid, R. W., Walters, E. E., & Zaslavsky, A. M. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry*, *60*(2), 184. https://doi.org/10.1001/archpsyc.60.2.184

Krafft, A. M., & Choubisa, R. (2018). Hope in the Indian psychology context: Philosophical foundations and empirical findings. *Social indicators research series*. Springer Nature (Netherlands. https://doi.org/10.1007/978-3-319-78470-0_7

Kraft, M., Piechocki, M., Ptak, B., & Walas, K. (2021). UAVVaste dataset [Dataset]. In *Zenodo (CERN European Organization for Nuclear Research)*. https://doi.org/10.5281/zenodo.8214061

Lazarus, R. S. (1999). Hope: An emotion and a vital coping resource against despair. *Social Research: An International Quarterly*, *66*. https://philpapers.org/rec/LAZHAE

McCullough, M. E., Emmons, R. A., & Tsang, J. (2002). The grateful disposition: A conceptual and empirical topography. *Journal of Personality and Social Psychology*, *82*(1), 112–127. https://doi.org/10.1037/0022-3514.82.1.112

Namastu, S. (2018). *The riddle of fate and free-will solved*: https://www.advaita-vedanta.org/articles/The_Riddle_of_Fate_and_Free.htm

Nooripour, R., Farmani, F., Brand, S., Ghanbari, N., Ilanloo, H., Matacotta, J. J., & Sadeghi-Bahmani, D. (2022). Psychometric characteristics of Persian version of adult hope scale (AHS) in Iranian females with multiple sclerosis (MS). *Journal of Kermanshah University of Medical Sciences*, *26*(2). https://doi.org/10.5812/jkums-123276

Paloutzian, R. F., & Ellison, C. W. (1982). Loneliness, spiritual well-being, and quality of life. In L. A. Peplau, & D. Perlman (Eds.), *Loneliness: A sourcebook of current theory, research and therapy* (pp. 224–237). Wiley.

Pettit, P. (2004). Hope and its place in mind. *Annals of the American Academy of Political and Social Science*, *592*(1), 152–165. https://doi.org/10.1177/0002716203261798

Redlich-Amirav, D., Ansell, L. J., Harrison, M., Norrena, K. L., & Armijo-Olivo, S. (2018). Psychometric properties of Hope Scales: A systematic review. *International Journal of Clinical Practice, 72*(7), e13213. https://doi.org/10.1111/ijcp.13213

Schrank, B., Woppmann, A., Sibitz, I., & Lauber, C. (2011). Development and validation of an integrative scale to assess hope. *Health Expectations*, *14*(4), 417–428. https://doi.org/10.1111/j.1369-7625.2010.00645.x

Scioli, A., Ricci, M. A., Nyugen, T., & Scioli, E. R. (2011). Hope: Its nature and measurement. *Psychology of Religion and Spirituality*, *3*(2), 78–97. https://doi.org/10.1037/a0020903

Snyder, C. R. (2002). TARGET ARTICLE: Hope theory: Rainbows in the mind. *Psychological Inquiry*, *13*(4), 249–275. https://doi.org/10.1207/s15327965pli1304_01

Snyder, C., Harris, C. A., Anderson, J. R., Holleran, S. A., Irving, L. M., Sigmon, S. T., Yoshinobu, L., Gibb, J., Langelle, C., & Harney, P. (1991). The will and the ways: Development and validation of an individual-differences measure of hope. *Journal of*

Personality and Social Psychology, *60*(4), 570–585. https://doi.org/10.1037/0022-3514.60.4.570

Susarla, S. M., Fernandes, R., & Parmar, S. (2021). Gratitude and hope in challenging times. *Craniomaxillofacial Trauma & Reconstruction*, *14*(3), 173. https://doi.org/10.1177/19433875211015634

Vinueza-Solórzano, A. M., Campoverde, R. E., Portalanza-Chavarría, C. A., De Freitas, C. P. P., Hutz, C. S., & Vazquez, A. C. S. (2023). Adaptation and validation of the adult dispositional hope scale in the Ecuadorian context. *Psicologia*, *36*(1). https://doi.org/10.1186/s41155-023-00246-2

Vivekananda, S. (2017). *The complete works of Swami Vivekananda, volume 1: Great master series*. Discovery Publisher.

Warber, S. L., Ingerman, S., Moura, V. L., Wunder, J., Northrop, A., Gillespie, B. W., Durda, K., Smith, K., Rhodes, K. S., & Rubenfire, M. (2011). Healing the heart: A randomized pilot study of a spiritual retreat for depression in acute coronary syndrome patients. *Explore-the Journal of Science and Healing*, *7*(4), 222–233. https://doi.org/10.1016/j.explore.2011.04.002

Waterworth, J. M. (2004). A philosophical analysis of hope. In *Palgrave Macmillan UK eBooks*. https://doi.org/10.1057/9780230506022

Witvliet, C. V., Richie, F. J., Luna, L. M. R., & Van Tongeren, D. R. (2018). Gratitude predicts hope and happiness: A two-study assessment of traits and States. *The Journal of Positive Psychology*, *14*(3), 271–282. https://doi.org/10.1080/17439760.2018.1424924

Yaghoobzadeh, A., Soleimani, M., Allen, K., Chan, Y. H., & Herth, K. (2018). Relationship between spiritual well-being and hope in patients with cardiovascular disease. *Journal of Religion & Health*, *57*(3), 938–950. https://doi.org/10.1007/s10943-017-0467-0

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

I. J. Computer Network and Information Security, **2025, 1, 34-50**
Published Online on April 8, 2025 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijcnis.2025.02.03

GMM-based Imbalanced Fractional Whale Particle Filter for Multiple Object Tracking in Surveillance Videos

Avinash Ratre
Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, 110042, India
E-mail: avinashratre@dtu.ac.in
ORCID iD: https://orcid.org/0000-0001-9803-2665

Received: 19 December 2024; Revised: 11 February 2025; Accepted: 04 March 2025; Published: 08 April 2025

Abstract: The imbalanced surveillance video dataset consists of majority and minority classes as normal and anomalous instances in the nonlinear and non-Gaussian framework. The normal and anomalous instances cause majority and minority samples or particles associated with high and low probable regions when considering the standard particle filter. The minority particles tend to be at high risk of being suppressed by the majority particles, as the proposal probability density function (pdf) encourages the highly probable regions of the input data space to remain a biased distribution. The standard particle filter-based tracker afflicts with sample degeneration and sample impoverishment due to the biased proposal pdf ignoring the minority particles. The difficulty in designing the correct proposal pdf prevents particle filter-based tracking in the imbalanced video data. The existing methods do not discuss the imbalanced nature of particle filter-based tracking. To alleviate this problem and tracking challenges, this paper proposes a novel fractional whale particle filter (FWPF) that fuses the fractional calculus-based whale optimization algorithm (FWOA) and the standard particle filter under weighted sum rule fusion. Integrating the FWPF with an iterative Gaussian mixture model (GMM) with unbiased sample variance and sample mean allows the proposal pdf to be adaptive to the imbalanced video data. The adaptive proposal pdf leads the FWPF to a minimum variance unbiased estimator for effectively detecting and tracking multiple objects in the imbalanced video data. The fractional calculus up to the first four terms makes the FWOA a local and global search operator with inherent memory property. The fractional calculus in the FWOA oversamples minority particles to be diversified with multiple imputations to eliminate data distortion with low bias and low variance. The proposed FWPF presents a novel imbalance evaluation metric, tracking distance correlation for the imbalanced tracking over UCSD surveillance video data and shows greater efficacy in mitigating the effects of the imbalanced nature of video data compared to other existing methods. The proposed method also outshines the existing methods regarding precision and accuracy in tracking multiple objects. The consistent tracking distance correlation near zero values provides efficient imbalance reduction through bias-variance correction compared to the existing methods.

Index Terms: Multiple Object Tracking, Particle Filter, Fractional Whale Optimization, Gaussian Mixture Model, Imbalanced Surveillance Video Data.

1. Introduction

Visual tracking is a critical task in vision-based applications in which the actions of humans are identified, and the human-computer interfaces are developed. Stochastic point tracking methods employ a state-space approach to model the linear or non-linear dynamics of the objects using the Kalman filter or particle filter. The Kalman filter fails when vision tracking problems are non-linear, non-Gaussian, and multi-modal. The PF solves such vision problems, which rely on the recursive construction of the state space's posterior probability density function (pdf). A set of random particles is produced to estimate the posterior pdf of the states of the tracked objects. In estimating the states of objects, these sets of particles are updated and propagated based on sequentially arriving observations. The particle filter better estimates hidden state variables than the other non-linear filters [1, 2]. However, due to the resampling procedure, the particle filter suffers from the problem of sample degeneration and impoverishment propagating in state space [3]. The resampling technique involves replacing the old set of particles with related weights with the new set of particles with related weights in more probable regions of state space when sample degeneration increases above a threshold. The resulting sample impoverishment pertains to the frequent selection of particles with high weights and the dying out of other particles with negligible weights when particles propagate after each time step [4, 5].

This work is open access and licensed under the Creative Commons CC BY 4.0 License.

The imbalanced surveillance video dataset consists of normal and anomalous instances. The normal and anomalous instances cause majority and minority samples associated with high and low probable regions. In the standard particle filter, the minority particles tend to be at high risk of being suppressed by the majority particles, as the biased proposal or importance pdf encourages only majority particles in the highly probable regions of the input data space. However, resampling tries to correct the bias to a small extent. The difficulty in designing the correct proposal or importance pdf prevents standard particle filter-based tracking in the imbalanced surveillance video data. Further, the problem of sample degeneration and sample impoverishment occurs due to the choice of a biased proposal or importance pdf ignoring the minority particles [6, 7].

Traditional tracking methods for imbalanced data suffer from the problem of occlusion and identity switching between objects. This problem worsens when the cost of tracking false negatives and false positives is the same. The resampling methods apply three actions for the imbalanced data: data, algorithmic, and hybrid-level. In this paper, multiple object tracking is based on hybrid-level actions on the majority and minority particles. This paper discusses converting the biased proposal or importance pdf into the unbiased proposal or importance pdf using hybrid-level action for the particle filter.

The authors [8] provided a theoretical justification for an evolutionary particle filter to prevent the problem of sample impoverishment without the application of visual tracking of the objects. The authors [9] proposed a whale optimization algorithm (WOA), which is a highly efficient swarm-based metaheuristic algorithm stimulated by the hunting behaviour of humpback whales (H-whales) and determine the best solution. The WOA avoids local optima and gets a global optimal solution. The WOA solves different constrained or unconstrained optimization problems for practical applications without reformulation in the algorithm. This algorithm is highly efficient compared to other metaheuristic algorithms and is found to help handle imbalanced data.

In keeping with the advantages of metaheuristic algorithms and fractional calculus in dealing with the problem of sample degeneration and impoverishment, this paper proposes a new fractional whale particle filter (FWPF) which fuses a novel fractional whale optimization algorithm (FWOA) [9] with the standard particle filter for multiple object tracking (MOT) in the imbalanced surveillance video data with tracking challenges. Multiple objects from the initial video frame are extracted through human intervention, and the particles and related weights are generated. Then, the particles are updated using the FWPF method. This operation is repeated until the best solution is determined. In this paper, contributions are as follows:

- Development of a complementary fractional whale particle filter (FWPF) that fuses the fractional calculus-based whale optimization algorithm and the standard particle filter under weighted sum rule fusion. Integrating the FWPF with an iterative Gaussian mixture model (GMM) with unbiased sample variance and sample mean allows the proposal or importance probability density function (pdf) to be adaptive to deal with the imbalanced nature of surveillance video data. The adaptive importance or proposal pdf leads the FWPF to a minimum variance unbiased estimator (MVUE) for effectively detecting and tracking multiple objects in the imbalanced surveillance video data. The novelty of the work is to convert the fixed biased proposal pdf into an unbiased adaptive proposal pdf to deal with imbalanced data.
- A novel imbalanced evaluation metric, tracking distance correlation (TDC), is employed to evaluate the bias and variance correction (imbalance reduction) caused by an unbiased proposal pdf in the imbalanced FWPF-based tracking over the biased proposal pdf in the standard particle filter-based tracking. Apart from TDC, other performance measures, such as multiple objects tracking precision and accuracy (MOTP and MOTA), the sample mean and variance are used for comparative evaluation with other existing methods in the presented work.

This paper is structured as follows: The second and third sections relate to the related works and the challenges, respectively. The fourth section recounts the mathematical formulations for the standard PF. The fifth section introduces the proposed FWPF for MOT in imbalanced surveillance video data. The sixth section details the results and discussion. Finally, the seventh section winds up the paper as the conclusion.

2. Related Works

In particle filter-based visual tracking, the object's states are unknown and can only be deduced from visual feature observations. Visual tracking has two main difficulties: the state hypothesis evaluation of the observed image and the search for the state space. Metaheuristic algorithm-based particle filter tries to solve such problems in visual tracking. The two main visual tracking procedures are bottom-up and top-down procedures. The former typically favors rebuilding the states of the objects by examining the objects in the video sequences. At the same time, the latter is used to build and examine a state hypothesis set based on object models[10]. The MOT has many tracking challenges, such as occlusion, confusion between objects, scale changes of objects tracked, tracking resumption, and variability in the number of objects. There are two steps in MOT: In the first step, the number and position of targets are inferred from the current frame at every iteration, and this step requires either a generative or discriminative model, whereas the second step refers to the

modelling of detection errors and the movement of the target [11]. The Markov-Chain-Monte-Carlo (MCMC) framework approximates the posterior pdf adaptively and estimates the best state of objects using the Markov chain. This method encounters the problems of sample degeneration and impoverishment while tracking multiple objects [12].

The authors [13] proposed an evolutionary particle filter using modified grey-wolf metaheuristics to resolve sample impoverishment. The method enhances exploitation and exploration properties in tracking. However, the technique had limited exploration of dynamic target object behaviors. In [14], authors proposed a chaotic local search model for irregular motion with fewer particles. A chaotic sampling procedure enhances particle distribution in high-likelihood areas. The method effectively handles occlusion and fast-motion challenges by introducing a histogram of dynamic information for motion representation. The method exhibits superior performance in comparison to leading-edge approaches in tracking accuracy. The method has limited exploration of chaotic optimization in higher dimensions. The authors [15] proposed the combination of quantum particle swarm and adaptive genetic metaheuristics to focus on interference factors affecting tracking efficiency and accuracy. The method effectively handles occlusion and fast motion. In [16], the authors proposed a Mean shift algorithm for particle prediction and modified particle swarm metaheuristics for detecting occlusions in dynamic environments. The anti-occlusion module in this method improves tracking performance during occlusions.

The authors [17] proposed the covariance matrix-based learning with biogeography-based metaheuristics for real-world object tracking utilizing suitability index variables and color histograms. In [18], the authors proposed mayfly position-weighting strategies and dynamic control for multi-target tracking. Additionally, chaotic optimization strategies enhanced velocity parameters in mayflies. This method required addressing new and faded targets. The authors [19] proposed an improved firefly algorithm for optimization for multitarget tracking addressing the particle filter's weight degradation issue. The method fell into premature convergence quickly. The authors [20] utilized convolutional features and superpixel segmentation to address challenges in target association during complex scenarios of intelligent transportation. The method lacks extensive evaluation of diverse datasets. In [21], the authors proposed an iterative particle filtering based on ant colony metaheuristics and locality-sensitive histogram models to address the target appearance under illumination variations. The method discussed an iterative unscented Kalman filter for proposal distribution. However, the technique does not explore real-time tracking applications.

3. Challenges

The most difficult problem in visual tracking is the occlusion of objects, in which the portions of occluded objects are only seen, and the tracking becomes very difficult after the occurrence of object occlusion. The entry and exit of multiple objects in video sequences also pose problems for tracking systems. Tracking 3D visual persons has also been a challenging task in computer vision. Multiple object confusion leads to mismatches in tracking. Another difficulty is handling variable-size objects and their inter-occlusion in video sequences, affecting the tracking process in video surveillance and monitoring systems. The other challenges include noisy video data, changes in scene illumination, multi-camera tracking, etc. The current literature does not address the imbalanced data issues of surveillance video data but suggests different techniques for multiple object tracking (MOT).

4. Particle Filter Fundamentals

The main aim of the PF method is the estimation of a posterior pdf $p\left(x_n \middle| y_{1:n}\right)$ from a time-varying state-space model consisting of motion model specified by a state transition pdf $p\left(x_n \middle| x_{n-1}\right)$ and observation model specified by the observation likelihood pdf $p\left(y_n \middle| x_n\right)$ where, x_n is the state vector of an object at discrete time-instant n, $y_{1:n}$ is the set of observations i.e. video frames from discrete time-instant 1 to the current discrete time-instant n, and y_n denotes the video frame at discrete time-instant n [1, 2, 3]. In the PF method, a posterior pdf or filtering pdf is computed according to prediction step and filtering or update step under the two-step recursion when a new video frame y_n is considered [4, 5]:

$$p\left(x_n \middle| y_{1:n-1}\right) = \int p\left(x_n \middle| x_{n-1}\right) p\left(x_{n-1} \middle| y_{1:n-1}\right) dx_{n-1} \tag{1}$$

$$p\left(x_n \middle| y_{1:n}\right) = \frac{p\left(y_n \middle| x_n\right) p\left(x_n \middle| y_{1:n-1}\right)}{\int p\left(y_n \middle| x_n\right) p\left(x_n \middle| y_{1:n-1}\right) dx_n} \tag{2}$$

But, the recursive computation of the above posterior pdf is not possible analytically. Instead, it is numerically approximated by a set of M random particles with related weights w at discrete time-instant n, $\left\{x_n^m, w_n^m\right\}_{m=1}^{M}$ and it is given as follows:

$$p\left(x_n \left| y_{1:n}\right.\right) \approx \sum_{m=1}^{M} w_n^m \delta\left(x_n - x_n^m\right) \tag{3}$$

Where $\delta(\cdot)$ denotes the Kronecker delta function. The particles are random samples of the state vector, and the weights have non-negative values that sum up to unity, i.e. $\sum_{m=1}^{M} w_n^m = 1$ and are interpreted as probabilities of the particles. In practice, we often cannot draw samples directly from a posterior pdf because it may not be a standard pdf. Instead, particles are easily drawn from an importance or proposal pdf $\pi\left(x_n \left| x_{n-1}^m, y_{1:n}\right.\right)$ which is a biased pdf approximating the important regions of input variables. In (3), RHS is characterised as follows:

$$x_n^m \sim \pi\left(x_n \left| x_{n-1}^m, y_n\right.\right) \tag{4}$$

$$w_n^j = \frac{\tilde{w}_n^m}{\sum_{m=1}^{M} \tilde{w}_n^m} \tag{5}$$

$$\tilde{w}_n^m = w_{n-1}^m \frac{p\left(y_n \left| x_n^m\right.\right) p\left(x_n^m \left| x_{n-1}^m\right.\right)}{\pi\left(x_n^m \left| x_{n-1}^m, y_n\right.\right)} \tag{6}$$

If the importance pdf is made equal to the state transition pdf, i.e. $\pi\left(x_n^m \left| x_{n-1}^m, y_n\right.\right) = p\left(x_n^m \left| x_{n-1}^m\right.\right)$ which, then, simplifies to the above equation as $\tilde{w}_n^m = w_{n-1}^m p\left(y_n \left| x_n^m\right.\right)$. The initialization of PF requires the generation of a set of random particles, $\left\{x_0^m, w_0^m\right\}_{m=1}^{M}$ from the prior pdf $p\left(x_0\right)$. Equations (4) to (6) show the recursive computing of the set of random particles with related weights for the representation of $p\left(x_n \left| y_{1:n}\right.\right)$ from the previous posterior pdf $p\left(x_{n-1} \left| y_{1:n-1}\right.\right)$ corresponding to a new observation y_n. The number of efficacious particles is expressed as

$$M_n^{eff} = \frac{1}{\sum_{m=1}^{M} \left(\tilde{w}_n^m\right)^2} \tag{7}$$

In the PF method, all the particles are assigned the exact weights, and at each discrete time-instant, the particles are propagated and assigned weights. As time evolves, degeneration of a set of particles occurs, and only a few particles are left with meaningful weights. A well-biased importance pdf and a resampling procedure can reduce the sample degeneration to a small extent. Still, it presents the supplementary challenges of sample impoverishment, resulting in a deficiency of sample diversity in the state space. Recursive update calculations of ineffective particles seriously degrade the performance of particle filters. The resampling procedure eliminates the small weighted random particles and replicates the large weighted random particles. This resampling procedure poses the problem of sample impoverishment due to the over-replication of large weighted random particles and the reduction in the number of meaningful particles. Finally, it becomes difficult for the new set of particles to keep track of the true states of targets. To overcome this problem, this paper proposes FWPF for MOT in surveillance video data.

5. Proposed FWPF for Multi-object Tracking

Fig. 1 represents the proposed FWPF-based hybrid model for MOT in surveillance videos. The first step is object extraction, in which the objects are extracted from the video through manual intervention by considering the features of the objects. The next step is particle generation, in which the particles are generated randomly for the extracted object in each video frame. Then, particle updation is performed to select the position of the multiple objects in the video frame. The particle updation uses a hybrid model of standard particle filter and FWOA methods. The standard particle filter updates the particles by calculating the weight of each particle in the extracted object and discarding the particles with small weights. In this way, the particles are resampled, and the optimal position of the object is determined. WOA is a swarm-based metaheuristic that assumes the present solution is the best prey [9]. Then, the search agents update their position to the high likelihood region of the best search agent to find the optimal solution.

5.1. Object Extraction from the Input Video

Let us assume that the proposed MOT method's input is a video consisting of n frames. The input video taken as input to the proposed method is represented as $V=\{V_i;1\le i\le n\}$ where V represents the video and n represents the number of frames in the video. The primary task in the visual tracking is to discover the moving objects from the initial frame by extracting the foreground objects. The extracted object is then given for the tracking process, which involves the identification of the object's trajectories. So, the MOT method requires an object extraction mechanism either from the initial frame of the video or when the object appears initially in the video. In this MOT method, the object is extracted manually from the video sequence at the initial frame. The detected objects from the video through manual intervention are denoted as $O=\{O_j;1\le j\le k\}$ where O represents the multiple objects set and k represents the number of objects presented in the input video V .

The preprocessing technique is based on weakly supervised object detection (WSOD). An adaptive background subtraction approach is employed to initialize the detection of moving objects over the frames. After moving objects are detected in the initial frames with little human intervention, they are weakly labelled as objects or non-objects by applying a distance-based classification method, and the same is continued in successive frames within the small spatiotemporal neighbourhood. With known positions and weak labelling for objects or non-objects, real-world video datasets are always unlabelled and imbalanced because they contain anomalies somewhere in the frames, and their exact locations are unknown. These weakly labelled objects are refined and localized by the weighted sum rule fusion of the complementary trackers for multiple objects over the frames. The WSOD-based complementary trackers (FWPF and standard PF) are suitable for unlabelled and imbalanced video datasets where limited data availability is a major concern.

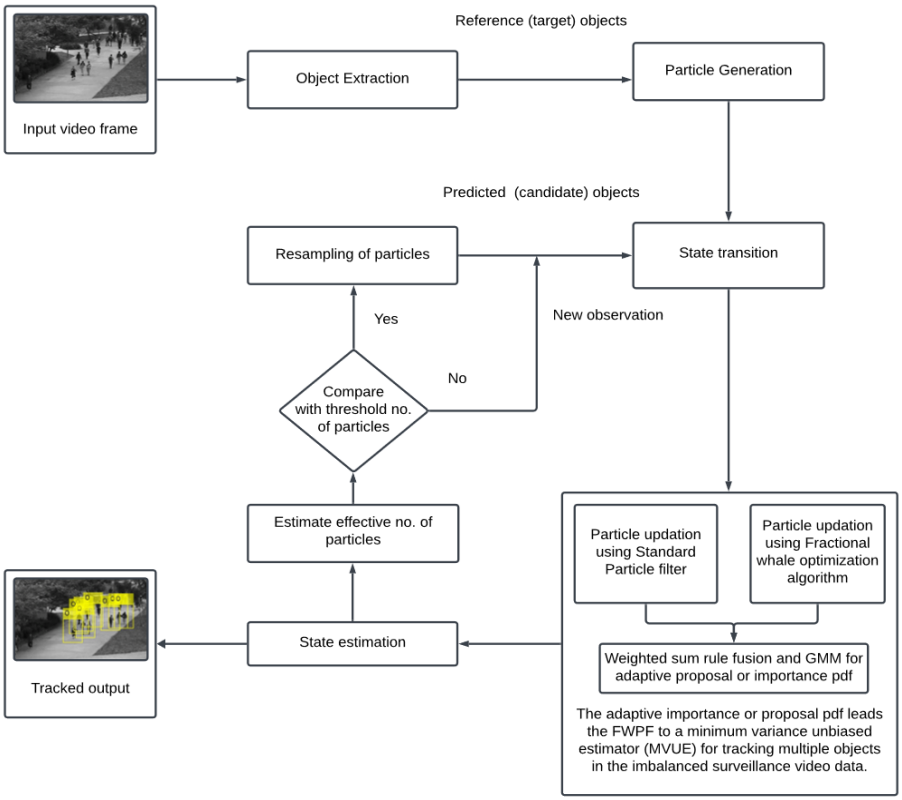


Fig.1. A proposed fractional whale particle filter method for MOT in the surveillance video

5.2. Multiple Object Tracking by Fractional Whale Particle Filter

This section describes the particle updating process in which the particles are updated using the standard particle filter and fractional calculus-based WOA method. Given the problem of sample degeneration and impoverishment in standard particle filters, the weighted sum rule fusion of FWOA and standard particle filter allows the biased proposal or importance pdf to be adaptive using the iterative Gaussian mixture model (GMM) with unbiased sample variance and sample mean.

A. Generation of Initial Particles

The particles are generated for every object after detecting the position of every object in the initial frame. Here, every object O_j contains v number of particles. The particles corresponding to every object are generated randomly to

evaluate the system's initial states. Here, the particles are used to describe the spatial and temporal attributes of moving objects in the video frames, and the particles are sampled based on the importance of the pdf.

B. Updating of Particles by Particle Filter for Visual Tracking

Once the particles are located in the initial frame, the movement (updating) of the particle is to be found to obtain the objects' location in the next frame. In a standard particle filter, the weight of each particle is determined based on the spatial and pixel correspondences. The spatial correspondence is computed based on the location of the objects using the Euclidean distance, and the pixel correspondence is computed based on the object's intensity using the Bhattacharyya distance. Once the weights of every particle are computed, they are multiplied by the previous position of the particle to find the location in the next frame. The new position of the object O_j at time T is determined as,

$$P_{O_j}^T(F) = \sum_{u=1}^{v} M_u^T C_u^T \tag{8}$$

Where v is the number of particles, is the centre of the particle u at time T and M_u^T is the weight of the particle u at time T. Here, the determination of the weight of each particle depends on the spatial pdf, $p\left(B_{O_j}^T \middle| C_u^T\right)$ and the colour pdf, $p\left(D_{O_j}^T \middle| C_u^T\right)$. The weight of the particle u at time T is determined as follows:

$$M_u^T = \frac{p\left(B_{O_j}^T \middle| C_u^T\right) p\left(D_{O_j}^T \middle| C_u^T\right)}{\sum_{e=1}^{v} p\left(B_{O_j}^T \middle| C_e^T\right) p\left(D_{O_j}^T \middle| C_e^T\right)} \tag{9}$$

The spatial pdf, $p\left(B_{O_j}^T \middle| C_u^T\right)$ is determined as follows:

$$p\left(B_{O_j}^T \middle| C_u^T\right) = \frac{1-k\left(y\left(C_u^T\right), y\left(P_{O_j}^{T-1}\right)\right)}{H} \tag{10}$$

Where, $y\left(C_u^T\right)$ is the centre of C_u^T and $y\left(P_{O_j}^{T-1}\right)$ is the centre of $P_{O_j}^{T-1}$. $k(\cdot)$ represents the Euclidean distance between $y\left(C_u^T\right)$ and $y\left(P_{O_j}^{T-1}\right)$. The denominator H in the above equation is computed based on the mean distance of the object region S located in the previous video frame with the height q and width p.

$$H = \frac{\sqrt{S_p\left(P_{O_j}^{T-1}\right)^2 + S_q\left(P_{O_j}^{T-1}\right)^2}}{2} \tag{11}$$

Then, the color pdf, $p\left(D_{O_j}^T \middle| C_u^T\right)$ is calculated as follows,

$$p\left(D_{O_j}^T \middle| C_u^T\right) = 1-W\left(g\left(P_{O_j}^{T-1}\right), g\left(C_u^T\right)\right) \tag{12}$$

Where, $W(\cdot)$ represents the Bhattacharyya distance obtained from (13). $g\left(P_{O_j}^{T-1}\right)$ and $g\left(C_u^T\right)$ denote the number of pixels in the bounding box surrounding the position of the object O_j at time $(T-1)$, i.e. $P_{O_j}^{T-1}$ and C_u^T, respectively.

Once the weights are computed, the particles which have lower weights are eliminated to improve the tracking accuracy. The remaining particles are the new set of particles for the next iteration. When a new set of particles is sampled, failure in the tracking gets repaired in the next few frames. In this way, the particles are resampled, and the object's new position is determined iteratively. The estimated result is applied to minimize the search area and the computation time during tracking.

C. Updating of Particles by Fractional-whale Optimization

The original whale optimization algorithm (WOA) copies the social behavior of the H-whales, which identifies the prey's location and encircles them. Since the placement of the optimal solution is unfamiliar at the beginning, the WOA

assumes the present solution as the target solution, beginning with a set of random solutions. Each iteration updates the search agent's position by either a randomly selected agent or the optimal solution. The process is repeated till it reaches the best solution. However, the WOA has a limited global search and a slow convergence rate [9].

The imbalanced surveillance video dataset introduces the drawbacks of overlapping objects during multiple object tracking, which causes problems with between-object imbalance and within-object imbalance with some missing or incomplete data. Such object imbalances occur when minority particles associated with low probable regions are not well separated from majority particles associated with high probable regions in specific areas of the input data space due to incomplete data. To counter this problem, the position of the particles is updated by the Fractional-whale optimization algorithm (FWOA), which aims to search the nearby position of the particles in the next frame to find the optimal position of the objects. This algorithm integrates the fractional calculus up to the first four terms, also known as adaptive multiple variable polynomials (AMFP), with the position update process of the whale optimization algorithm [22, 23].

The FWOA is a local and global search operator with inherent memory property. The FWOA adjusts the distribution of between-object and within-object generated particles using multiple imputation-based oversampling of minority particles in the targeted specific area of the input data space. The AMFP generates multiple different imputations for each missing data. If the missing or incomplete data are multivariate, they can be imputed from a joint multivariate normal or log-linear model. The AMFP imputes each majority particle only once for its missing data, whereas it oversamples each minority particle by replacing it with multiple imputations. Thus, all incomplete data are imputed with the oversampling of minority particles. One significant advantage of AMFP is that the oversampled minority particles are diversified with multiple imputations to eliminate data distortion, and all the observed data are kept unmodified. The FWOA approach results in low-bias and low-variance with greater efficacy for tracking precision and tracking distance in the imbalanced surveillance video dataset [24, 25].

i. Solution Encoding: Every solution is represented by a two-dimensional vector that indicates the position of the particles in the next frame.

ii. Fitness Calculation: The fitness is used to evaluate whether the generated solution is better. The fitness of every solution is calculated using the Bhattacharyya distance, which is given as follows:

$$Fitness = \sqrt{1-\frac{\sum_j \sqrt{l(j)m(j)}}{\sqrt{\sum_j l(j)\sum_j m(j)}}} \tag{13}$$

Where l and m represent the position of search agents in the present video frame and previous video frame, respectively.

iii. Algorithm: The algorithm for FWOA is explained below:

a. Initialization: The search agents are initially populated randomly and initiate their position randomly by surrounding the object's position. This process kicks up the search agents to move toward the likelihood region of the objects placed in the previous frame.

b. Fractional encircling prey: H-whales identify the prey's location and encircle them to find the position in the next iteration. In the shrinking encircling phase, the new position of the search agents is computed as follows:

$$\vec{G}(r+1) = \left(1-\vec{H}.\vec{F}\right)\vec{G^*}(r) - \vec{H}\cdot\left(\alpha\vec{G}(r) + \frac{1}{2}\alpha\vec{G}(r-1) + \frac{1}{6}\alpha(1-\alpha)\vec{G}(r-2) + \frac{1}{24}\alpha(1-\alpha)(2-\alpha)\vec{G}(r-3)\right) \tag{14}$$

Where $\vec{G^*}$ and $\vec{G}$ are the position vectors corresponding to the best solution and the search agent with r as the current iteration. $\vec{H}$ and $\vec{F}$ are both coefficient vectors. The encircling prey is modified with fractional calculus up to the four terms with fractional order of α to be in the range of $0 < \alpha < 1$. The fractional calculus is based on the generalized Grunwald-Letnikov Equation for the discrete fractional derivative with fractional order of α [24, 25].

Then, the updation of the coefficient vectors $\vec{H}$ and $\vec{F}$ are computed as follows:

$$\vec{H} = 2\vec{h}\cdot\vec{s} - \vec{h} \tag{15}$$

$$\vec{F} = 2\cdot\vec{s} \tag{16}$$

Where, $\vec{h}$ linearly decreases in the range $[2 \to 0]$ over the iterations, and $\vec{s}$ is a random vector in the range $[0,1]$.

c. Fractional exploitation phase: In the spiral updating position phase, the new position of the search agents is computed based on the following equation:

$$\vec{G}(r+1) = \left(1+e^{df}\cdot\cos(2\pi f)\right)\overrightarrow{G^*}(r) - e^{df}\cdot\cos(2\pi f)\cdot\left(\begin{array}{l}\alpha\vec{G}(r)+\frac{1}{2}\alpha\vec{G}(r-1)+\frac{1}{6}\alpha(1-\alpha)\vec{G}(r-2)+\\ \frac{1}{24}\alpha(1-\alpha)(2-\alpha)\vec{G}(r-3)\end{array}\right) \quad (17)$$

Where d is a constant denoting the state of the logarithmic helix in the range, f is a random number [-1,1] and '$\cdot$' is the element-wise multiplication. In the shrinking circling mechanism, the position of H-whales is updated as per (13) and (17) when the random probability of occurrence is less than or greater than $p = 0.5$.

d. Fractional exploration phase: The search agents search for prey (perfect solution) randomly, and the position of the search agent is updated based on the position of other search agents. This can be represented as:

$$\vec{G}(r+1) = (1-\vec{H}.\vec{F})\vec{G}_m - \vec{H}\cdot\left(\alpha\vec{G}(r)+\tfrac{1}{2}\alpha\vec{G}(r-1)+\tfrac{1}{6}\alpha(1-\alpha)\vec{G}(r-2)+\tfrac{1}{24}\alpha(1-\alpha)(2-\alpha)\vec{G}(r-3)\right) \quad (18)$$

Where, $\vec{G}_m$ is specified as a random position vector.

e. Termination: The FWOA is performed between either a circular or spiral movement, considering the value of p. Then, the process is repeated until it reaches the maximum number of iterations the user gives. After reaching the maximum iteration, the best prey location is assigned to $P_{O_j}^T(FW)$, which is the position of the particle in the next frame.

D. Weighted Sum Rule Fusion for Updating the New Position

The new position obtained through the PF and FWOA is combined with the weighted sum rule fusion to track multiple objects in the video sequence. The fusion provides the mean position of the PF and FWOA as the final tracking paths of the multi-objects. The following equation determines the new position of the object:

$$P_{O_j}^T = x_n = \frac{P_{O_j}^T(PF) + P_{O_j}^T(FW)}{2},$$

$$Z = \sum_{i=1}^{N=2}\alpha_i X_i \text{ such that } \sum_{i=1}^{N=2}\alpha_i = 1, \text{ and} \quad (19)$$

$$\sigma_Z^2 = \sum_{i=1}^{N=2}\alpha_i^2\sigma_{X_i}^2 = \alpha_1^2\sigma_{X_1}^2 + \alpha_2^2\sigma_{X_2}^2$$

Where $P_{O_j}^T(PF)$ is the new position of the object obtained by the PF method and $P_{O_j}^T(FW)$ is the new position of the object obtained by the FWOA. Since $P_{O_j}^T(PF)$ and $P_{O_j}^T(FW)$ are originated from the same video frame (data) distribution; they are not independent but are uncorrelated due to the FWOA-based tracker being complementary to the PF-based tracker. When treated as uncorrelated random variables, these complementary trackers produce a minimum variance, undergoing weighted sum rule fusion with equal weighting coefficients. If these complementary trackers achieve minimum variance for multiple objects, it will result in higher precision. The other score-level fusion techniques, like the maximum rule, minimum rule, product rule, median rule, etc., are not applied because these trackers are not independent [26, 27].

The weighted sum rule fusion is based on optimization by complementarity. One tracking model's incorrect detection and localization are complemented by another. If the trackers complement each other, they cancel out each other's random errors and help output correct or unbiased detection. These complementary trackers are uncorrelated but not independent because they originate from the same video frame (data) distributions. These trackers cause two complementary events which are mutually exclusive to each other. These trackers undergo the weighted sum rule fusion as per (19), where the terms X_i and α_i are uncorrelated random variables (two complementary trackers) and real weighting constants. The term Z is another random variable formed through weighted sum rule fusion. Further, we obtain the expression of the variance of Z in terms of the variances of X_1 and X_2 as given in (19). In statistics, precision is the reciprocal of the variance. In order to verify the approximated values of the precision, we assume $\sigma_{X_1}^2 = \sigma_{X_2}^2 = \sigma^2$ and consider different

values of α_1 and α_2. With $\alpha_1 = \alpha_2 = 0.5$, it is found that the averaging of two complementary trackers $\left(Z = \frac{X_1 + X_2}{2} \right)$ results in higher precision $\left(MOTP \propto \frac{1}{\sigma_Z^2} \right)$.

E. Adaptive Proposal pdf for the Imbalanced Data

The imbalanced surveillance video dataset consists of majority and minority classes as normal and anomalous instances. The normal and anomalous instances cause majority and minority samples associated with high and low probable regions. The minority particles tend to be at high risk of being suppressed by the majority particles, as the proposal or importance pdf encourages the highly probable regions of the input data space to remain a "biased distribution" in the standard particle filter. The proposal or importance sampling depends entirely on the choice of biased distribution, which results in a biased estimation of the discrete random measure. However, true and normalized importance weights are introduced to correct the bias to a small extent. The difficulty in designing the correct proposal or importance pdf prevents the use of particle filtering-based tracking in the imbalanced surveillance video dataset. Further, the standard particle filter tracker suffers from the problem of sample degeneration and sample impoverishment due to the biased proposal or importance pdf ignoring the minority particles [28, 29].

To alleviate this problem, The FWOA is fused with the standard particle filter under the weighted sum rule fusion to represent the proposal or importance pdf as a mixture of Gaussian distributions to deal with highly non-Gaussian true posterior pdf in the imbalanced data. Thus, FWOA adapts each component's weights and parameters in the GMM to result in the adaptive proposal or importance pdf for the FWPF, which balances exploring new areas of the state space with exploiting promising regions [30, 31].

i. Further to (19), a set of M random particles with associated weights u at discrete time-instant n, $\left\{ x_n^m, u_n^m \right\}_{m=1}^{M}$ is considered to denote x_n^1 and x_n^M as best (with a small selection pressure coefficient, α) and worst (with a large selection pressure coefficient, α) initial solution, whereas the associated weight u_n^m is defined to be Gaussian distributed with constraint as under:

$$u_n^m \propto \frac{1}{\sqrt{2\pi}\alpha M} \exp\left\{ -\frac{1}{2}\left(\frac{m-1}{\alpha M} \right)^2 \right\} \quad \text{such that} \quad \sum_{m=1}^{M} u_n^m = 1 \tag{20}$$

ii. The GMM, as the importance or proposal pdf, is generated for every p^{th} observations with Gaussian probability density function, characterized by the mean (μ) and variance (σ^2) as follows:

$$G_n^p\left\{ x_n^m(p) \right\} = \sum_{m=1}^{M} u_n^m \mathrm{N}\left(x_n^m(p); \mu_n^m(p), \left(\sigma_n^m(p) \right)^2 \right) \tag{21}$$

$$\mathrm{N}\left(x; \mu, \sigma^2 \right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{1}{2}\left(\frac{x-\mu}{\sigma} \right)^2 \right\} \tag{22}$$

Where $x_n(p)$ denotes the pth element of the vector x_n. The unbiased sample mean and sample variance for the observations drawn from the Gaussian pdf are defined as follows:

$$\mu_n^m(p) = \bar{x}_n^m(p) = \frac{1}{M} \sum_{m=1}^{M} x_n^m(p) \tag{23}$$

$$\left\{ \sigma_n^m(p) \right\}^2 = \frac{1}{(M-1)} \sum_{m=1}^{M} \left\{ x_n^m(p) - \bar{x}_n^m(p) \right\}^2 \tag{24}$$

iii. Thus, the GMM generates N new samples that represent the progeny of the primary population. The most optimal function is assessed for all the newly generated samples. The M_n^{eff} fittest solution among the M current samples and N new samples are formed to find the best solution. The adaptive importance or proposal pdf, $G_n\left(x_n^m \right)$ now modifies (6) with (21) and (22) as follows:

$$G_n\left(x_n^m\right)=\prod_{p=1}^{Q}G_n^p\left(x_n^m\left(p\right)\right)=G_n^1\left(x_n^m\left(1\right)\right)\times G_n^2\left(x_n^m\left(2\right)\right)\times\cdots\times G_n^Q\left(x_n^m\left(Q\right)\right) \tag{25}$$

$$\tilde{w}_n^m=\frac{\frac{1}{M}\sum_{l=1}^{M}p\left(y_n\left|x_n^m\right.\right)p\left(x_n^m\left|x_{n-1}^l\right.\right)}{G_n\left(x_n^m\right)} \tag{26}$$

Once the weights are modified according to adaptive importance or proposal pdf, the particles $\left\{x_n^m,\tilde{w}_n^m\right\}_{m=1}^{M}$ are completely updated/resampled based on the modified posterior pdf from (3) till the completion of the epoch n as follows:

$$p\left(x_n\left|y_{1:n}\right.\right)\approx\sum_{m=1}^{M}\tilde{w}_n^m\delta\left(x_n-x_n^m\right) \tag{27}$$

6. Results and Discussion

The proposed approach is simulated using the "MATLAB2019b" software tool on a system with an "Intel" core processor and 2 GB RAM, as well as the "Windows-11" operating system. The imbalanced UCSD video data – 1 and 2 [32] are obtained using a stationary camera mounted on the building, observing the pedestrian walkways where the density varies from sparse to crowded. Some occlusions and scale changes cause anomalies like cyclists, carts, and skateboarders to move at nonlinear speed. Each video has 100 frames with an image resolution of 158×238 pixels. The reason behind considering this video data is that the camera's field of view leads to a heightened degree of perspective distortion and pronounced variations in the scale of moving objects. Furthermore, anomalous events encompass motion anomalies elicited by cyclists, carts, and skateboarders, as well as contextual anomalies. The simulation results are discussed in this section.

6.1. Evaluation Metrics

Multiple objects tracking accuracy (MOTA) can reveal how class imbalance affects the overall tracking performance. For instance, a disproportionately high false negative rate for a minority class could indicate that the tracker consistently struggles to detect or maintain tracks of these minority or anomalous objects. In statistics, precision is often defined as the inverse of variance. Intuitively, variance measures how spread out a set of values is from its mean, while precision indicates how close or concentrated those values are to the mean. As a minimum variance unbiased estimator, the proposed method considers the Multiple objects tracking precision (MOTP), which can assess how accurately the tracker localizes those objects, regardless of their frequency in the video. Tracking distance correlation (TDC) can assess object tracks' temporal consistency and reliability. This is particularly important when dealing with imbalanced video data, where the performance of minority or anomalous objects might be compromised.

A. Multiple Object Tracking Accuracy (MOTA)

This metric incorporates three types of errors: false negatives, false positives, and identity mismatches. For the most accurate tracking, a high MOTA value is desired. The total count of ground truth objects is GT_t at frame (time) t. The terms FN_t, FP_t, and $IDMM_t$ are the number of false negatives, false positives, and identity mismatches at frame (time) t, respectively [33, 34].

$$MOTA=1-\frac{\sum_t\left(FN_t+FP_t+IDMM_t\right)}{\sum_t GT_t} \tag{28}$$

B. Multiple Object Tracking Precision (MOTP)

The metric refers to the average difference between all true positives and their ground truth hypothesis. A high MOTP value is preferred for the most precise tracking. The average dissimilarity among all true positives and their corresponding ground-truth hypothesis is the Multiple Object Tracking Precision [33, 34].

$$MOTP=\frac{\sum_{t,i}d_{t,i}}{\sum_t c_t} \tag{29}$$

Where the terms $d_{t,i}$ and c_t are the bounding box overlap of the target i with its assigned ground truth object and the count of matches at frame (time) t, respectively. It is used to represent the ability of the tracker to evaluate the correct position of the object positions in a video frame. MOTP and MOTA values vary between 0 (no tracking) and 1 (fully tracked). Higher values of both MOTP and MOTA are desirable for accurate and precise tracking.

C. Tracking Distance Correlation (TDC)

This metric measures both linear and nonlinear association between the original tracking path (X) and algorithmic tracked path (Y) in the imbalanced particle filter-based tracking. TDC value varies between 0 (when two paths are the same) and 1 (when two paths are independent). This metric depends on distance covariance, and its minimum value shows the bias and variance correction from tracking based on a biased proposal pdf to tracking based on an unbiased proposal pdf as required for FWPF. The tracking distance correlation is represented as follows [35]:

$$TDC(X,Y) = 1 - \frac{TDCov(X,Y)}{\sqrt{TDVar(X)TDVar(Y)}} \tag{30}$$

Where $TDCov(\cdot)$ and $TDVar(\cdot)$ represent the paths's distance covariance and distance variance.

6.2. Experimental Results

The performance of the proposed method is evaluated using two videos obtained from the UCSD dataset [32], each containing 100 frames. The UCSD video datasets are considered limited, unlabelled, real-world imbalanced surveillance video datasets. Figs. 2 (a) and (c) represent the 7th, 21st, and 35th frames of video-1 (Bicycle) and video-2 (Cart). The experimentation is performed to track the different numbers of multiple objects (from one to eight) from video-1 and video-2, as shown in Figs. 2 (b) and (d).

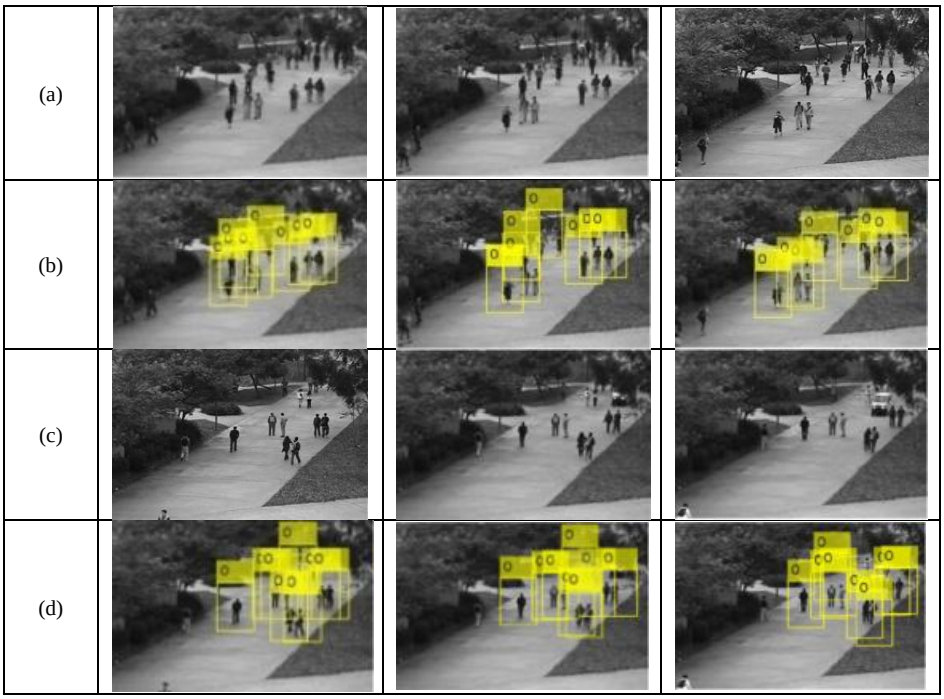


Fig.2. Proposed FWPF-based tracking over the original frames 7, 21, and 35 of the (a, b) UCSD video data – 1 [32] and (c, d) UCSD video data – 2 [32]

6.3. Performance Analysis

This section presents the comparative performance evaluation of the proposed FWPF with the existing methods like standard PF [1], Modified grey wolf particle filter (MGWPF) [13], Adaptive chaotic sampling particle filter (ACPF) [14], and Quantum particle swarm genetic particle filter (QPGPF) [15] when applying on the 100 frames extracted from UCSD video data – 1 and 2 [32].

A. Evaluation based on MOTP

Table 1 compares the average MOTP values of the proposed FWPF method with the existing methods [1, 13, 14, 15]

for the different numbers of individual tracked objects (from one to eight) in UCSD video data-1 and 2. Figs. 3 (a) and (c) depict the MOTP analysis curve for UCSD video data -1 and 2, respectively. The proposed method attains an average MOTP of 92.71% for UCSD video data -1, outperforming the MOTP values of 56.85%, 84.58%, 76.80%, and 70.63% achieved by PF, MGWPF, ACPF, and QPGPF, respectively. Similarly, the proposed method attains an average MOTP of 91.43% for UCSD video data -2, outperforming the MOTP values of 71.55%, 85.93%, 81.75%, and 78.82% achieved by PF, MGWPF, ACPF, and QPGPF, respectively.

B. Evaluation based on MOTA

Table 1 compares the average MOTA values of the proposed FWPF method with the existing methods [1, 13, 14, 15] for the different numbers of individual tracked objects (from one to eight) in UCSD video data-1 and 2. Figs. 3 (b) and (d) depict the MOTA analysis curve for UCSD video data -1 and 2, respectively. The proposed method attains an average MOTA of 90.73\% for UCSD video data -1, outperforming the MOTA values of 53.03%, 81.76%, 73.27%, and 67.31% achieved by PF, MGWPF, ACPF, and QPGPF, respectively. Similarly, the proposed method attains an average MOTA of 89.58% for UCSD video data -2, outperforming the MOTA values of 65.93%, 82.67%, 76.86%, and 75.46% achieved by PF, MGWPF, ACPF, and QPGPF, respectively.

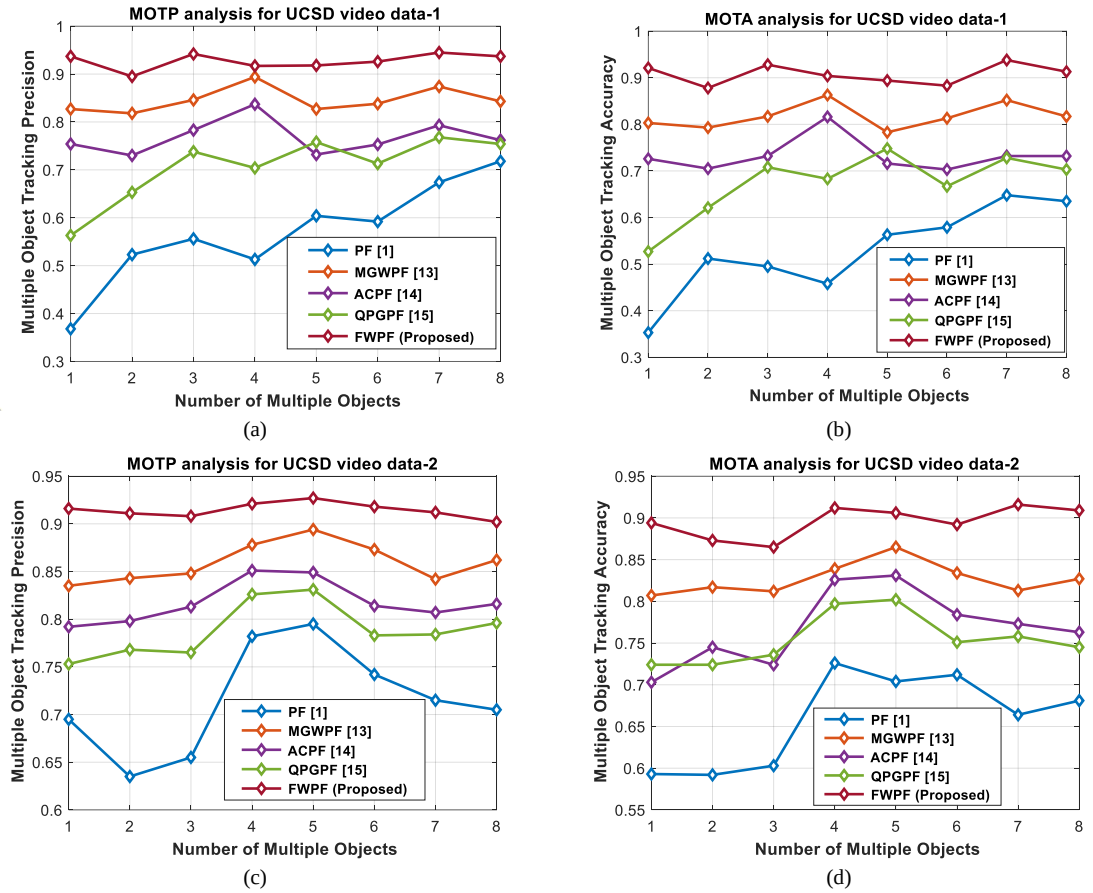


Fig.3. (a, b) MOTP and MOTA analysis curve and (c, d) MOTP and MOTA analysis curve for the UCSD video data - 1 and 2 [32]

C. Evaluation based on TDC

Table 2 compares the tracking distance correlation (TDC) values of the proposed FWPF method with the existing methods [1, 13, 14, 15] for the different numbers of individual tracked objects (from one to eight) in UCSD video data-1 and 2. Figs. 4 (a) and (b) depict the TDC analysis curve for imbalanced tracking over UCSD video data -1 and 2, respectively. The proposed method attains TDC value of 0.0038 (for tracking of eight multiple objects) for UCSD video data -1, outperforming the TDC values of 0.0388, 0.0166, 0.0214, and 0.0318 achieved by PF, MGWPF, ACPF, and QPGPF, respectively. Similarly, the proposed method attains TDC value of 0.0035 (for tracking of eight multiple objects) for UCSD video data -2, outperforming the TDC values of 0.0550, 0.0128, 0.0175, and 0.0319 achieved by PF, MGWPF, ACPF, and QPGPF, respectively.

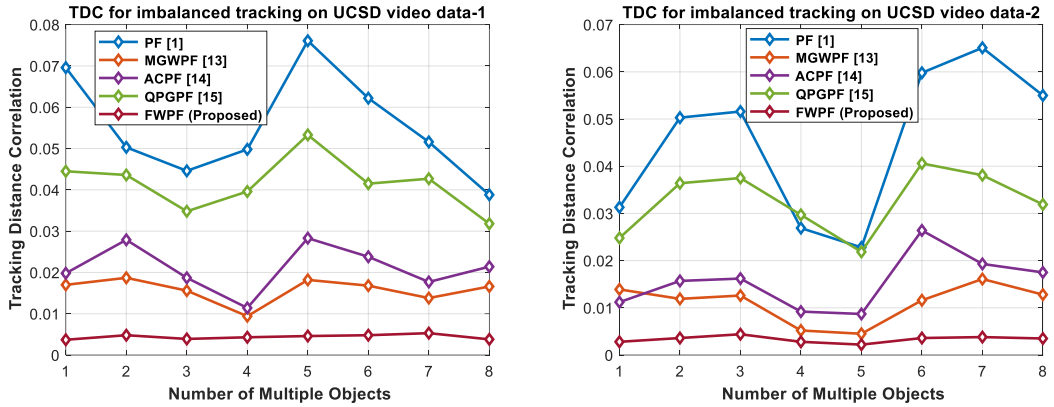


Fig.4. (a, b) Tracking distance correlation for imbalanced tracking on the UCSD video data – 1 and 2 [32]

Table 1. MOTP and MOTA comparison of methods in UCSD video data-1 and 2

Video	Methods	Average MOTP	Average MOTA
UCSD video data - 1 [32]	PF [1]	56.85	53.03
	MGWPF [13]	84.58	81.76
	ACPF [14]	76.8	73.27
	QPGPF [15]	70.63	67.31
	FWPF (Proposed)	92.71	90.73
UCSD Video data - 2 [32]	PF [1]	71.55	65.93
	MGWPF [13]	85.93	82.67
	ACPF [14]	81.75	76.86
	QPGPF [15]	78.82	75.46
	FWPF (Proposed)	91.43	89.58

Table 2. Tracking distance correlation for imbalanced tracking of (1 to 8) multiple objects over UCSD video data – 1 and 2

Video data	Methods	Tracking distance correlation for imbalanced tracking of (1 to 8) multiple objects over video data							
		1	2	3	4	5	6	7	8
UCSD-1 [32]	PF [1]	0.0696	0.0503	0.0446	0.0498	0.0761	0.0622	0.0516	0.0388
	MGWPF [13]	0.0170	0.0187	0.0156	0.0094	0.0182	0.0168	0.0138	0.0166
	ACPF [14]	0.0198	0.0279	0.0187	0.0114	0.0283	0.0238	0.0177	0.0214
	QPGPF [15]	0.0445	0.0436	0.0348	0.0396	0.0533	0.0415	0.0427	0.0318
	FWPF (Proposed)	0.0037	0.0048	0.0039	0.0043	0.0046	0.0048	0.0053	0.0038
UCSD-2 [32]	PF [1]	0.0313	0.0503	0.0516	0.0269	0.0228	0.0598	0.0651	0.0550
	MGWPF [13]	0.0139	0.0119	0.0126	0.0052	0.0045	0.0116	0.0161	0.0128
	ACPF [14]	0.0112	0.0157	0.0162	0.0092	0.0087	0.0264	0.0193	0.0175
	QPGPF [15]	0.0248	0.0364	0.0375	0.0297	0.02185	0.0406	0.0381	0.0319
	FWPF (Proposed)	0.0028	0.0036	0.0044	0.0028	0.0022	0.0036	0.0038	0.0035

D. Evaluation based on Sample Mean and Sample Variance

Tables 3 and 4 compare the sample mean and sample variance of the proposed FWPF method with the existing methods [1, 13, 14, 15] when applied over the imbalanced UCSD video data-1 and 2. The proposed method attains minimum sample mean and variance of 0.0044 (close to zero) and 3.25714 x 10^{-7} for UCSD video data - 1, as compared to sample mean (sample variance) values of 0.0554 (1.62203 x 10^{-4}), 0.0157 (8.8856 x 10^{-6}), 0.0211 (3.1324 x 10^{-5}), and 0.0415 (4.25221 x 10^{-5}) achieved by PF, MGWPF, ACPF, and QPGPF, respectively. Similarly, the proposed method attains minimum sample mean and variance of 0.0033 (close to zero) and 4.82679 x 10^{-7} for UCSD video data - 2, as compared to sample mean (sample variance) values of 0.0454 (2.57158 x 10^{-4}), 0.0110 (1.68079 x 10^{-5}), 0.0155 (3.45568 x 10^{-5}), and 0.0326 (4.56215 x 10^{-5}) achieved by PF, MGWPF, ACPF, and QPGPF, respectively.

Table 3. Sample mean and sample variance for imbalanced tracking of multiple objects (multiple observations) over UCSD video data -1

Methods	Sample mean	Sample variance
PF [1]	0.0554	1.62203×10^{-4}
MGWPF [13]	0.0157	8.88560×10^{-6}
ACPF [14]	0.0211	3.13240×10^{-5}
QPGPF [15]	0.0415	4.25221×10^{-5}
FWPF (Proposed)	0.0044	3.25714×10^{-7}

Table 4. Sample mean and sample variance for imbalanced tracking of multiple objects (multiple observations) over UCSD video data -2

Methods	Sample mean	Sample variance
PF [1]	0.0454	2.57158×10^{-4}
MGWPF [13]	0.0110	1.68079×10^{-5}
ACPF [14]	0.0155	3.45568×10^{-5}
QPGPF [15]	0.0326	4.56215×10^{-5}
FWPF (Proposed)	0.0033	4.82679×10^{-7}

E. Evaluation based on Bias-variance Correction for Imbalance Reduction

Tables 5 and 6 compare the bias-variance correction for imbalance reduction by the proposed FWPF method to the existing methods [1, 13, 14, 15] when applied over the imbalanced UCSD video data-1 and 2. The proposed method attains the correction of bias (variance) by -92.06% (-99.79%), -72.12% (-96.33%), -79.21 (-98.96%), and -89.40 (-99.23%) for imbalance reduction over UCSD video data - 1 when compared to PF, MGWPF, ACPF, and QPGPF, respectively. Similarly, the proposed method attains the correction of bias (variance) by -92.64% (-99.81%), -69.90% (-97.12%), -78.53 (-98.60%), and -89.77 (-98.94%) for imbalance reduction over UCSD video data - 2 when compared to PF, MGWPF, ACPF, and QPGPF, respectively.

Table 5. Bias-variance correction (imbalance reduction shown with negative sign) by the proposed method (FWPF as minimum variance unbiased estimator) in comparison with other methods for imbalanced tracking over UCSD video data -1

Methods	Bias correction (%)	Variance correction (%)
PF [1]	-92.06	-99.79
MGWPF [13]	-72.12	-96.33
ACPF [14]	-79.21	-98.96
QPGPF [15]	-89.40	-99.23

Table 6. Bias-variance correction (imbalance reduction shown with negative sign) by the proposed method (FWPF as minimum variance unbiased estimator) in comparison with other methods for imbalanced tracking over UCSD video data -2

Methods	Bias correction (%)	Variance correction (%)
PF [1]	-92.64	-99.81
MGWPF [13]	-69.90	-97.12
ACPF [14]	-78.53	-98.60
QPGPF [15]	-89.77	-98.94

F. Evaluation based on Computational Time and Limitations of Methods

Table 7 compare the computational time of the proposed FWPF method as compared to the existing methods [1, 13, 14, 15] when applied over the imbalanced UCSD video data-1 and 2. The proposed method processes the 100 frames each of UCSD video data-1 and 2 in 55.993 seconds and 25.775 seconds, respectively. The computational time is the lowest compared to the existing methods. The proposed method is suitable for limited (short-duration) video data obtained from fixed surveillance cameras under weakly supervised object detection settings. The limitations of other competing methods are as follows. In PF-based method, grid-based methods require a predefined state space, limiting uneven partitioning for high-probability density regions without prior knowledge. It highlights that the variance of importance weights increases over time, making degeneracy unavoidable. The computational cost of grid-based methods increases dramatically with state space dimensionality [1]. The MGWPF-based method may face challenges in motion blur and fast motion scenarios, leading to sample degeneracy. The convergence of the MGWO-PF algorithm can vary across different datasets, affecting performance consistency [13]. In the ACPF-based method, local representations fail to capture complex appearances effectively. Global representations are not robust to partial occlusions. Errors tend to accumulate during object model updates, leading to tracking failures under occlusions. The complexity and cost of global search pose an issue [14]. In the QPGPF-based method, the complexity of feature selection strategies is mentioned as a

burden on computational efficiency. The position updating equation has some limitations in optimizing the particle distributions with imbalanced video data. This method incurs higher computational costs despite improving target positioning accuracy [15].

Table 7. Computational time (in seconds per 100 frames of UCSD video data-1 and 2)

Methods	UCSD-1 [32]	UCSD-2 [32]
PF [1]	493.756	176.107
MGWPF [13]	119.621	54.758
ACPF [14]	145.072	67.198
QPGPF [15]	404.676	168.769
FWPF (Proposed)	55.993	25.775

6.4. Discussion

The proposed method is not meant for real-time tracking. Instead, it is used for limited, unlabelled real-world imbalanced video datasets. The limited video data from CCTV cameras for a particular time duration can be of important investigation by the law-enforcement authorities. Further, weakly labelled training data is generated using weakly supervised object detection. The weighted sum rule fusion of complementary trackers can further refine the localization of multiple objects. The proposed method tries to convert imbalanced tracking with a biased proposal pdf into balanced tracking with an unbiased proposal pdf in limited, unlabelled, real-world imbalanced video datasets. The conditions of extreme data imbalance with noisy input are generally found in the limited video data footage obtained from the video surveillance camera for a particular time duration. Weakly supervised object detection and refined localization by fusion of complementary trackers result in a better tracking model that deals with the imbalance in the limited video data footage [36]. As a minimum variance unbiased estimator, the proposed method exhibits bias-variance correction as an imbalance reduction factor, as shown in Tables 5 and 6.

The higher values of MOTP and MOTA are attained by the proposed FWPF method because it fuses the fractional calculus-based whale optimization algorithm and the standard particle filter under weighted sum rule fusion. Integrating the FWPF with an iterative Gaussian mixture model (GMM) with unbiased sample variance and sample mean allows the proposal or importance pdf to adapt to the imbalanced nature of surveillance video data. Thus, there are two trackers, one based on particle filter with a biased proposal or importance pdf and another based on FWOA with the adaptive proposal or importance pdf. Since these two trackers operate on the same video frame (data) distribution, they are not independent but are uncorrelated due to their complementary nature. These trackers deal with two complementary video data distributions that are mutually exclusive. These two trackers' weighted sum rule fusion leads the FWPF to a minimum variance unbiased estimation for effectively detecting and tracking multiple objects in the imbalanced surveillance video data. Since multiple objects are extracted with human intervention in the initial frame, making them weakly labelled (object or non-objects), the weighted sum rule fusion improves both MOTP and MOTA (with lowest TDC) over the frames due to the minimum variance achieved by these trackers being complementary. The FWPF, being an unbiased estimator, also results in the higher MOTP and MOTA overcoming the challenges of the imbalanced surveillance videos. The proposed method solves the problems of sample degeneration and impoverishment along with other tracking problems, such as scale change of objects tracked, occlusion, tracking resumption, etc [30, 31]. To justify the claim of the proposed method acting as a minimum variance unbiased estimator, the sample mean obtained from the multiple observations with different numbers of objects (from one to eight) is determined to be close to zero (unbiased) with minimum variance. The proposed method justifies the imbalance reduction through the bias-variance correction based on sample mean and sample variance obtained for multiple observations. The bias-variance correction (imbalance reduction) is attributed to an unbiased proposal pdf in the imbalanced FWPF-based tracking.

7. Conclusions

This paper fuses the fractional calculus-based whale optimization algorithm (FWOA) and the standard particle filter under weighted sum rule fusion. The fractional calculus up to the first four terms, known as adaptive multiple variable polynomials, makes the FWOA a local and global search operator with inherent memory property. Integrating the FWPF with an iterative Gaussian mixture model (GMM) with unbiased sample variance and mean allows the proposal or importance pdf to be adaptive to deal with the imbalanced surveillance video data, which is highly nonlinear and non-Gaussian. Thus, there are two trackers, one based on particle filter with a biased proposal or importance pdf and another based on FWOA with the adaptive proposal or importance pdf. In view of operating on the same video frame (data) distribution, these trackers are not independent but are uncorrelated due to their complementary nature. These two complementary trackers' weighted sum rule fusion leads the FWPF to a minimum variance unbiased estimation for effectively detecting and tracking multiple objects in the imbalanced surveillance video data. Since multiple objects are extracted with human intervention in the initial frame are weakly labelled (object or non-objects), the weighted sum rule fusion of these complementary trackers causes minimum variance, improving both MOTP and MOTA over the frames. The proposed method handles sample degeneration, impoverishment, and other tracking challenges, such as scale change

of objects tracked, occlusion, tracking resumption, etc. The resultant tracker, FWPF, achieves a higher MOTP of 92.71% (91.43%), MOTA of 90.73% (89.58%), TDC of 0.0038 (0.0035) for tracking of eight multiple objects in the imbalanced UCSD surveillance video data-1 (2) and outperforms the existing methods. The consistent tracking distance correlation near zero values infers the bias-variance correction for efficient imbalance reduction by the proposed method.

Acknowledgement

I confirm that the presented research work is the author's individual work. I also attest to the validity and legitimacy of the data and its interpretation. There are no conflicts of interest present, either with any individual or the organization.

References

[1] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking," IEEE Transactions on Signal Processing, vol. 50, no. 2, pp. 174–188, 2002, doi: 10.1109/78.978374.

[2] O. Cappe, S. J. Godsill, and E. Moulines, "An overview of existing methods and recent advances in sequential Monte Carlo," Proceedings of the IEEE, vol. 95, no. 5, pp. 899–924, 2007, doi: 10.1109/JPROC.2007.893250.

[3] M. Isard and A. Blake, "CONDENSATION-Conditional Density Propagation for Visual Tracking," Int J Comput Vis, vol. 29, no. 1, pp. 5–28, 1998.

[4] G. M. Rao and Ch. Satyanarayana, "Visual Object Target Tracking Using Particle Filter: A Survey," International Journal of Image, Graphics and Signal Processing, vol. 5, no. 6, pp. 57–71, May 2013.

[5] F. Gustafsson et al., "Particle Filters for Positioning, Navigation, and Tracking," IEEE Transactions on Signal Processing, vol. 50, no. 2, pp. 425–437, 2002.

[6] S. Wang, C. Shao, S. Xu, X. Yang, and H. Yu, "MSFSS: A whale optimization-based multiple sampling feature selection stacking ensemble algorithm for classifying imbalanced data," AIMS Mathematics, vol. 9, no. 7, pp. 17504–17530, 2024, doi: 10.3934/math.2024851.

[7] E. M. Hassib, A. I. El-Desouky, L. M. Labib, and E. S. M. El-kenawy, "WOA + BRNN: An imbalanced big data classification framework using Whale optimization and deep neural network," Soft comput, vol. 24, no. 8, pp. 5573–5592, Apr. 2020, doi: 10.1007/s00500-019-03901-y.

[8] S. Park, J. P. Hwang, E. Kim, and H. J. Kang, "A new evolutionary particle filter for the prevention of sample impoverishment," IEEE Transactions on Evolutionary Computation, vol. 13, no. 4, pp. 801–809, 2009, doi: 10.1109/TEVC.2008.2011729.

[9] S. Mirjalili and A. Lewis, "The Whale Optimization Algorithm," Advances in Engineering Software, vol. 95, pp. 51–67, May 2016, doi: 10.1016/j.advengsoft.2016.01.008.

[10] Y. Wu and T. S. Huang, "Robust Visual Tracking by Integrating Multiple Cues Based on Co-Inference Learning," Int J Comput Vis, vol. 58, no. 1, pp. 55–71, 2004.

[11] J. Berclaz, F. Fleuret, E. Türetken, and P. Fua, "Multiple object tracking using k-shortest paths optimization," IEEE Trans Pattern Anal Mach Intell, vol. 33, no. 9, pp. 1806–1819, 2011, doi: 10.1109/TPAMI.2011.21.

[12] X. Hai-Xia, W. Yao-Nan, Z. Wei, Z. Jiang, and Y. Xiao-Fang, "Multi-object visual tracking based on reversible jump Markov chain Monte Carlo," IET Computer Vision, vol. 5, no. 5, pp. 282–290, Sep. 2011, doi: 10.1049/iet-cvi.2010.0086.

[13] M. Narayana, H. Nenavath, S. Chavan, and L. Koteswara Rao, "Intelligent visual object tracking with particle filter based on Modified Grey Wolf Optimizer," Optik - International Journal for Light and Electron Optics, vol. 193, pp. 1–21, Sep. 2019, doi: 10.1016/j.ijleo.2019.06.013.

[14] M. Firouznia, J. A. Koupaei, K. Faez, G. A. Trunfio, and H. Amindavar, "Adaptive chaotic sampling particle filter to handle occlusion and fast motion in visual object tracking," Digital Signal Processing: A Review Journal, vol. 134, pp. 1–13, Apr. 2023, doi: 10.1016/j.dsp.2023.103933.

[15] X. Yuqi, W. Yongjun, and Y. Fan, "A scale adaptive generative target tracking method based on modified particle filter," Multimed Tools Appl, vol. 82, no. 20, pp. 31329–31349, Aug. 2023, doi: 10.1007/s11042-023-14901-4.

[16] S. Xu and H. Chen, "Visual Tracking via a Novel Adaptive Anti-occlusion Mean Shift Embedded Particle Filter," Circuits Syst Signal Process, pp. 1–26, Feb. 2024, doi: 10.1007/s00034-024-02882-0.

[17] J. Li, Z. Cao, F. Liu, Y. Fu, X. Li, and F. Tian, "An adaptive biogeography-based optimization with integrated covariance matrix learning for robust visual object tracking," Expert Syst Appl, vol. 234, pp. 1–24, Dec. 2023, doi: 10.1016/j.eswa.2023.121110.

[18] Y. Xiao and Y. Wu, "Robust visual tracking based on modified mayfly optimization algorithm," Image Vis Comput, vol. 135, pp. 1–8, Jul. 2023, doi: 10.1016/j.imavis.2023.104691.

[19] M. Tian, Y. Bo, Z. Chen, P. Wu, and C. Yue, "Multi-target tracking method based on improved firefly algorithm optimized particle filter," Neurocomputing, vol. 359, pp. 438–448, Sep. 2019, doi: 10.1016/j.neucom.2019.06.003.

[20] Z. Li, J. Chen, and J. Bi, "Multiple Object Tracking With Appearance Feature Prediction and Similarity Fusion," IEEE Access, vol. 11, pp. 52492–52500, 2023, doi: 10.1109/ACCESS.2023.3279868.

[21] F. Wang, Y. Wang, J. He, F. Sun, X. Li, and J. Zhang, "Visual object tracking via iterative ant particle filtering," IET Image Process, vol. 14, no. 8, pp. 1636–1644, Jun. 2020, doi: 10.1049/iet-ipr.2019.0967.

[22] S. Arora, T. Mathur, S. Agarwal, K. Tiwari, and P. Gupta, "Applications of fractional calculus in computer vision: A survey," Neurocomputing, vol. 489, pp. 407–428, Jun. 2022, doi: 10.1016/j.neucom.2021.10.122.

[23] T. P. Morris, I. R. White, J. R. Carpenter, S. J. Stanworth, and P. Royston, "Combining fractional polynomial model building with multiple imputation," Stat Med, vol. 34, no. 25, pp. 3298–3317, Nov. 2015, doi: 10.1002/sim.6553.

[24] M. S. Couceiro, R. P. Rocha, N. M. F. Ferreira, and J. A. T. Machado, "Introducing the fractional-order Darwinian PSO," Signal Image Video Process, vol. 6, no. 3, pp. 343–350, Sep. 2012, doi: 10.1007/s11760-012-0316-2.

[25] E. J. Solteiro Pires, J. A. Tenreiro MacHado, P. B. De Moura Oliveira, J. Boaventura Cunha, and L. Mendes, "Particle swarm optimization with fractional-order velocity," Nonlinear Dyn, vol. 61, no. 1–2, pp. 295–301, Jul. 2010, doi: 10.1007/s11071-009-9649-y.

[26] M. Dunnhofer, K. Simonato, and C. Micheloni, “Combining complementary trackers for enhanced long-term visual object tracking,” Image Vis Comput, vol. 122, pp. 1–10, Jun. 2022, doi: 10.1016/j.imavis.2022.104448.
[27] W. Dai, L. Jin, and G. Li, “Long-term adaptive tracking via complementary trackers,” IET Image Process, vol. 13, no. 9, pp. 1569–1577, Jul. 2019, doi: 10.1049/iet-ipr.2018.6142.
[28] Y. Xie, L. Peng, Z. Chen, B. Yang, H. Zhang, and H. Zhang, “Generative learning for imbalanced data using the Gaussian mixed model,” Applied Soft Computing Journal, vol. 79, pp. 439–451, Jun. 2019, doi: 10.1016/j.asoc.2019.03.056.
[29] M. Javvad ur Rehman, S. C. Dass, and V. S. Asirvadam, “A weighted likelihood criteria for learning importance densities in particle filtering,” EURASIP J Adv Signal Process, vol. 2018, no. 1, pp. 1–19, Dec. 2018, doi: 10.1186/s13634-018-0557-5.
[30] R. Gurajala, P. B. Choppala, J. S. Meka, and P. D. Teal, “A fast and unbiased minimalistic resampling approach for the particle filter,” in 2021 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), Kuala Terengganu, Malaysia, Sep. 2021, pp. 227–232.
[31] C. Liu, Y. Wang, D. Zhou, and X. Shen, “Minimum-Variance Unbiased Unknown Input and State Estimation for Multi-Agent Systems by Distributed Cooperative Filters,” IEEE Access, vol. 6, pp. 18128–18141, Mar. 2018, doi: 10.1109/ACCESS.2018.2815662.
[32] “UCSD Anomaly Detection Dataset.” [Online]. Available: www.svcl.ucsd.edu/projects/anomaly/dataset.htm
[33] K. Bernardin and R. Stiefelhagen, “Evaluating multiple object tracking performance: The CLEAR MOT metrics,” EURASIP J Image Video Process, vol. 2008, pp. 1–10, 2008, doi: 10.1155/2008/246309.
[34] Y. Park, L. M. Dang, S. Lee, D. Han, and H. Moon, “Multiple object tracking in deep learning approaches: A survey,” Electronics (Switzerland), vol. 10, no. 19, pp. 1–31, Oct. 2021, doi: 10.3390/electronics10192406.
[35] S. Feng, J. Keung, P. Zhang, Y. Xiao, and M. Zhang, “The impact of the distance metric and measure on SMOTE-based techniques in software defect prediction,” Inf Softw Technol, vol. 142, pp. 1–14, Feb. 2022, doi: 10.1016/j.infsof.2021.106742.
[36] X. Yang and L. J. Latecki, “Weakly Supervised Shape Based Object Detection with Particle Filter,” in 2010 11th European Conference on Computer Vision (ECCV), Springer Berlin Heidelberg, 2010, pp. 757–770.

Authors’ Profiles



Avinash Ratre received his Ph.D. in Electronics and Communication Engineering from the Indian Institute of Technology, Roorkee, India. Presently, he is working as an Assistant Professor in the Department of Electronics and Communication Engineering at Delhi Technological University, Delhi, India (ORCID ID: https://orcid.org/0000-0001-9803-2665). His research interests include computer vision, machine learning, signal processing, and wireless communication. He can be contacted at email: avinashratre@dtu.ac.in.

How to cite this paper: Avinash Ratre, "GMM-based Imbalanced Fractional Whale Particle Filter for Multiple Object Tracking in Surveillance Videos", International Journal of Computer Network and Information Security(IJCNIS), Vol.17, No.2, pp.34-50, 2025. DOI:10.5815/ijcnis.2025.02.03

2025 First International Conference on Advances in Computer Science, Electrical, Electronics, and Communication Technologies (CE2CT)

2025 First International Conference on Advances in Computer Science, Electrical, Electronics, and Communication Technologies (CE2CT) | 979-8-3315-1857-8/25/$31.00 ©2025 IEEE | DOI: 10.1109/CE2CT64011.2025.10939636

GRU Based Autoencoder with Maximum Overlap Discrete Wavelet Transfrmers for brain tumor detection

Bhawana Sachdeva
Electronics and communication engineering
Delhi Technological University
Delhi, India
er.barora@gmail.com

Jeebananda Panda
Electronics and communication engineering
Delhi Technological University
Delhi, India
jpanda@dce.ac.in

***Abstract*— A new approach based on autoencoder with GRU (gated recurrent unit) and discrete transforms for multi scale features of MRI images. Most of the systems fail to adequately handle sequentially dependencies or effectively leverage multi resolution scale. Therefore, in this paper we are analyzing a spatial and temporal features of brain tumor MRI images by imposing modified wavelet transforms for multi-scale analysis with a sequential modeling approach. Masoud Nickparver's Kaggle is used to train and test this model. In this study, 80% of the data are used for training and 20% of the data is used for testing. The implementation outputs give strong validation by achieving 98.88% classification accuracy for four classes. The F1 score and Precision of our novel framework are 98% and 98.30% respectively.**

Keywords*— Index Terms— MRI images, Spatial and Temporal features, Wavelet transform, GRU (gated recurrent unit), Autoencoder

I. INTRODUCTION

Brain tumors are masses created when abnormal growths in brain lesions occur inside of humans, directly affecting how well the brain functions to manage both voluntary and involuntary processes. It is a serious illness that contributes significantly to cancer mortality globally. The number of deaths from various cancers, including brain cancer, has increased, according to Cancer Statistics 2019 [1]. One effective non-invasive method for identifying brain anomalies is neuroimaging. The most widely used neuroimaging methods for diagnosing brain tumors are magnetic resonance imaging (MRI) and computed tomography (CT) scans. Using radio waves and powerful magnetic fields, MRI examines a body's physiology and anatomy [2] [3]. As a result, MRI is an effective method for identifying illnesses and physical abnormalities in the body. It can be used in many different medical imaging contexts. Nevertheless, it might be challenging to identify a tumor in a brain MRI; a specialist must review the pictures to ascertain whether a tumor is present. Additionally, determining the type of tumor and its location is crucial. CNN models are now a commonly utilized paradigm for MRI brain tumor identification, among other body part tumor detection. But these models give only spatial feature. In order to apply Computer Vision (CV) technology that can precisely detect and identify the existence of a tumor in an MRI, a strong solution is required [4][5][6]. Therefore, we are using the autoencoder and DWT to preserve the key characteristics of the time-series data while reducing the quantity of information. The autoencoder framework can handle complicated patterns in MRI images more effectively if DWT is incorporated. For local and global feature extraction, a probabilistic neural network (PNN) is designed where DCT, DWT, and PCA methods are used as decomposition and reduction [7]. The human hand is extremely agile and performs many different tasks. Therefore, author proposed recognition algorithms with the combination of wavelet neural network (WNN) and discrete wavelet transform (DWT) [8]. To convert the low resolution into high resolution, a combination of DWT for decomposition for enhancing the edges is designed to get the interpolation of high frequency sub bands The wavelet transform is used to break down the input data into its approximate and detailed components during the encoding process. This allows the autoencoder to concentrate on important characteristics while disregarding noise. In order to improve data structure preservation and minimize distortion, the signal can be rebuilt from its compressed representation during the decoding stage using the inverse DWT [9]. The detection and classification of MRI brain tumors are implemented using different wavelet transforms and support vector machines to differentiate between normal brain and tumor brain. But due to lack of scalability this method achieved low classification frequency [10]. Berkeley's wavelet transformation (BWT) and deep learning classifier have been used by the researchers to enhance performance and streamline the medical image segmentation procedure because of the large range of tumor tissues present in various individuals [11]. Utilizing the gray-level-co-occurrence matrix (GLCM) approach, significant features are retrieved from each segmented tissue and then optimized using a genetic algorithm. But this method was lack due to computational complexity and noise sensitivity [12]. A new technique for differentiating between benign and malignant brain tumours in MRI images uses contourlet transform and time adaptive self-organizing map, optimized by the whale optimization algorithm [13]. A single evaluation system that combines localisation, detection, and classification with other elements of image processing. Additionally, it makes use of the special powers of Ridgelet transformations to pinpoint the exact position of the cancer [14]. The Ridgelet-based approach used the Ridgelet transform in combination with the Grey Level Co-occurrence matrix for feature extraction and classification using Multi-Layer Preceptron (MLP) to classify MRI data into tumorous and non-tumorous, as well as meningioma and glioma.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:13:55 UTC from IEEE Xplore. Restrictions apply.

This guarantees a comprehensive assessment of the localisation procedure even further.

II. Proposed Method

This section discusses a proposed autoencoder for brain tumor detection using maximal overlap discrete wavelet transforms, based on GRUs (Gated Recurrent Units) [15]. The diagram representing the proposed methodology is shown in Fig 1. Data collection, pre-processing, autoencoder-based feature extraction, and classification are some of its four stages. Each step's comprehensive description is given here.

A. Data Collection

Masoud Nickparver's Kaggle platform provided the original input brain tumor data for this model [16]. There are 7022 brain MRI pictures available for analysis of multiple brain tumors. Together with 1621 Meningioma, 1600 No-tumor, and 1757 Pituitary classes, the dataset also contains 1645 Glioma. This dataset combines the Fighshare [17], SARTAJ [18], and Br35H [19] datasets.

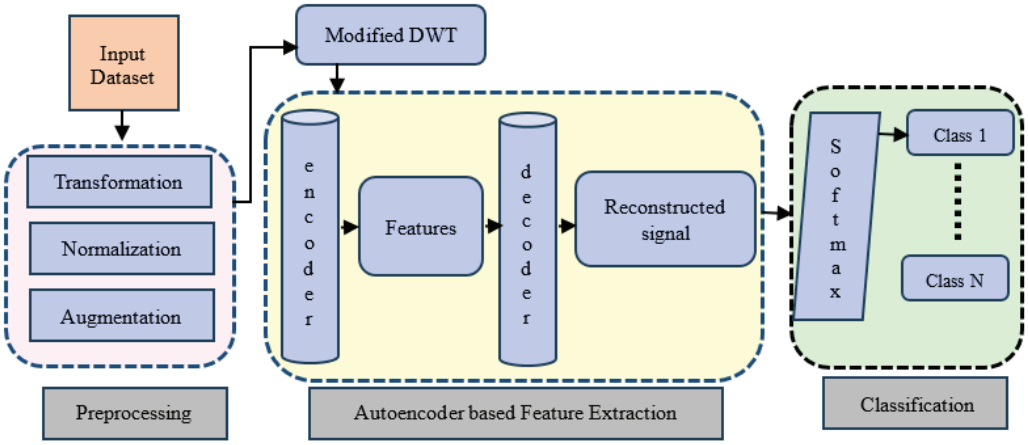


Fig. 1. Diagram representing proposed methodology

B. Data preprocessing

In this stage, the uniform resolution of all the images in the dataset is checked, and the pixel values are normalized to make processing them easier. An augmentation technique is used to further limit the problem of overfitting, as the size of the dataset in medical images is limited.

C. Autoencoder-Based Feature Extraction

After preprocessing the data, the images are first decomposed using maximum overlap DWT. To extract multi-scale features, this important step is completed. By doing this, images are transformed into wavelet coefficients, which are used to record frequency and spatial information. Here, using decomposed wavelet coefficients as features instead of raw pixel values improves the performance of the model. The decomposition of the image at various scales is provided as follows when maximum overlap DWT is applied to the input I_t:

$$I_t = \sum_{s=0}^{S} w_s(t) + C_S(t)$$

Where the wavelet coefficient is $w_s(t)$ and the scaling coefficient is $C_S(t)$.

Different resolutions of images are represented by set of coefficients given in the above transformation.

$$\{w_s(t)\}_{s=1}^{S}$$

These wavelet coefficients are features used by G RU-based encoders. GRU layers are now used to reduce the dimensions of wavelet transformed features.

In an encoder, GRU layers are stacked. Let the following be an input sequence:

$$I_t = \{w_s(t)\}_{s=1}^{S}$$

Where

$t = 1, \ldots\ldots, T$

Time steps are represented by T, whose vector of wave coefficient is given by I_t, because an MRI scan consists of multiple slices. Fig 2 illustrates the various hidden states z_t that each GRU layer possesses. GRU effectively captures sequential dependencies very efficiently at various levels. A reset gate and an update gate are the two gates. The update gate determines how much past state will be remembered, while the reset gate determines how much past state will be forgotten.

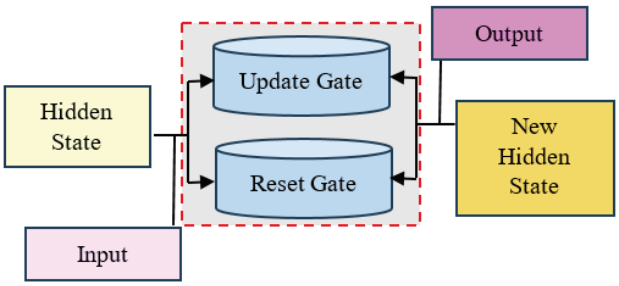


Fig. 2. GRU architecture

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:13:55 UTC from IEEE Xplore. Restrictions apply.

The update gate, represented by the symbol ug_t, and the reset gate, represented by rg_t, are computed as follows:

$$ug_t = \sigma(w_{ug}.I_t + U_{ug}.z_{t-1} + b_{ug})$$
$$rg_t = \sigma(w_{rg}.I_t + U_{rg}.z_{t-1} + b_{rg})$$

The latent vector is given to the classifier for brain tumor classification once the autoencoder has been trained.

III. Experiment Results

A. Experiment Results

The original input brain tumor dataset, Masoud Nicaraven's Kaggle [16] used to train and assess the proposed model. Eighty percent of each dataset's data is used for training and twenty percent is used for testing in the deep learning approach. A batch size of 16 photos and a learning rate of 0.001 were used to run the trained model 100 times. This model is optimized using Adam. The testing data's validation criteria yield the best classification results after 100 iterations. The classifier output of our model is displayed in Figs. 3 and 4, where we can view the classification accuracy and loss outcomes for various parameters. This graph shows that overall classification accuracy in Fig. 3 achieves excellent performance for the validation of the proposed model. In training, it attains a classification accuracy of 98.88 while in testing, it reaches 98%.

The resulting confusion matrix displays the classification outcomes of a model with 98.88% accuracy in predicting pituitary, meningioma, glioma, and no tumor pictures. The correct predictions are represented as diagonal values. The 272 glioma images were correctly categorized as well. Misclassifications are shown by off-diagonal numbers for example, 26 glioma images were mistakenly identified as meningiomas.

B. Discussion

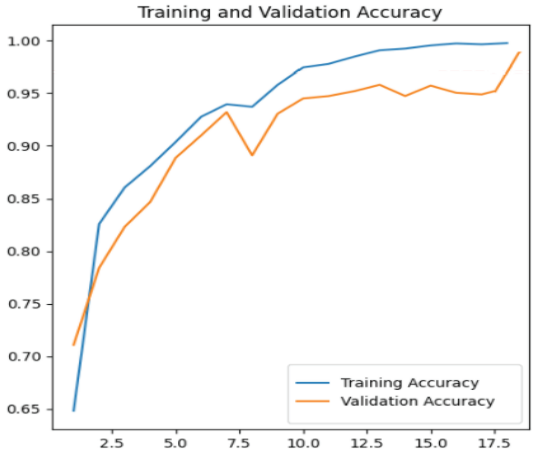


Fig. 3. Train and Test Accuracy

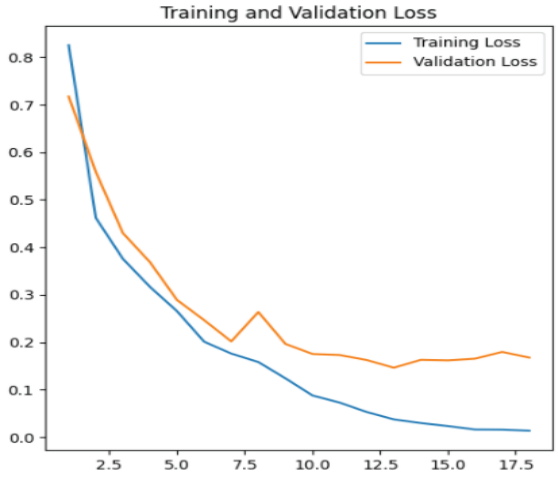


Fig. 4. Train and Test Loss

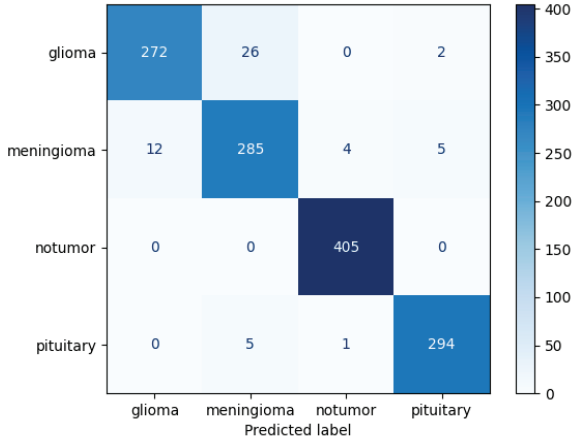


Fig. 5. Confusion Matrix

Table 1 contains a number of studies where academics have worked with various MRI picture difficulties to develop segmentation and feature extraction approaches. In order to identify tumour, it has been discovered with 86% classification accuracy using a support vector machine for classification and a discrete wavelet transform for decomposition. A model for MRI-based brain tumour identification by orthogonal wavelet transform and classification using deep learning achieved 98.5% classification accuracy and 98% precision. The contourlet transform based time adaptive self-organizing map is a unique approach to MRI brain tumour detection that obtained 98.5% classification accuracy, 97.9% F1 score and approximately 96.1% precision. Ridgelet transform, a revolutionary technique, was used to detect and localize brain tumours more accurately with classification accuracy of 98.65% of MRI images. Our proposed model, GRU with maximal overlap discrete wavelet has excellent results in comparison of existing models. The classification accuracy of our model is 98.88%, F1 score is 98% and precision is 98.3%.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:13:55 UTC from IEEE Xplore. Restrictions apply.

TABLE I. COMPARISON ANALYSIS OF DIFFERENT WAVELET MODELS

References	Year	Technique Used	Classification Accuracy	F1 Score	Precision
[20]	2017	DWT + SVM	86%	NR	NR
[11]	2022	Orthogonal Wavelet Transform	98.5	NR	0.98
[13]	2023	Contourlet Transform	98.5%	0.979	0.961
[14]	2023	Ridgelet transform	98.65%	NR	NR
Proposed	-	GRU with Maximal Overlap Discrete Wavelet	98.88%	0.98	0.983
NR-Not Reported DWT- Discrete wavelet transform SVM- support vector machine GRU- Gated recurrent unit					

IV. CONCLUSIONS

This paper provides a thorough examination of the issue of extracting tiny features from tumors in order to handle sequential dependencies or use multi-resolution scaling efficiently. The analysis of the spatial and temporal features of MRI images of brain tumors is being done using a new method based on autoencoder with GRU (gated recurrent unit) and discrete transforms for multi-scale analysis using a sequential modelling approach. This model is trained and tested using Masoud Nickparver's Kaggle platform. Eighty percent of the data in this study are utilized for training, while twenty percent are used for testing. Strong validation is provided by the implementation outputs, which achieve 98.88% classification accuracy across four classes. Our innovative framework has an F1 score of 98% and precision of 98.30%.

In future, this work can be extended to incorporate multi-modal data, such as integrating MRI with other medical imaging techniques. This can help to improve the robustness of brain tumor detection.

REFERENCE

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2019," CA. Cancer J. Clin., vol. 69, no. 1, pp. 7–34, Jan. 2019, doi: 10.3322/caac.21551.

[2] C. Westbrook and J. Talbot, Mri in Practice Fifth Edition. 2019.

[3] A. Comelli et al., "Tissue Classification to Support Local Active Delineation of Brain Tumors," 2020, pp. 3–14. doi: 10.1007/978-3-030-39343-4_1.

[4] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," Insights Imaging, vol. 9, no. 4, pp. 611–629, Aug. 2018, doi: 10.1007/s13244-018-0639-9.

[5] O. Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge," Sep. 2014, [Online]. Available: http://arxiv.org/abs/1409.0575

[6] M. Toğaçar, B. Ergen, and Z. Cömert, "BrainMRNet: Brain tumor detection using magnetic resonance images with a novel convolutional neural network model," Med. Hypotheses, vol. 134, p. 109531, Jan. 2020, doi: 10.1016/J.MEHY.2019.109531.

[7] F. Xu, T. Jia, and R. Jing, "Recognition of Key Information in Non-Stationary Signals Based on Wavelet Threshold Denoising and Back Propagation Neural Network Optimized by Manta Ray Foraging Optimization Algorithm," IEEE Access, vol. 10, pp. 118156–118166, 2022, doi: 10.1109/ACCESS.2022.3220365.

[8] A. M. Atto, E. Trouve, J.-M. Nicolas, and T. T. Le, "Wavelet Operators and Multiplicative Observation Models—Application to SAR Image Time-Series Analysis," IEEE Trans. Geosci. Remote Sens., vol. 54, no. 11, pp. 6606–6624, Nov. 2016, doi: 10.1109/TGRS.2016.2587626.

[9] H. Demirel and G. Anbarjafari, "IMAGE Resolution Enhancement by Using Discrete and Stationary Wavelet Decomposition," IEEE Trans. Image Process., vol. 20, no. 5, pp. 1458–1460, May 2011, doi: 10.1109/TIP.2010.2087767.

[10] M. Gurbina, M. Lascu, and D. Lascu, "Tumor Detection and Classification of MRI Brain Image using Different Wavelet Transforms and Support Vector Machines," in 2019 42nd International Conference on Telecommunications and Signal Processing (TSP), IEEE, Jul. 2019, pp. 505–508. doi: 10.1109/TSP.2019.8769040.

[11] M. Arif, F. Ajesh, S. Shamsudheen, O. Geman, D. Izdrui, and D. Vicoveanu, "Brain Tumor Detection and Classification by MRI Using Biologically Inspired Orthogonal Wavelet Transform and Deep Learning Techniques," J. Healthc. Eng., vol. 2022, 2022, doi: 10.1155/2022/2693621.

[12] Y. Hao et al., "Breast cancer histopathological images classification based on deep semantic features and gray level co-occurrence matrix," PLoS One, vol. 17, no. 5, p. e0267955, May 2022, doi: 10.1371/journal.pone.0267955.

[13] A. Farzamnia, S. H. Hazaveh, S. S. Siadat, and E. G. Moung, "MRI Brain Tumor Detection Methods Using Contourlet Transform Based on Time Adaptive Self-Organizing Map," IEEE Access, vol. 11, pp. 113480–113492, 2023, doi: 10.1109/ACCESS.2023.3322450.

[14] K. Chomu Basi and A. Ajay Yajnik, "Enhanced brain tumour detection and localization using ridgelet transform in MRI," Multimed. Tools Appl., Apr. 2024, doi: 10.1007/s11042-024-18923-4.

[15] G. Shen, Q. Tan, H. Zhang, P. Zeng, and J. Xu, "Deep Learning with Gated Recurrent Unit Networks for Financial Sequence Predictions," Procedia Comput. Sci., vol. 131, pp. 895–903, 2018, doi: 10.1016/j.procs.2018.04.298.

[16] "Brain MRI Masoud Dataset", [Online]. Available: https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset/data

[17] "Figshare Brain Tumor Dataset | Kaggle." Accessed: Apr. 19, 2023. [Online]. Available: https://www.kaggle.com/datasets/ashkhagan/figshare-brain-tumor-dataset

[18] "SARTAJ Brain tumor Dataset." [Online]. Available: https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri

[19] AHMED HAMADA, "Br35H :: Brain Tumor Detection 2020," kaggle. Accessed: Apr. 19, 2023. [Online]. Available: https://www.kaggle.com/datasets/ahmedhamada0/brain-tumor-detection?select=no

[20] A. R. Mathew and P. B. Anto, "Tumor detection and classification of MRI brain image using wavelet transform and SVM," in 2017 International Conference on Signal Processing and Communication (ICSPC), IEEE, Jul. 2017, pp. 75–78. doi: 10.1109/CSPC.2017.8305810.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 06:13:55 UTC from IEEE Xplore. Restrictions apply.

2025 IEEE 14th International Conference on Communication Systems and Network Technologies (CSNT) | 979-8-3315-3193-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/CSNT64827.2025.10968490

Hybrid Anomaly Detection Framework for Matter-Enabled IoT Devices

Vaibhav Krishan, Shivam Dhar, Kapil Sharma
Department of Information Technology
Delhi Technological University
Delhi, India
vaibhavkrishan3108@gmail.com, shivamdhar02@gmail.com, kapil@ieee.org

***Abstract*— Communications reliability and security of Matter-enabled IoT networks are now gaining even more prominence with their widespread acceptance and deployment. Nonetheless, such advancements indeed introduce further challenges, most notably in the detection of the diverse set of anomalies caused by device faults and DDoS attacks that may destabilize the network. This paper presents a hybrid anomaly detection system specifically designed for Matter-enabled IoT environments where rule-based filtering is incorporated with an ML-based detection approach leveraging Extreme Gradient Boosting (XGBoost). The rule-based module detects potential anomalies by applying predefined threshold conditions, and the XGBoost aims to curtail this detection into high detection capability and reduced false positives. The system is experimented with making the prediction on IoT telemetry data, achieving 97% precision, recall, and accuracy-greatly exceeding the lone detection strategies. The results prove how our approach can help mitigate the occurrence of false alarms, alongside imbalanced datasets being major obstacles in IoT networks' anomaly detection. Future research will further include the real-time detection application of sophisticated temporal models, scalable improvements, and security optimization within the confines of resource-constrained Matter-enabled IoT deployment.**

***Keywords*- Anomaly Detection, Matter-enabled IoT, Rule Based Filtering, XGBoost, Hybrid Detection System, Machine Learning**

I. INTRODUCTION

Matter-enabled IoT devices across diverse domains such as smart homes, industrial automation, and healthcare have created a new wave of connectivity and interoperability. These are the devices that provide cross-manufacturer interoperability, which means improved usability and performance in their operation. However, the greater the dependency on these systems, the consideration shifts to that of reliability of the network, security, and scale-out. Hardware failures, communication failures, and denial-of-service attacks have, therefore, to be found, as early as possible, in order to ensure the integrity of the system. The spread of IoT through adoption in smart homes, healthcare, and industrial automation serves to give an explosion of devices, hence terrific data throughput requiring advanced analysis [1]. Despite all these, IoT still faces many problems, such as many security measures are not implemented well in the devices in the resource-constrained environment, hence exposing the network to threats of DDoS attacks and data breaches [2]. These factors indicate the requirement of robust security and anomaly detection systems. IoT networks, especially those using protocols such as Thread, solve some of the challenges by enabling secure, low power IPv6 communication. The layers of Thread include 6LoWPAN and CoAP, which enhance scalability and efficiency. However, issues like standardization and power consumption remain a challenge for general adoption. Sensor reliability in dynamic IoT environments further requires advanced techniques, such as multi-agent distributed deep reinforcement learning, to guarantee correct anomaly detection and fault management [3].

This paper presents a hybrid anomaly detection system suitable for Matter-enabled IoT networks by combining rule-based filtering with XGBoost. Such a balance between simplicity and adaptability leads the system toward scalability and real-time anomaly detection, thus promising safe IoT ecosystems.

II. LITERATURE REVIEW

Mesh network for smart homes offers integrated advanced technology to provide more automation. This paper shows increased importance given by IoT to large grids to household devices. A four-layered sensor, gateway and network unit, man-servicing unit, and applications on the IoT laid increasing emphasis on mesh networks for reliability. Wireless Sensor Networks using low-cost devices, like the XBee module, showed improved communication robustness. The other solution is Message Queuing Telemetry Transport (MQTT), identified as a more suited solution for resource-constrained environments, low cost, and enriches smart home functionality.[4]

Most of the developments related to the theoretical foundations of Digital Twin technology in healthcare, implementation challenges, and combination of AI and IoT in creating Digital Twins of patients have been summarized. The barriers were first related to a lack of real-time data application and second a lack of empirical data to support the proposed models. The research reveals the revolutionary potential of Digital Twins in transforming patient monitoring and healthcare delivery, stressing the need for standardization in practice and stronger data protection

979-8-3315-3193-5/25/$31.00 ©2025 IEEE
DOI: 10.1109/CSNT.2025.66

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:29:57 UTC from IEEE Xplore. Restrictions apply.

standards to ensure confidence in these emerging technologies.[5]

The IoT unification of physical objects with an Internet instrument opens boundaries for numerous applications, yet they become of great concern for cybersecurity attacks owing to the adjacent and dispersed devices in their cyberspace. The state-of-the-art survey on IDS and the Internet of Things highlights issues not only with interoperability but also with immature research and pseudo coverage of the various IoT attack vectors. Anomaly-based IDS, signature-based IDS, and specification-based IDS lack a consensus on the optimal placement of detection strategies and on how to establish them. This paper recommends giving further attention to extending any form of attack coverage, equipping the same with emerging technologies, and developing adaptable frameworks.[6]

The rapid proliferation of IoT devices creates a strong demand for effective, secure authentication protocols that can handle resource-constrained environments. Classic schemes involving time synchronization and nonce storage are highly impractical on low-power devices, leading to vulnerability against replay attacks. The innovative approaches introduced here - for example, a new challenge-response mechanism that achieves security with lower overhead compared to the use of time synchronization and nonce storage - strengthen authentication but minimize overhead, this represents a critical step in securing the Internet of Things (IoT) and its application-related interconnectivity.[7]

IoT has integrated itself into the daily life, but due to cost, size, and power constraints, the systems remain insecure against attacks. Distributed Denial of Service attacks are a severe threat to the IoT networks. This paper is a review of IDS and its ability to provide a boost with machine learning and deep learning. Gaps in literature of IDS datasets and anomaly detection methods are identified, and a resilient security framework is proposed against the evolving threats of IoT.[8]

This paper explores the deployment of honeypots on Raspberry Pi devices aims to increase IoT security in areas like health, where privacy and data must be safeguarded. Cowrie, Dionaea, and Glastopf are the honeypots used by the authors for attack pattern detection. Dionaea scans several protocols, which gives it the opportunity to detect a threat by looking into FTP login attempts and Denial of Service attacks, whereas Glastopf detects web attacks. Cowrie is a high-interaction honeypot, since it records interactions with the attacker. Observations regarding the low-cost honeypot configuration indicate that this honeypot is effective for detecting threats and gathering forensic data on IoT security in the face of evolving cyber threats. [9]

This paper examines a Sensor Fault Detection paper in Internet of Things (IoT) examines the issues occurring due to hostility and wrong reads. Although wireless sensor networks' research is mainstream, IoT would require specialized approaches because the context is different from WSN, and this discussion will give knowledge about outlier methods, advantages and disadvantages, of which the entire systematic literature study and industry review will be found. The study suggests a multi-agent distributed deep reinforcement learning approach to improve sensor reliability, addressing faults and anomalies in IoT environments.[10]

This paper covers the advantages and deployment techniques of wireless mesh networks in empowering IoT systems. As, wireless mesh networks are self-organizing topologies based on the ad hoc networking model and facilitate scale-up, robustness, and cost-effectiveness. They provide extended seamless coverage with minimal costs and eliminate single points of failure, thus increasing robustness. Transitioning from star networks to hybrid wireless mesh networks provides an enhanced performance of IoT systems. Challenges like possible network delays and infrastructural compatibility issues, still necessitating further research for large-scale deployment, exist.[11]

Data splitting is essential for unbiased model validation in statistics and machine learning. Holdout data, calibration, and validation increase accuracy. The R package Split is designed to optimize training/testing ratios that enhance inference reliability. On the other side, high-dimensional classifiers heavily rely upon resampling and optimal partitioning to generalize data. Random matrix theory provides advances in data splitting. Thereby assuring robust practices in data sciences and the development of trustworthy predictive models. [12]

III. DATA DESCRIPTION

The dataset adopted for this experiment sourced from telemetry information from an IoT communication system, targeted toward performance metrics for anomaly detection.

It was also developed to resemble real-world conditions within IoT and assist in classification models for detection of system behaviour. Some major points regarding the dataset are elaborated below.

*A. **Structure and Metrics***

The dataset includes key telemetry metrics necessary for network performance analysis. **Packet Drops** refer to dropped packets, which indicate both network instability and congestion. **Retry_Counts** refer to the number of retransmissions, which are useful in the analysis of interference and failures in delivery. **Message Delivery Time**, expressed in milliseconds, is the time required for delivery of a message when it is passed and successful in transmission, indicating system latency. The network states are classified using a Label (with 0 signifying a normal state and 1 indicating an anomaly) to enable proper anomaly detection and performance diagnosis.

*B. **Composition and Balance***

The dataset consists of normal and anomalous traffic and thus provides a balanced representation of communication system behaviours. Normal Traffic, or Class 0, reflects the standard patterns of communication when everything is

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:29:57 UTC from IEEE Xplore. Restrictions apply.

functioning normally and forms the bulk of the dataset. Class 1 or anomalous traffic, therefore captures instances of the behaviour of an anomalous system that fails to communicate or the inefficiency in network operation and represents a much smaller proportion of the dataset. Robustness and relevance were ensured through the dataset being curated from a pool of 1,000 records. Preprocessing was conducted to remove about 15% of noise and outliers. Thus, the final dataset has 850 high-quality instances, and normal and anomalous traffic is well balanced.

C. *Realism and Future Scope*

The dataset is synthetically generated but structured and designed to mimic the behaviour of a real-world IoT network in close detail. Nevertheless, there remains a major challenge in acquiring large, high-quality datasets from real IoT deployments.

IV. Proposed Methodology

A. *Architecture Overview*

This proposed hybrid anomaly detection system is for accurately and with very few false positives identifying anomalies on Matter-enabled IoT devices. There are four core modules in its architecture:

1) Data Collection Module:

- It collects the telemetry data coming from Matter enabled IoT devices.
- All of the monitored metrics, packet drops, retry counts, and delivery times of the message, will be logged continually.
- The gathered data is also put into structured formats and analysed.

2) Pre-Processing Module: The Pre-Processing Module is used for preparing raw telemetry data, which is noisy and inconsistent, to be fed into the anomaly detection module. The module begins by using Min-Max ormalization, which scales features so that all the data falls in the range

[0, 1], improving the performance of machine learning models. The data is divided into training and testing sets for supervised learning so that the model is trained on one portion while validated on another. For analysing temporal patterns, the data can also be divided into fixed-length windows, which enables more accurate insights when future extensions of the project necessitate handling sequential data. This pre-processing step ensures the data is clean, well-structured, and ready for anomaly detection algorithms.

3) Detection Module: It has its two fundamental modules namely, Rule-Based Filtering and XGBoost Classifier. Rule-Based filtering is meant for fast and quasi-real-time identification of suspect anomalies by applying operators established through previously set statistical thresholds called rules determined by analysts. For example, there could be an indication of anomalies once packet drops cross a threshold of mean + 2σ (packet_drops > mean + 2σ). Non-flagged data denotes those data points that do not cross above these thresholds and therefore are considered normal with no analytical measure. Based on choosing key metrics such as packet drops, retry counts, and message delivery time, those that exceeded a certain threshold were flagged as anomaly (1); the rest would be marked normal (0). This obviously helps in avoiding unnecessary processing and improving computational efficiency. The rule-based filtering is both computationally easy and efficient for such real-time detection but lacks adaptability to the changes in network behaviour over time. Since this limitation should not lead to the crumbling of the very foundation of anomaly detection work, some flagged data are resorted to by XGBoost Classification in the sense that it will be training the model using telemetry data for classifying anomalies (label 1) from false positives (label 0). Performance is assessed on XGBoost against classification metrics like precision, recall, F1-score, and overall accuracy. The dataset is split into 80% training and 20% testing, with the actual values, y_test, compared against predictions, y_pred. Precision indicates the total percentage correctly labelled as anomalies; recall estimates the model's capacity for finding all real anomalies; whereas the F1-score is that balanced view of both. Overall accuracy is how much data is correctly predicted: the proportion of total observed data this rate captures still makes sense.

4) Decision Engine: The Decision Engine aggregates the result of the Detection Module to finally produce the anomaly detection outputs.Non-flagged data is automatically categorized as normal (0) since it does not fit any abnormal activity. For data that is flagged, the XGBoost model processes it and gives it a classification either as true anomalies (1) or false positives (0). Other than classification, the engine also updates system logs to aid in post-incident analysis; each anomaly detected is logged for future checking. It also fires immediate alerts by showing the anomalies noticed at the end for the relevant stakeholders of the respective problems, ensuring response time.

B. *Anomaly Detection Pipeline Digram*

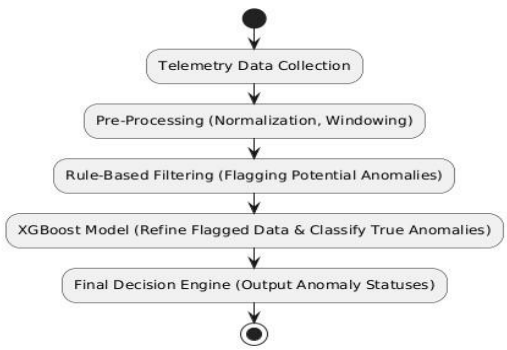


Figure 2: Detection Pipeline for Anomaly Detection

C. *Hybrid Anomaly Detection Algorithms*

1) Rule-Based Detection: The rule-based detection system identifies potential anomalies by applying predefined thresholds to telemetry data.

2) XGBoost Algorithm: The XGBoost-based anomaly detection algorithm refines the results of the rule-based system by leveraging supervised learning. It trains on

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:29:57 UTC from IEEE Xplore. Restrictions apply.

flagged data points to classify them as true anomalies or false positives. The use of gradient-boosted decision trees ensures high accuracy, recall, and precision, making it suitable for IoT telemetry data analysis.

3) Hybrid Detection System: The hybrid system integrates rule-based filtering and XGBoost refinement for robust anomaly detection. The pseudocode for the hybrid detection system is as follows.

Algorithm: Hybrid Anomaly Detection System

Procedure:hybrid_anomaly_detection(dataset): [A hybrid approach combines rule-based detection and machine learning for improved anomaly detection]

Step 1: Rule-Based Detection

1. Define thresholds for anomaly detection:

packet_drops_threshold := mean(packet_drops) + 2 std(packet_drops)

retry_counts_threshold := mean(retry_counts) + 2 std(retry_counts)

message_delivery_time_threshold := mean(message_delivery_time) + 2 std(message_delivery_time)

2. For each data point row in the dataset:

if

packet_drops > packet_drops_threshold **or**

retry_counts > retry_counts_threshold **or** message_delivery_time > message_delivery_time_threshold:

Flag the data point as anomalous (1).

else:

Flag the data point as normal (0).

3. Store the flags in the rule_based_flag column.

Step 2: Data Splitting

4. Separate data based on rule-based flags:

Approach	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
Rule-Based Filtering	92	87	89	90
XGBoost	100	99	99.5	99
Hybrid Model	97	97	96	97

flagged_data := data points where rule_based_flag = 1

non_flagged_data := data points where rule_based_flag = 0

5. Prepare flagged_data for machine learning:

Features: X := [packet_drops, retry_counts, message_delivery_time]

Labels: y := label.

6. **if** flagged_data has less than two unique classes:

Use rule-based predictions directly:

predictions:= rule_based_flag.

else: Proceed to Step 3.

Step 3: Train XGBoost on Flagged Data

7. Perform train-test split:

Divide X and y into X_train, X_test, y_train, and y_test using an 80-20 ratio.

8. Train an XGBoost model with the following parameters:

- n_estimators := 100
- max_depth := 3
- learning_rate := 0.1
- random_state := 42.

9. Fit the model on X_train and y_train.

10. Predict on X_test:

- y_pred := predictions on X_test.

11. Evaluate XGBoost performance using a classification report comparing y_test and y_pred.

12. Use the model to generate predictions for all flagged_data.

Step 4: Handle Non-Flagged Data

13. Assign predictions for non_flagged_data:

predictions := 0 (normal).

14. Combine flagged_data and non_flagged_data to create final_results.

Step 5: Evaluate Overall System

15. Compute the classification report comparing final_results[label] and final_results[predictions].

16. Return the overall system performance metrics.

V. EXPERIMENTAL RESULTS

A. Performance of the Anomaly Detection Model: The proposed hybrid anomaly detection system, tailored for Matter-enabled IoT devices. The rule-based detection method, although efficient in identifying anomalies based on simple thresholds, suffered from limited recall (87%) and high false positives due to its rigid nature. Conversely, the XGBoost model, trained on flagged anomalies, achieved perfect precision, recall, and F1-Score (100%), showcasing its ability to refine classifications effectively. The combined **hybrid system** improved overall performance by leveraging the strengths of both approaches. With precision at 97%, recall at 97%, and an F1-Score of 96%, the hybrid system achieved a balanced trade-off between sensitivity and specificity, making it suitable for real-time anomaly detection in IoT environments.

Table 1: Performance Metrics by the System

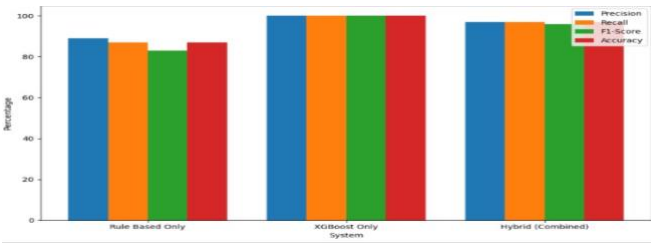


Figure 3: Performance Metrics by System

B. Parameter Tuning and Feature Analysis:

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:29:57 UTC from IEEE Xplore. Restrictions apply.

The parameter tuning process for the XGBoost model was key parameters of hyperparameter optimization, which were to maximize detection accuracy. The learning rate was set to 0.1 to maintain a balance between speed and accuracy in training; the maximum depth was adjusted to 5, which prevents overfitting, and more subtle interactions among features can be detected; and the number of estimators was set to 100 to retain efficiency without compromising performance. By applying these settings, the model showed good generalization to unseen data, with F1-score, as well as overall accuracy to be high. In addition, performing feature importance analysis demonstrated the importance of telemetry metrics in terms of anomaly identification. A packet drop was found to be 40% responsible for the aggregate importance, after retry counts account for 35%, and for message delivery times, 25%.

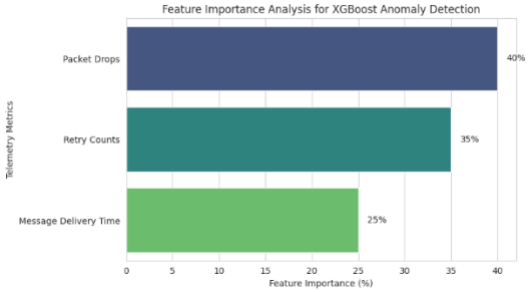


Figure 4: Feature Importance Analysis

VI. DISCUSSION

The hybrid approach successfully overcame the limitations of standalone methods since it effectively implemented real time anomaly detection for Matter-enabled IoT networks. Combining rule-based filtering for quick flagging of anomalies with XGBoost for more accurate classification helped achieve higher accuracy and minimized false positives. However, challenges such as data imbalance, where anomalies were significantly outnumbered by normal instances, required mitigation using oversampling techniques like SMOTE. The system still requires further optimization for real-time deployment on the resource-constrained IoT devices in order to run seamlessly. With these challenges aside, the system's performance puts it as one of the potential robust frameworks in anomaly detection on IoT environments with balance between precision and recall and still being adaptable to dynamic network behaviours.

VII. CONCLUSION

The main challenge addressed here is that in anomaly detection by Matter-enabled IoT networks, system reliability and security are important requirements. The number of current approaches, including the purely rule-based systems or standalone machine learning models, are deficient in their potential applications with serious limitations, like high false positives, lack of adaptability, and computational inefficiencies, under the harsh environment of resource constraints in IoT scenarios. This calls for a hybrid model that will bridge the gap of the accuracy and efficiency. The studied anomaly detection system is a hybrid system combining rule-based filtering and classification with XGBoost. It applies the rule-based system to the first level of anomaly detection by using predefined statistical thresholds and refines the flagged anomalies by using XGBoost to limit false positives and improve the accuracy of detection. The system was tested on telemetry data from IoT; it obtained 97% precision, 97% recall, and 96% F1-score, significantly outperforming the standalone rule-based approach.

Key Findings of the Study:

1. Rule-Based Detection is a computationally anomaly filtering mechanism, but its low recall (87%) causes many anomalies to be missed.
2. The XGBoost Model delivered 99 percent accuracy on flagging data that shows it worked well for perfecting the anomalous classification accuracy.
3. The Hybrid System has excellent balance between sensitivity and specificity.

This work contributes to a novel hybrid anomaly detection approach specifically suited to Matter-enabled IoT networks, pipelining rule-based filtering with machine learning, and through analysis of features, insight as to which is most critical of the anomaly indicators: it is message delivery time, scalable, and flexible toward static rule-based methods while also responsive toward computationally intensive ML models. Therefore, this work lays the foundation for increasing security in IoT networks. However, future work aims to adapt the threshold values, deep learning techniques like LSTM, and real-time deployment on edge devices, thereby increasing the anomaly detection capabilities.

State-of-the-art anomaly detection for the Matter-enabled Internet of Things networks suffers from limitations such as very high false positives, dependency on static thresholds, and difficulty dealing with imbalanced datasets. Classical rule-based techniques are simple, but they don't adapt much, and individual machine learning algorithms require large sets of labelled training data and aren't robust with real-world variability. Most of the present works do not experiment with hybrid methods that combine rule-based filtering with machine learning for achieving precision. Furthermore, the datasets used are also quite limited in terms of diversity. This work addresses these gaps by proposing a hybrid anomaly detection system that integrates rule-based filtering for fast anomaly flagging with XGBoost for finer classification, achieving a balance between the detection accuracy, computational efficiency, and real-world applicability in Matter-enabled IoT networks.

Future Work

The proposed concurrent anomaly detection system is efficient, but there is plenty of scope for scalability and adaptability improvements. Future work will include the investigation of more sophisticated deep learning techniques such as LSTMs to capture complex time-dependent patterns in telemetry data, allowing for better anomaly detection. The

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:29:57 UTC from IEEE Xplore. Restrictions apply.

system will also be optimized to work at scale for IoT deployments where IoT devices may vary in the order of tens of thousands, ensuring robustness. Real-time dashboards will be added for easier visualization of anomalies for actionable insights to the operators of the networks. Secure communication protocols will also eventually be integrated for the protection of telemetry data as it goes through the various points of transmission. Expanding support beyond Matter to multiple IoT protocols will give versatility and applicability. If these challenges are addressed well, we might be left with a very robust IoT network that's secure and resilient in changing environments.

VIII. References

[1] J. Zhang, R. Gardner and I. Vukotic, "Anomaly detection in wide area network meshes using two machine learning algorithms," *Future Generation Computer Systems,* vol. 93, p. 418–426, 2019.

[2] I. Unwala, Z. Taqvi and J. Lu, "Thread: An iot protocol," in *2018 IEEE Green Technologies Conference (GreenTech)*, 2018.

[3] L. Santos, C. Rabadao and R. Gonçalves, "Intrusion detection systems in Internet of Things: A literature review," in *2018 13th Iberian conference on information systems and technologies (CISTI)*, 2018.

[4] M. U. Qadeer, S. Liaquat, M. A. Shafique and A. R. Kashif, "Implementation of Wireless Mesh Network for IoT based Smart Homes," in *2019 International Symposium on Recent Advances in Electrical Engineering (RAEE)*, 2019.

[5] V. Sharma, K. Sharma, and A. Kumar, "AI and digital twins transforming healthcare IoT," in Proc. 14th International Conf. on Cloud Computing, Data Science & Engineering, Noida, India, Jan. 18-19, 2024.

[6] Y. Liu, K.-F. Tong, X. Qiu, Y. Liu and X. Ding, "Wireless mesh networks in IoT networks," in *2017 International workshop on electromagnetics: applications and student innovation competition*, 2017.

[7] V. R. Joseph, "Optimal ratio for data splitting," *Statistical Analysis and Data Mining: The ASA Data Science Journal,* vol. 15, p. 531–538, 2022.

[8] A. Gaddam, T. Wilkin, M. Angelova and J. Gaddam, "Detecting sensor faults, anomalies and outliers in the internet of things: A survey on the challenges and solutions," Electronics, vol. 9, p. 511, 2020.

[9] U. Goyal, A. Krishna, A. Kumar, and K. Sharma, "Detection and prevention of cyber-attacks on multi-purpose IoT devices using honeypot," in Proc. 2nd Intl. Conf.on Advancement in Electronics & Communication Engineering (AECE 2022), Ghaziabad, India, Jul. 14-15, 2022, pp. 324-329.

[10] A. A. Cook, G. Miserly and Z. Fan, "Anomaly detection for IoT time-series data: A survey," IEEE Internet of Things Journal, vol. 7, p. 6481–6494, 2019.

[11] N. Mishra and S. Pandya, "Internet of things applications, security challenges, attacks, intrusion detection, and future visions: A systematic review," IEEE Access, vol. 9, p. 59353–59377, 2021.

[12] Y. Feng, W. Wang, Y. Weng and H. Zhang, "A replayattack resistant authentication scheme for the internet of things," in 2017 IEEE international conference on computational science and engineering (CSE) and IEEE international conference on embedded and ubiquitous computing (EUC), 2017.

[13] M. Singh, "Matter IoT Telemetry Dataset," GitHub repository, 2024. [Online]. Available: https://github.com/manjits265/Matter-IOT-Telementary-Dataset

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:29:57 UTC from IEEE Xplore. Restrictions apply.

iHDoc: Secure NFC Mobile-based Digital Document Signatures Prototype

Himanshu Sharma
Dept. of Software Engineering,
Delhi Technological University,
Delhi, India
sdothimanshu@gmail.com

Divyashikha Sethia
Dept. of Software Engineering,
Delhi Technological University,
Delhi, India
divyashikha@dtu.ac.in

2025 International Conference on Intelligent Control, Computing and Communications (IC3) | 979-8-3315-2749-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/IC363308.2025.10956421

Abstract—Digital signatures are essential for secure data interchange. However, most document signing solutions depend on cloud-based infrastructures, which may expose users to security vulnerabilities and significant computational requirements. This paper proposes designing and implementing a proximity mobile-based iHDoc digital document signing prototype to sign documents over a Near Field Communication (NFC) tap. It consists of Signer and Signee applications on two different mobile devices that tap to first mutually authenticate each other and then sign the document. iHDoc utilizes iDSign, a previously proposed secure and lightweight design framework based on Identity-based Cryptography (IBC). IBC eliminates the need for certificates by associating public keys directly with user identities, reducing computational and communication overhead and assisting in lightweight mutual authentication over HCE and resource-constrained mobile devices. NFC proximity reduces the risk of man-in-the-middle attacks and ensures the locality of reference. Performance comparison of iHDoc with IBC and RSA-based mutual authentication indicates that the iHDoc prototype performs better with iDSign due to reduced time and storage. It is, hence, suitable for the proximity-based, resource-constrained digital signature of documents.

Keywords—digital signatures, host card emulation (HCE), near field communication (NFC), identity-based cryptography (IBC), mobile devices, lightweight mutual authentication, session key establishment

I. INTRODUCTION

Digital document signing has become essential for secure data exchange in various fields, such as healthcare, legal contracts, and financial transactions [1]. The need for robust digital signature solutions has never been more critical. However, many existing solutions rely on cloud-based infrastructures, which expose users to potential vulnerabilities, high computational requirements and threats such as Man-in-the-Middle (MiTM) attacks. Cloud-based Public Key Infrastructure (PKI) systems, commonly used in popular signing platforms, often necessitate complex certificate management and online connectivity, creating security risks and limiting their usability in resource-constrained, offline, or mobile environments [2], [3]. The growing number of Android mobile devices presents the potential to digitize conventional paper-based processes that require in-person encounters, including government document approvals, legal contracts and delivery confirmations. These situations benefit from proximity-based solutions that guarantee secure, localized communication, facilitating in-person digital signature systems.

Near-field communication (NFC) offers a strong alternative for robust in-person, proximity-based digital communication.

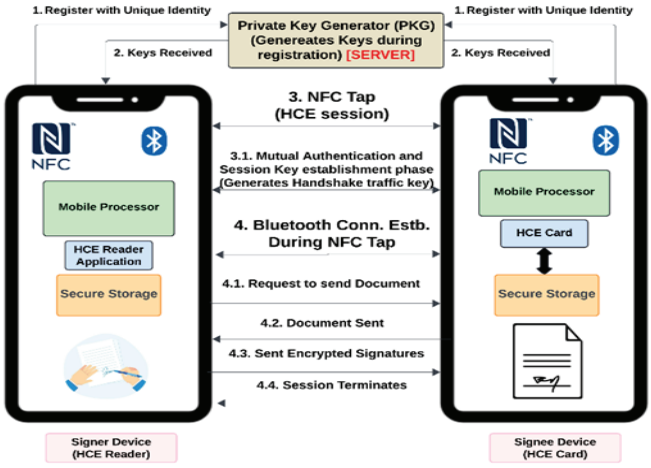


Fig. 1. Block Diagram of iHDoc.

It is a wireless communication technology that allows two devices to transfer data within short distance, typically within a few centimetres [4]. It operates in three modes: reader/writer mode for reading NFC tags, peer-to-peer mode for exchanging data between mobile devices, and Host Card Emulation mode (HCE), where the device function as a contactless smart card. The HCE mode enables secure two-way communication, allowing mobile devices to mimic contactless cards and reader applications. The versatility of HCE has made it a popular choice in applications such as smart health cards [5], payment solutions [6], and secure authentication systems. HCE employs Application Protocol Data Units (APDU) commands [7] based on ISO/IEC 14443 standards [8] for secure communication between the emulated card and reader application. NFC's locality ensures communication only within proximity, greatly reduces the possibility of proximity-based attacks, such as MiTM attacks.

Hence, this paper proposes the design and implementation of a novel iHDoc android mobile prototype application that makes use of NFC's HCE mode [4] and Bluetooth to address the shortcomings of existing solutions. It is a proximity-based digital document signing prototype designed for the in-person digital document signing on Android mobile devices. In iHDoc, the Signer and Signee tap mobile devices over NFC HCE mode to first mutually authenticate each other based on the lightweight iDSign framework [9]. After authentication, the applications establish an automated Bluetooth connection during the NFC tap to address the HCE mode's data transfer limitations and enable faster transfer of larger files. Finally, the document is securely transferred and digitally signed, as illustrated in the block diagram in Fig. 1.

TABLE I. COMPARISON OF EXISTING DIGITAL SIGNATURE SOLUTIONS

979-8-3315-2749-5/25/$31.00 ©2025 IEEE
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:37:46 UTC from IEEE Xplore. Restrictions apply.

Existing Solution	Signature Type	Authentication Method	Internet Required	Encryption Model	Limitations
Adobe Sign [14]	PKI-based digital, Cloud Cert. or tokens	PKI certs, Hardware tokens, Cloud-based tokens	Required	RSA encrypt., High computational demand	High processing, Internet dependent
Visma Sign [15]	Strong ID signatures, cloud servers	External ID Verification (Banking, Mobile ID)	Mandatory for ID verification	Govt. ID verification	Cloud server dependent, Limited offline cap.
DocuSign [16]	E-sign, PKI-based digital	SMS, email, third-party integration	Required for auth. and cert. access	RSA encrypt, High comp. demand	Limited offline, Cloud dependency

A solid secure protocol is crucial for authenticating devices engaged in NFC HCE mode interactions to guarantee that only authorized entities participate in document signing. Conventional Rivest, Shamir, and Adleman (RSA) based authentication methods, such as those used in Transport Layer Security (TLS) and HTTPS protocols [10], necessitate large certificate exchanges which are computationally demanding and inappropriate for resource-constrained mobile devices. The low data transmission capability of HCE complicates the challenges associated with frequent, large-scale exchanges in certificate-based systems. To address these challenges, iHDoc prototype is based on the iDSign [9], a previously proposed framework that leverages Identity-based Cryptography (IBC) [11] along with a modified Lightweight Identity-based Transport Layer Security (iTLS) protocol [12] for lightweight mutual authentication and session key establishment. IBC uses a certificate-free methodology by associating public keys directly with user identities (e.g., email or phone), making it suitable for mobile and resource-constrained devices with reduced computational overhead. Additionally, the Hess Identity-based Signatures (IBS) scheme [13], built on the principles of IBC, is used for the digital signing of documents.

A. Main Contributions: The key contributions are as follows:

- Design and implementation of iDSign Framework [9] with NFC HCE mode [4] and Bluetooth: Demonstrates in-person signing on Android devices using the iDSign framework with low-energy NFC-based HCE for secure proximity authentication, mitigating risks like man-in-the-middle attacks, and Bluetooth for efficient post-authentication data transfer. This approach also has the potential for future remote signing capabilities.
- Comparative Analysis: Performance comparison of the proposed iHDoc prototype with IBC [9] and RSA-based mutual authentication [10] and digital signatures. Results indicate, IBC achieves faster execution, reduced storage, and improved suitability for mobile devices.
- Roadmap: The remaining paper's organization is as follows: Section II contains an overview of related work. Section III describes the system architecture and implementation details. Section IV presents a comparative performance evaluation. In the end, Section V concludes the paper with key findings and suggests future work to extend the system's capabilities to remote signing and improve cross-platform compatibility.

II. Related Work

In recent years, mobile digital signing solutions have become increasingly popular, with Adobe Sign [14], Visma Sign [15], and DocuSign [16] emerging as prominent choices [17]. These platforms largely depend on cloud-based architecture and PKI, which, although offering strong security, impose constraints that may impact usability and security, especially in mobile and offline environments. TABLE I presents the comparison of existing digital signature solutions.

Adobe Sign [14] employs PKI for authentication and document signing, which cloud-based certificates or hardware tokens can implement. Nevertheless, PKI systems frequently necessitate complex certificate management and incur considerable processing requirements, making them less suitable for resource-constrained devices. The dependence on a centralized authority for certificate validation makes Adobe Sign prone to network-based attacks and restricts offline capabilities.

Visma Sign [15] depends on strong ID verification techniques, such as banking or mobile ID, which require constant internet access for authentication. The dependence on external identity providers and centralized verification creates network weaknesses, making it inappropriate for environments with limited connectivity. Although Visma Sign provides secure, regulated authentication, it needs greater flexibility for offline or in-person signing.

DocuSign [16], a PKI-based service, supports many authentication techniques like SMS, email, and third-party connections. However, its dependence on cloud infrastructure limits its use in offline settings and introduces security risks related to remote authentication.

Although these technologies provide reliable document signing for remote and high-security scenarios, they often rely on constant internet access, making them inappropriate for offline usage. In contrast, the proposed iHDoc prototype overcomes these drawbacks by utilizing IBC with NFC HCE mode for tap-based in-person digital signing. Unlike PKI-based solutions, iHDoc's IBC prototype eliminates certificate administration, reducing computational load and simplifying key management for mobile devices. The NFC HCE mode tap requirement minimizes proximity attacks such as MiTM attacks. Moreover, iHDoc's offline signing capability makes it well-suited for environments with limited or no connectivity.

TABLE II. Symbols, Abbreviations, And Operators Description

Notations	Definitions
ID_S	Signer's Identity
ID_D	Signee's Identity
EK_S	Ephemeral Public Key of Signer
EK_D	Ephemeral Public Key of Signee
N_S	Large Random Nonce chosen by Signer
N_D	Large Random Nonce chosen by Signee
HTK	Handshake Traffic Key (Session Key)
EFM	Encrypted Finished Message
P_{Pub}	Master Public Key
P	Generator in IBC

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:37:46 UTC from IEEE Xplore. Restrictions apply.

III. System Architecture And Implementation

The proposed iHDoc system is a functional prototype designed for in-person digitally signing documents on Android mobile devices. This prototype features an intuitive user interface (see Fig. 3.) designed to streamline the signing process, with distinct workflows for the Signer and Signee roles. The system utilizes the notations outlined in Table II to define the cryptographic operations and interactions between devices. A demonstration of the prototype in action is available in the Technical Report [18]. iHDoc builds on the theoretical foundation of the iDSign framework [9]. This section breaks down the architecture and implementation specifics, demonstrating how each component facilitates secure and efficient in-person digital signing.

A. System Components

The proposed iHDoc prototype comprises the following components:

- Private Key Generator (PKG): It is a server component that generates and distributes identity-based cryptographic keys.
- Signer Device: This device preserves the confidential signing credentials and commences the signing procedure. It functions as the NFC reader in Host Card Emulation mode.
- Signee Device: This device requests document signing and receives the signed document. It operates in NFC HCE mode to serve as a virtual card for data transmission.

B. Workflow Overview

The iHDoc workflow adopts a structured sequence from device registration to completing a digitally signed document. Fig. 2. details the prototype, which utilizes multiple libraries and technologies that work in unison to ensure secure, efficient document signing. The key steps are as follows:

1) Device Registration with PKG: Both the Signer and Signee devices register their unique identities (such as email or phone number) with the PKG. It uses these identities to generate corresponding public and private keys, leveraging IBC.

2) Keys Reception: The devices receive their respective keys (private and public) from the PKG.

3) Mutual Authentication and Session Key Establishment (NFC HCE Session) (For a detailed explanation, refer to iDSign Framework [9]):

a) During this tap-based interaction, both devices authenticate each other using iDSign's mutual authentication [9] and exchange ephemeral public keys (EKS, EKD), unique identifiers (IDS, IDD), nonces (NS, ND), and PKG details to establish trust.

b) Once registered, both devices initiate mutual authentication by tapping the devices, utilizing NFC HCE mode and IBC (as shown in Fig. 3 -B).

c) Both devices independently compute a shared secret using bilinear pairing curve operations. The shared secret is then processed through an HMAC-based key derivation function (HKDF) [19] to derive the session key HTK.

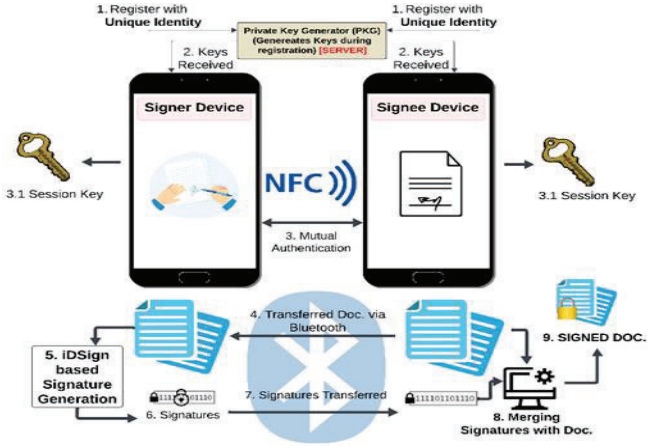


Fig. 2. iHDoc Prototype.

d) The Signee sends an encrypted message (EFMD) to the Signer, containing a random string (RS), a timestamp, and the incremented nonce (NS +1). The Signer decrypts EFMD using the session key, validates the parameters and responds with its encrypted message (EFMS) containing similar parameters. Successful validation completes mutual authentication, and the session key secures subsequent communication.

e) In addition to authentication, the Signer's Bluetooth name is shared with the Signee during the NFC tap, facilitating automatic Bluetooth pairing.

4) Document Transfer via Bluetooth: The Signee establish an automated Bluetooth connection and sends the document to the Signer (as shown in Fig. 3-C).

5) Digital Signature Generation: On receiving the document, the Signer uses the Hess IBS [13] scheme to generate the digital signature by utilizing the document's hash, the Signer's private key, and cryptographic parameters managed through the jPBC library [20], which facilitates pairing-based operations for IBC.

6) Signer then encrypts the signatures with the previously generated session key to guarantee secure transmission.

7) Signature Transmission: The Signer transmits the encrypted signature to the Signee device over the secure Bluetooth channel.

8) Signature Integration: The Signee decrypts signatures using the session key, incorporates it into the original document and completes the digital signing process. As shown in Fig. 3-C, the digital signature includes identifying information and a timestamp for authenticity within the document. The user can then choose to continue signing more documents or end the session (as shown in Fig. 3-C).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:37:46 UTC from IEEE Xplore. Restrictions apply.

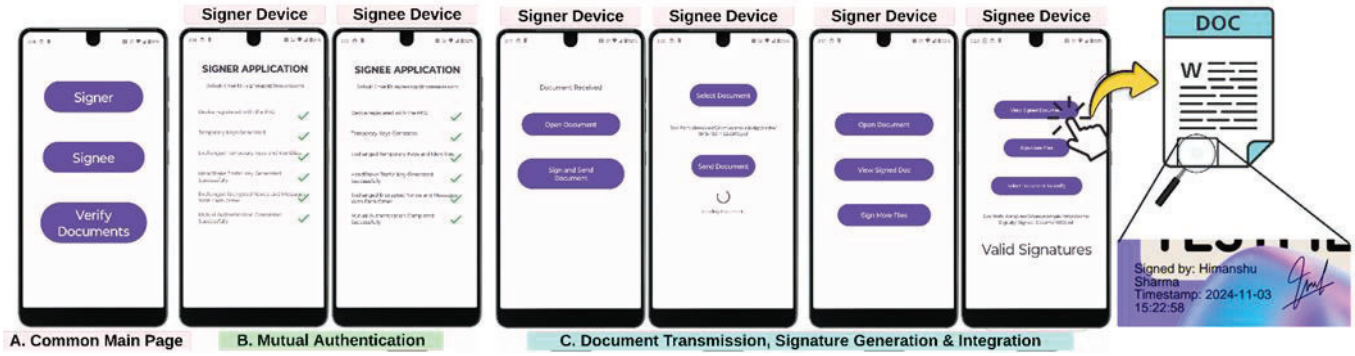


Fig. 3. Screenshots of iHDoc User Interface.

9) Signed Document Storage and Verification: The final signed document is stored locally and can be verified by any device using the Signer's public key.

C. Technologies and Libraries Used

This section breaks down the technologies and libraries incorporated into the iHDoc prototype, clarifying their functions in secure digital document signing on Android devices.

1) Android NFC with HCE: NFC technology in HCE mode [4] enables proximity, tap-based interactions between Android devices, simulating a physical smart card. This setup ensures robust session initiation by reducing vulnerability to man-in-the-middle attacks.

2) Bluetooth for Data Transfer: Bluetooth is used for effective data transfer following NFC-based authentication. This technique makes use of Bluetooth's stability and bandwidth for transferring larger files, such as PDF documents, post-authentication.

3) jPBC Library [20]: The Java Pairing-Based Cryptography (jPBC) library is an crucial element of the iHDoc system, facilitating all cryptographic functions based on IBC. The jPBC library enables bilinear pairing operations crucial for IBC, facilitating secure connections and digital signatures based on user identities instead of conventional public-key certificates. This pairing-based methodology, essential to iHDoc, facilitates the framework's lightweight yet safe cryptographic requirements, from mutual authentication to document signing. The library reduces the computational and storage demands commonly linked to certificate-based systems, enabling iHDoc to offer both security and efficiency, making it a perfect option for mobile applications.

4) Bouncy Castle Library [21]: Bouncy Castle provides essential cryptographic operations, including the HKDF (HMAC-based Key Derivation Function), which generates the symmetric session key (HTK) during the mutual authentication phase.

5) iText Library [22]: iText provides fast and seamless PDF file manipulation within the iHDoc system, offering tools for incorporating and verifying digital signatures.

6) Spring Framework for Local Server Setup [23]: A local server built with the Spring framework operates as the PKG, creating and distributing keys based on identity. When devices send identity information (e.g., email or phone number) to the server, it provides the corresponding identity-based cryptographic key pairs, facilitating iHDoc's IBC prototype.

IV. Performance Evaluation

This section evaluates the performance of the proposed iHDoc prototype, comparing its efficiency when using IBC versus RSA certificates. Using the same iHDoc prototype for both configurations enables an accurate assessment of the advantages of IBC in proximity-based mobile environments. Key metrics include timing and storage requirements, emphasizing the lightweight nature of IBC for mobile use. Testing and evaluation are conducted on devices including the Samsung Galaxy Tab Active 2, Motorola G84, OnePlus Nord 2T, and Poco X6 Pro, covering diverse hardware combinations.

A. Timing Comparison of Core Components

TABLE III compares the approximate time taken for different components in iHDoc with IBC [9] and RSA [24] when processing a document of 1MB size. The components analyzed include Mutual Authentication, Bluetooth File Transfer, Signature Generation, and Signature Integration.

- Mutual Authentication: iHDoc with IBC achieves a faster mutual authentication time (0.479 sec) than RSA (0.702 sec) due to the absence of certificate verification in IBC, thus reducing computational overhead.
- Bluetooth File Transfer: The time for Bluetooth transfer remains similar in both configurations, as Bluetooth communication is independent of the cryptographic method.
- Signature Generation: iHDoc with IBC also performs faster signature generation (0.207 sec) compared to RSA configuration (0.232 sec), reflecting the

TABLE III. Approximate Timing Comparison of Diff. Components for Document of Size 10 MB

Scheme	Mutual Authentication (sec)	Bluetooth File Trans. (sec)	Signature Generation (sec)	Signature Integration (sec)	Total (sec)
IBC [9]	0.479	68.32	0.273	0.521	69.59
RSA [24]	0.778	68.37	0.298	0.530	69.97

TABLE IV. APPROXIMATE CREDENTIALS STORAGE REQUIREMENTS

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:37:46 UTC from IEEE Xplore. Restrictions apply.

Scheme	Public Key Size (bytes)	Private Key Size (bytes)	Certificate Size (bytes)	Symmetric Session Key Size (bytes)	P_{Pub} (bytes)	P (bytes)	Total (bytes)
IBC [9]	**128**	**20**	**-**	**128**	**128**	**128**	532
RSA [24]	**256**	**256**	**1024**	**128**	**-**	**-**	1664

efficiency of the Hess IBS scheme used in iHDoc with IBC configuration.

- Signature Integration: The time required for signature integration is slightly lower in iHDoc with IBC (0.421 sec) compared to RSA (0.428 sec), benefiting from the streamlined IBC approach.

The results in TABLE III demonstrate that IBC offers reduced processing times in the mutual authentication and signature generation phases, emphasizing the advantages of lightweight cryptography for digitally signing documents on mobile devices.

B. Storage Requirements

TABLE IV compares storage requirements for credentials and cryptographic keys in iHDoc leveraging IBC and RSA certificate-based keys. The RSA configuration requires more storage due to including public/private key pairs and certificates, whereas the IBC configuration significantly reduces storage requirements by eliminating certificates.

- RSA Key and Certificate Storage: In RSA configuration, each device requires storage for RSA keys and certificates and requires approximately 1664 bytes for storage.
- IBC Key Storage: IBC configuration eliminates certificates, reducing storage needs to only 532 bytes for storing public and private keys, the master public key (PPub), and the generator (P). Hence, IBC is more suitable for mobile devices with limited storage capacity.

As shown in Table IV, the total storage requirement for iHDoc with IBC [9] is 532 bytes, which is significantly lower than the 1664 bytes required for the RSA [24]. This reduction in storage requirements is critical for mobile devices, where storage space is often limited.

C. Document Size vs. Signing Time

Fig. 4. illustrates the relationship between document size and total signing time in iHDoc with IBC. As document size increases, the time taken for signing also increases, largely due to Bluetooth transfer time. For smaller documents (e.g., 0.1 MB), the signing time is relatively low (around 14 seconds), but for larger documents (e.g., 60 MB), the signing time increases substantially (up to 404 seconds). This graph highlights iHDoc's scalability and the impact of document size on overall performance.

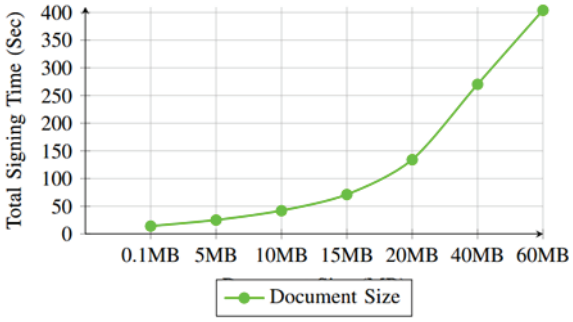


Fig. 4. Relationship between Total Signing Time and Doc. Size in iHDoc with IBC [9].

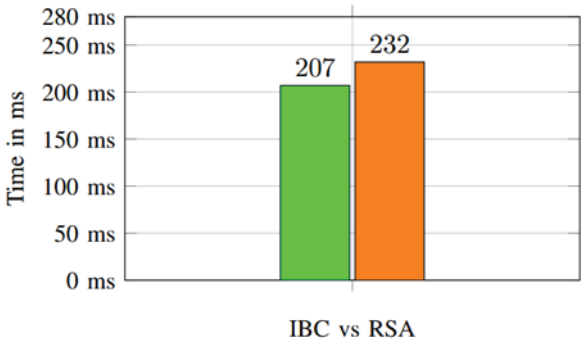


Fig. 5. Comparison of Sign. Generation Time for iHDoc with IBC [9] and RSA [24]

D. Signature Generation Time Comparison

Fig. 5. shows the comparison of signature generation time between iHDoc with IBC and RSA. The signature generation time in IBC is approximately 207 ms, while RSA requires 232 ms. This performance difference highlights the efficiency of the iDSign's IBC-based signature scheme [13] in iHDoc, making it well-suited for mobile environments with limited computational resources.

The performance evaluation shows that iHDoc with IBC outperforms RSA configuration in terms of time and storage efficiency. Mutual authentication and signature generation phases are faster with IBC due to the lightweight cryptographic approach, while storage requirements are lower since it does not use large certificates and keys. Additionally, the impact of document size on signing time demonstrates the scalability of iHDoc. In contrast, the overall reduced signature generation time illustrates the effectiveness of the Hess IBS scheme in mobile applications.

V. Conclusion And Future Work

This paper presents iHDoc, a secure and lightweight prototype based on the iDSign framework [9] for in-person digital document signing on Android smartphones. Utilizing NFC HCE mode [4] for proximity, tap-based authentication, iHDoc mitigates proximity-based threats, such as man-in-the-middle attacks, enhancing security in mobile settings. IBC enables iHDoc to function without digital certificates, reducing computing requirements and making it suitable for resource-limited devices. In contrast to conventional PKI-based systems, iHDoc facilitates offline signing, enhancing its applicability in regions with restricted connectivity. Performance evaluations demonstrate iHDoc's efficiency, with faster execution times and reduced storage requirements compared to RSA-based configuration, highlighting its capability in secure, lightweight mobile applications.

Future work will extend iHDoc for remote signing while maintaining its security advantages, improving platform compatibility, and exploring hybrid signing modes. Furthermore, incorporating a Secure Element [5] will improve credential security and optimize cryptographic functions. iHDoc offers a versatile and resilient solution for safe mobile

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:37:46 UTC from IEEE Xplore. Restrictions apply.

document signing by integrating proximity security with strong offline capabilities.

REFERENCES

[1] P. Kaur, N. Arora, "A Comprehensive Study of Cryptography and Digital Signature," International Journal of Science, Engineering and Computer Technology, vol. 5, no. 1, pp. 1. Indian Association of Health, Research and Welfare (2015).

[2] M. R. Ramadhan, S. Mandala, F. Yulianto, "Analysis and Implementation of Digital Signature Algorithm in PDF Document," In: 2023 11th International Conference on Information and Communication Technology (ICoICT), pp. 11–16. IEEE (2023). https://doi.org/10.1109/ICoICT58202.2023.10262708

[3] G. Saha, "DSign Digital Signature System for Paperless Operation," In: 2017 International Conference on Communication and Signal Processing (ICCSP), pp. 0324–0328. IEEE (2017). https://doi.org/10.1109/ICCSP.2017.8286370

[4] S. C. Alliance, "Host Card Emulation (HCE) 101," A Smart Card Alliance Mobile and NFC Council White Paper (2014).

[5] D. Sethia, D. Gupta, and H. Saran, "Smart health record management with secure NFC-enabled mobile devices," *Smart Health*, vol. 13, p.100063, 2019.

[6] S. Micallef and K. Markantonakis, "Mobile payments using host card emulation with NFC: security aspects and limitations," Royal Holloway Information Security Thesis Series, 2018.

[7] Application Protocol Data Unit (APDU): https://www.cardlogix.com/glossary/apdu-application-protocol-data-unit-smart-card/ (Last accessed: August 26, 2024)

ISO/IEC 14443: https://developer.android.com/develop/connectivity /nfc /hce (Last accessed: August 25, 2024).

[8] ISO/IEC 14443: ttps://developer.android.com/develop/connectivity/nfc/hce Last accessed: August 25, 2024.

[9] H. Sharma, D. Sethia, "iDSign: Secure Lightweight Digital Document Signatures Framework with Mutual Authentication and Identity-Based Cryptography," Accepted in: International Conference on Computational Intelligence and Communication Systems (ICCIS), 2024, in press.

[10] E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.3," (2018).

[11] A. Shamir: Identity-Based Cryptosystems and Signature Schemes. In: G. R. Blakley, D. Chaum (eds) Advances in Cryptology. CRYPTO 1984. Lecture Notes in Computer Science, vol. 196, pp. 47–53. Springer, Berlin, Heidelberg (1985). https://doi.org/10.1007/3-540-39568-7 5.

[12] P. Li, J. Su, X. Wang, "iTLS: Lightweight Transport-Layer Security Protocol for IoT With Minimal Latency and Perfect Forward Secrecy," IEEE Internet of Things Journal 7(8), 6828–6841 (2020). https://doi.org/10.1109/JIOT.2020.2988126

[13] F. Hess, "Efficient identity based signature schemes based on pairings," In: Selected Areas in Cryptography: 9th Annual International Workshop, SAC 2002, St. John's, Newfoundland, Canada, August 15–16, 2002 Revised Papers, vol. 9, pp. 310–324. Springer, Berlin, Heidelberg (2003). https://doi.org/10.1007/3-540-36492-7_20

[14] Adobe. "Digital Signatures with Adobe Sign," https://helpx.adobe.com/in/sign/using/digital-signatures.html, Last accessed: November 7, 2024.

[15] Visma. "Visma Sign." https://vismasign.fi/english/, Last accessed: November 7, 2024.

[16] DocuSign. "Electronic Signature." https://www.docusign.com/ products/electronic-signature, Last accessed: November 7, 2024.

[17] T. Puistosalo, "Introduction of Digital Signing to Case Company," (2023).

[18] Technical Report on iHDoc Digital Signing Prototype. Available at: https://sites.google.com/view/sdarldtu/ihdoc-document-signing?authuser=0 (Last accessed: December 07, 2024).

[19] H. Krawczyk, P. Eronen, "HMAC-based extract-and-expand key derivation function (HKDF)," Available online: https://www.rfc-editor.org/rfc/rfc5869.html Last accessed: September 20, 2024.

[20] G. Dia, "Java Pairing-Based Cryptography (JPBC)," http://gas.dia.unisa.it/projects/jpbc/, accessed November 7, 2024.

[21] BouncyCastle. "The BouncyCastle Crypto APIs." https://www.bouncycastle.org/, Last accessed: November 7, 2024.

[22] iTextPDF. Available at: https://itextpdf.com. Last accessed: November 4, 2024.

[23] Spring. "Spring Framework." https://spring.io/, Last accessed: November 7, 2024.

[24] H. Jaiswal, R. Mishra, H. Sharma, and D. Sethia, "SecureXfer NFC HCE Mobile Framework using RSA for Secure Digital Transfer and Signing," under submission to the IEEE International Conference on Next Generation Information System Engineering 2025. [Online]. Available: https://ngise.org/ unpublished.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:37:46 UTC from IEEE Xplore. Restrictions apply.

Incurred Changes in Water and Fat-Soluble Vitamin Profile of Alternative Ingredient Spirodela Polyrhiza Based Partial Fishmeal Substituted Aquafeed: Effect of Long-Duration Variable Temperature Storage

Parul Puri[1,2], Neha Tiwari[2], Vaibhav Puri[3], Ram Singh[4], Jai Gopal Sharma[2], Rina Chakrabarti[5]

[1]*Department of Zoology, Maitreyi College, University of Delhi, India*
[2]*Department of Biotechnology, Delhi Technological University, India*
[3]*Department of Economics, Sri Guru Gobind Singh College of Commerce, University of Delhi, India*
[4]*Department of Applied Chemistry, Delhi Technological University, India*
[5]*Aqua Research Laboratory, Department of Zoology, University of Delhi, India*
Email: parul_acemail11@rediff.com

Abstract: The study aims to evaluate effects of storage duration and storage temperature on changes in vitamin profile of formulated aquafeeds.Two aquafeeds are formulated one as extrusion processed fishmeal replacement diet comprising aquatic macrophyte greater duckweed(Spirodela polyrhiza) as D1, other as non-substituted fishmeal diet from pelettization D2. Diets are stored for sixmonths (180days) under four storage temperature conditions; -20o, 4o,ambient and 45oC namely T1,T2,T3 and T4, respectively. Changes in water-soluble and fat-solublevitaminprofile is assessed bimonthly up to sixmonth storage (at 0,60,120,180 days) for both the diets. Both diets at all storage temperatures,show substantial loss in water- miscible andfat-miscible vitamins namely A, E, K, B2, B12, C, 60-day further.Noteworthily in D2, intense diminutions in vitaminDcontent was observable during four month storage at T4 condition, while temperature, duration and interaction effects on vitamin D were significant ($P \leq 0.05$)for D1.Thiamin retentions in D2 are influenced by interference from antinutrients; whileco-elutants affect B6 determinations in D2. Evidently, incurred loss of most vitamins were highest at end of storage underT4 conditions.The paper highlights effect of temperature, duration on dietary storage profile of vitamins as essential nutrients,affecting quality and shelf-life of stored processed feeds suggesting at best utilizationregimes during storage.

Keywords: aquafeed, storage, temperature, vitamins, antinutrients, bioavailability.

1. Introduction

Fish is a valued source of quality dietary protein with reasonably significant extent of most indispensable amino acids, omega-n3 essential fatty acids - eicosapentaenoic, docosahexaenoic (EPA and DHA), minerals and vitamins. Digestibility and bioavailability of proteins, minerals and vitamins obtained from fish foods is higher compared to any plant-based food product. Fish is growingly becoming global staple; in fulfilling food requirement of world population rural as well as urban. Feed is the foremost important input for fish welfare and health in sustainable intensive aquaculture productions[1]. Demanding pursuit for alternative sources of protein and oil; wholly or partly substituting fish meal (FM) and fish oil (FO) in aquafeeds [2] accord to FAO recommendations on lessening their consumption from extract fishing, as well to meet sustainable UN goals for fisheries and aquaculture. In this direction much focus has shifted in finding and use of non-conventional alternative feed source towards aquaculture growth and developments [3,4]. To this effect potential of aquatic origin macrophytes such as duckweeds (belonging to family Lemnaceae) have been explored as alternative feed constituent. Spirodela polyrhiza and duckweeds at large offer abundant nutritional profiles significantly comparable to animal feed [5,6] with

their partial inclusion in fish diets known to appreciably augment growth responses as well as nutritional attributes of fish [7].

Fish require vitamins in trace quantities as vital organic substances for important physiological functions. Vitamins are classified as water and lipid-miscible forms due to their miscibility in aqueous environment. Water-soluble vitamins comprise vit. B-group including in them B1, 2, 3, 5, 6, 7, 9,12 (namely thiamin, riboflavin, niacin, pantothenic acid, pyridoxine, biotin, folate, cobalamin) and ascorbate (vit. C). Vitamins (vit.) function mainly as coenzymes contributing to metabolic processes (B1, B2); in red blood cell formation, fatty acid- and amino-acid metabolism (B6, B12), amino acid interconversions (B6); bone formation, collagen synthesis and immunoregulatory functions (vit. C) [8]. Due to their hydrophilic property these vitamins are miscible in aqueous solutions. Vit. C is the most labile water-soluble vitamin, known for its antioxidant attributes. Vit.C is a co-substrate for enzymes involved in the biosynthesis carnitine, and neurotransmission. Fish like humans lack capability of vit. C biosynthesis and can fulfill it through dietary supplementation. Dietary inclusions of vit. C is found to improve growth, hematological parameters, carcass nutrients and feed conversion ratio in Labeo rohita[9]. Fat-miscible vit. A, D (calciferols), E (tocopherol), K are bound to lipophilic fraction in feeds and are transported, as well absorbed, in similar fashion as fats [10]. These vitamins have antioxidant properties (vit.A,E), hematological (vit. K) and immunostimulatory functions (A,D) hence contribute to fish and consumer health.Vit. A has essential function in vision, embryonic development, growth, reproduction, and differentiation [11]. Vit. E can prevent fragility of erythrocytes and due to its anti-oxidant mechanism prevent free radical oxidation of dietary (PUFA) and membrane lipids [12]. Loss of vitamins from diet is evident during processing, storage and due to leaching in aquatic conditions.Vit A, E and C due to their free radical sequestration potential have strong antioxidant properties. Thus, are prone to similar oxidative effects during storage. Storage loss of vit. A is reported due to oxidative influences [13]. Cessation of vitamins during prolonged storage (>180 days) is related to lipidic rancidity ensuing peroxidation[14]. Storage loss of vitamins require over compensations as supplementations to balance available levels from diet.

Storage feeds are prone to oxidative loss of lipids, proteins and vitamins as essential dietary nutrients. Autooxidation of lipids, proteins in feed can also decrease abundance of vitamins as natural antioxidants in feed ingredients [15]. Largely, diminution of feed vitamins is accentuated due to high temperature, humidity, extreme pH, light exposure, presence of elements and lipidic free radicals. Water soluble forms of vitamin are prone to storage effects showing losses due thermal, processing, storage and feed matrix properties. Extrusion processing of feed can affect nutritional quality of diet at various degrees.Diet composition along with extruder parameters including temperature, shear, screw type, pressure, and size of die have great influence on physical and nutritional quality of extruded feed [16]. Depending on extrusion conditions heat susceptible vitamins can be lost, with protein-sugar reactions at high extrusion temperatures deteriorating feed nutrient profiles [17,18] evaluated vitamin stability due to extrusion, and pelleting procedures at three months storage, reporting extrusion loss of thiamine (88%) compared to pelleting loss (60- 96%). Even coated forms of ascorbate are largely lost at high rates to extrusion. Most susceptible vitamins to extrusion were thiamine (due to high shear rates of extrusion), vit. A, C and folate. [19] evaluated impact of extrusion temperatures and encapsulation procedure on vitamin stability to conclude high stability of micro-encapsulated vitamin compared to un-encapsulated forms in feed. Since many parameters need to be optimized for extrusion processing of feeds, the technique requires precise controlled operating conditions to restore maximal nutrient benefits from feedsonsubsequentutilisation.

Nutrient quality and safety are two most important attributes of healthy diet [4].There is a larger scope towards increase in fish production as well as their nutrient value for human consumption by judicious improvement in feed formulations and management of processed feed during storage, handlings; in farming environment. Storage help administer timely requirement of feed in aquaculture; nevertheless, improper storage or storage handlings can intensify feed spoilage [20]. Proper storage of feed is imperative for safeguard of feed quality, ensuring overall fish health and consumer beneficence. Knowledge of storage sequels help assess losses during prolonged feed utilizations, suggesting harmonization of stored processed feeds to maximise nutritional benefit for the fish and incidentally the consumers.The present study explores feed storage changes in composition of water- soluble, fat-soluble vitamins essential to fish nutrition in influencing overall feed quality during storage,at longduration variable temperature conditions. This work provides comparison of plant and fishmeal based formulated feeds to ascertain their response to the storage variables.

2. Materials and Methods

2.1Formulation of feed

Aquafeeds were formulated using Interactive Fish Feed Designer (IFFD) software version 2 [21] as per composition given in table 1. Extrusion based compounded diet (D1), was formulated comprising fishmeal as major protein source with substitution of S. polyrhiza and wheat flour as minor protein sources, plant-oil sunflower oil was substituted for fish oil. Dry ingredients were well mixed with subsequent addition of oil and sifted through fine mesh, subjecting to extrusion pelleting with addition of extra pure milli-Q water (Synergy ultrapure, Millipore, Germany). Two-kilogram sinking pellet extruded feed was prepared based on twin-screw extrusion technique (BTPL lab model twin-screw food extruder, Kolkata, India) using 2.5 mm die as per set conditions with extruder rpm 214, cutter rpm 545, feeder rpm 10, extruder torque 13.9, heater1 (65°C) and 2 (70°C), temperature of final mass 101°C. Extruded feed was dried in feed drier (Hicon, India) for 4h at 35oC, then kept at room conditions(1h) and packaged for storage in sealable bags. For D2 (non-extruded diet)hot water (60oC) was added to well kneaded feed ingredients,subjecting to pelletization (Kent, India) through 2mm die. Prepared diet pellets were driedand packed for storage as D1.

2.2 Feed storage

Three replicates (n=3) per temperature condition were kept for storage. Diet samples atT1,T2 and T3 conditions were stored in resealabletransparent plastic bags whereasT4 sampleswere kept in capped transparent glass bottles (Borosil, India) duringstorage. Daily temperature record was maintained towards T3 storage using thermo-hygrometer(ThermoPro TP53, US),table 2.Mean (± standard error, SE)temperature records for D1 ranged between 18.5oC (±3.25)to 29.5oC(±4.15) and D2 between 17.7oC (±1.90)to28.45oC (±2.55). Replicates (n=3) for four temperature conditions under study were drawn bimonthly (at 0, 60th, 120th, 180th day) during each assessment.

2.3 Preparation of feed sample

As per standard method 950.02, samples were milledto fine particles and sieved (1mm) to obtain fine powder [22]. Powdered diet was thoroughly mixed beforeanalysis of vitamin profile. Each assay was performed in replicatesfor calculation of mean for the evaluation of results.

2.4 Vitamin analysis

2.4.1Chemicals and Reagents

Vitamin standards, retinyl acetate (vit. A) (No.46958, Sigma-Aldrich Chemicals private limited, India), thiamine hydrochloride (B1), riboflavin (B2), pyridoxine hydrochloride (B6), cyanocobalamin (vit. B12) and L-Ascorbic acid (vit. C) (No. 47858, 47861, 47862, 47863, 47869 respectively), cholecalciferol (vit.D3), ergocalciferol (vit. D2), vit. K1 (No. 95271 Bio-xtra sum of isomers, HPLC grade) at high purity >98% were purchased from Sigma-Aldrich, USA. E-acetate was obtained from Himedia laboratories private limited, India at purity >97%. Taka-diastase, source Aspergillus oryzae (No. 86247) was obtained from Sigma-Aldrich (USA). n-butanol, chloroform (SRL, Sisco Research Laboratories, India), acetone, sodium acetate (extrapure AR SRL, Sisco Research Laboratories, India), sulphuric acid (Merck, India), methanol, acetonitrile (HPLC grade, Merck, India), metaphosphoric acid (No.239275, Sigma-Aldrich Chemie GmbH, Steinheim, Germany) were also procured. Standard solutions were prepared using extrapure Milli-Q water at pH7.

2.4.2 Sample preparation for vitamins

Water-soluble vit. B1, B2, B6, B12 and C, were assayed according to the method of [23]. Extraction was performed in triplicates. For vit. B1, B2, B6 and B12 estimation diet sample(2g) was added with 25mL H_2SO_4 (0.1N) solution in a conical flask. Sample solution was incubated for 30min at 121oC in oven. Contents were cooled at room temperature and using 2.5M sodium acetate, pH was adjusted to 4.5. Then, 50mg Taka-diastase enzyme powder was added to the sample. After gentle vortex for 1min, preparation was stored in oven at 35oC for 18h. The digested mixture was filtered through Whatman filter (No.4). Filtrate was diluted with 50mL Milli-Q ultrapure water. Extract was re-filtered through 0.45μm micropore syringe filter, into HPLC vial. For the estimation of vit. C,10g diet samplewas blended and homogenized with 25mL extracting solution comprising 0.3M metaphosphoric acid and 1.4M acetic acid. After 15min centrifugation at 10,640g (fixed-angle rotor 12181 Sigma 2-16K, Germany)mixture was filtered using Whatman filter (No. 4). Fat-soluble vitamins (A, D, E and K) were estimated following the method of [24]. Briefly,in a falcon tube 0.5g sample was mixed with 8mL chloroform and acetone mixture in the ratio 70:30. Nitrogen (N_2) gas was flushed through the sample and tube was tightly capped. Contents were vortexed for 1min and tube was left to stand for 5min then

vortexed again for 30s. Mixture was centrifuged at 1702g (fixed-angle rotor 12181 Sigma 2-16K, Germany), 25oC for 5min. After centrifugation, 1mL supernatant was transferred for drying in aN2 evaporator. Completely dried extract was dissolved in 1mL n-butanol. Reconstituted extract was filtered using 0.4□m syringe filter and kept in HPLC vial for further analysis.

2.4.3 Determination and quantification

Stock solutions of vitamin standards 1.0mgmL-1 were prepared. Standard calibration was performed to obtain linearity at concentrations 0.25, 0.5 and 1.0mgmL-1. Chromatographic analysis of vitamins was conducted by ultra-high pressure liquid chromatography (UHPLC Thermo Fisher, model Dionex Ultimate 3000) with stationary phase C18 column (3µm,150x4.6mm). 20□L of sample extract was injected into the HPLC system. Simultaneous separation for vit. B1, B2, B6, B12 was achieved at a flow rate of 0.5mLmin-1, wavelength 270nm using MeOH : 0.023M H3PO4 (33:67 ratio, pH 3.54) as mobile phase. For estimation of vit. C, flow rate of 1mLmin, wavelength 254 nm and acetonitrile (ACN, 100%) as mobile phase solvent was deployed. Fat-soluble vit. A,D,E,K were determined at flow rate 1.8mLmin-1, UV absorbance 290nm, 100% MeOH as mobile phase.

2.5Statistical analysis

For both fat and water-soluble vitamins, data is represented as mean□SE. Analysis of variance (two-way ANOVA) consideringmajor effects of temperature, storage duration and interaction on vitamin contentof formulated feedsis performed using SPSS statistics (IBM, version 25)[25]. Statistical differencesare reported by honestly significant difference (HSD). Level of significance $P \leq 0.05$based on Tukey's post-hoc analysis is considered[26].

3. Results and Discussion

3.1Water-soluble vitamin storage changes

3.1.1 Vitamin C

Among all storage temperatures, vit. C retention at end of 180 day storage duration is highest at T1 condition, (Fig.1a,1b), in order T1 (89% D1, 86.13% D2) > T2 (85.3% D1, 81.02% D2) >T3 (51.38% D1, 79.56% D2) >T4 (39.45% D1, 51.82% D2) retentions. Significant difference ($P \leq 0.05$) due to temperature, duration and interaction effects was found for vit. C values in D1. For D2, storage effects of duration were highly significant. Neither temperature nor interaction had significant effect on changes in vit. C for D2. Although D1 and D2 depict gradual loss of vit. C over low temperature storages (T1,T2 and T3); losses are evidently higher at T4 due to intense temperature effects. This is in agreement with works of [27, 28], reporting decreased vit. C levels in formulated feeds due to extrusion, pelletization processing based moisture and heat addition, causing rapid losses at extreme conditions. Vit. C in aquafeeds is susceptible to processing, storage losses [29]. Additionally, oxidative, non-oxidative and leaching losses of vit. C in fish feeds are reported. In fishes, dietary deficiency of vit. C can impair collagen synthesis leading to scurvy, lordosis, scoliosis, impaired wound healing, lethargy and anemia symptoms [29, 30]. [31], recommends 67.17 mgkg-1 of ascorbate in grass carp feed for optimum growth parameters of juveniles. Reducing potential of vit. C can be utilized by its supplementation to diets containing high levels of polyunsaturated fats thereby, preventing oxidation of feed lipids [32]. Vit. C potentiates antioxidant effects of vit. E in protecting lipids against peroxidative distress [33].

VITAMIN C

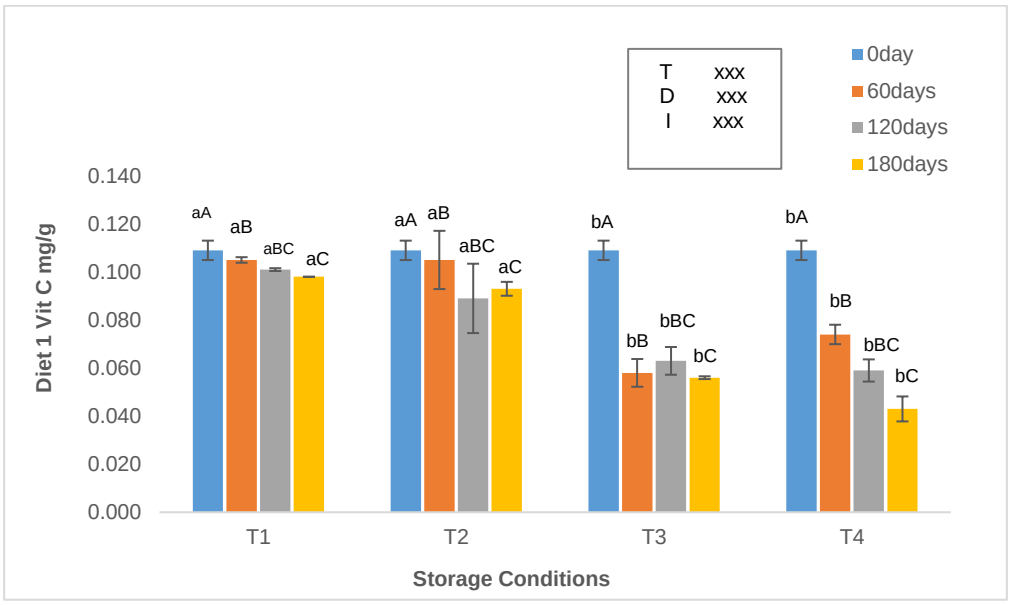


Fig.1a

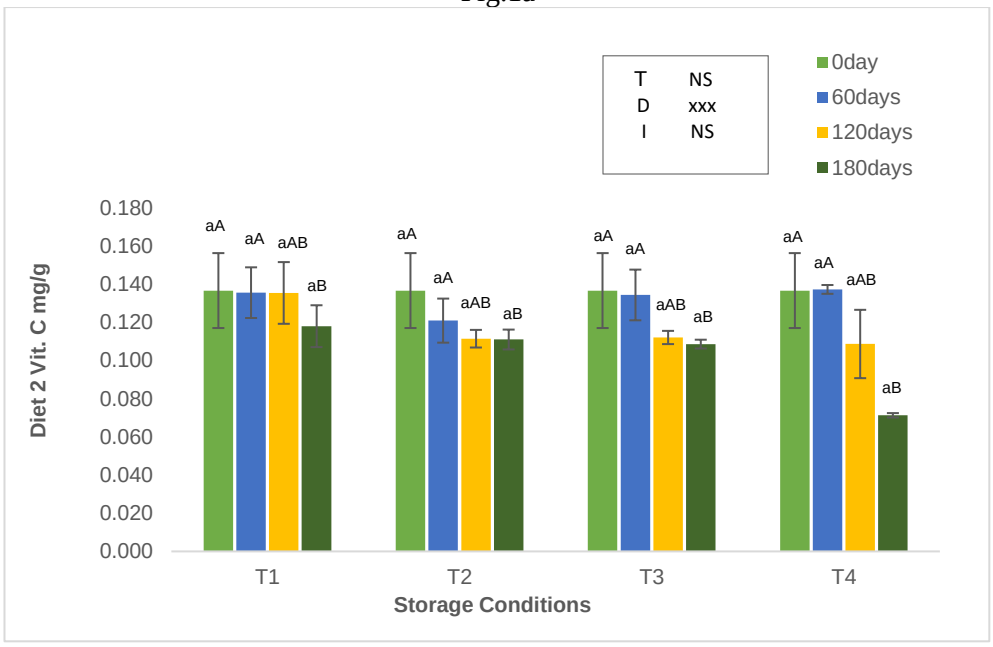


Fig.1b

3.1.2 Vitamin B1 (thiamin)

Gradual decrease in vit. B1 is evident bi-monthly at all storage temperatures for D1, (fig.1c). Incurred losses being higher at high temperatures T4 (55.81% vit. B1 present) >T3 (61.88% present) compared to higher retentions at low temperatures T1 (66.06%), T2 (66.17%) at end of storage duration. Thiamine is thermolabile and susceptible to losses at high temperature processing (extrusion) as well as low moisture conditions. Extrusion loss of thiamin up to 50% is reported by [18]. As per [8] thiamin retentions ranged between 60-80% in extruded feeds stored for 3months at ambient temperatures. Although, non-extruded D2, show initial increase in thiamin concentration up to 120 days, with decreases at later duration between fourth to sixth month across storage temperatures, (fig.1d). These findings are in good agreement to study based on frozen foods [34], reporting initial increase in thiamin content during storage, processing; followed by storage decrease of thiamin, yet no proposed mechanism could be explained. We believe that this effect may be due to interaction effects of thiamin with dietary antinutrients mainly phytochemicals such as tannins, which would have overestimated vit. B1 peaks in chromatographic run. Additional work by [35] provide support to this hypothesis on thiamin-tannin interaction indicating formation of adducts (weak bonds) between thiamin and oxidized tannin to be co-

eluted at same retention values during thin-layer chromatographic analysis. Compared to extruded diets, non-extruded diets are prone to interaction influence of polyphenols. Extrusion processing can reduce tannins and their interference on bioavailability of dietary nutrients (proteins, lipids, vitamins and minerals). Duration effects on storage thiamin are significant for both diets while temperature and interaction significantly (P≤0.05) impact thiamin concentration only in D2.

VITAMINB1

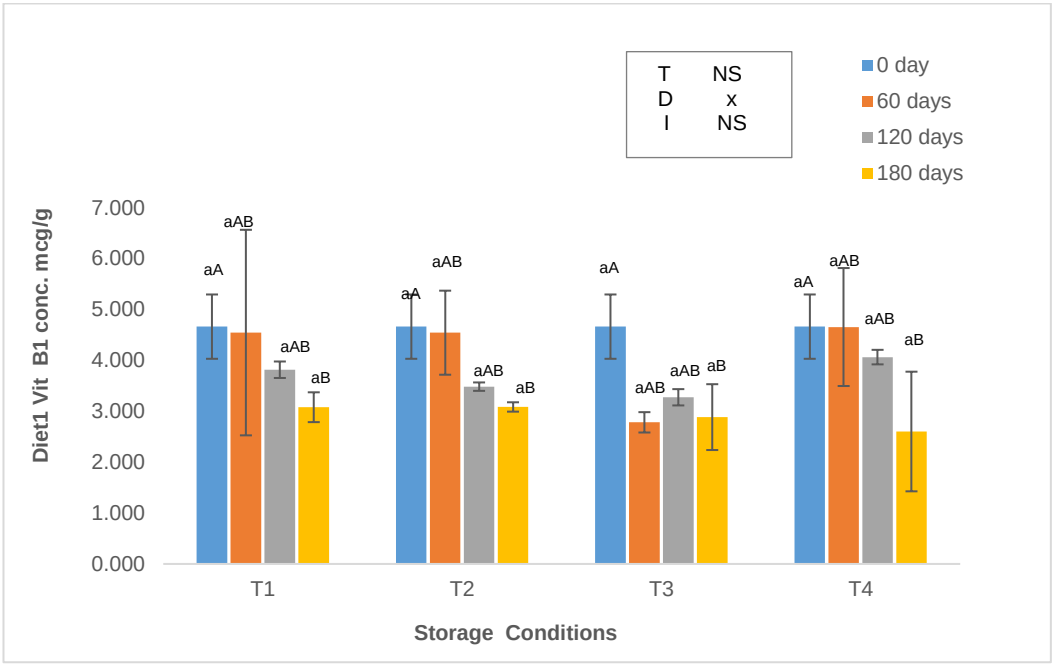


Fig.1c

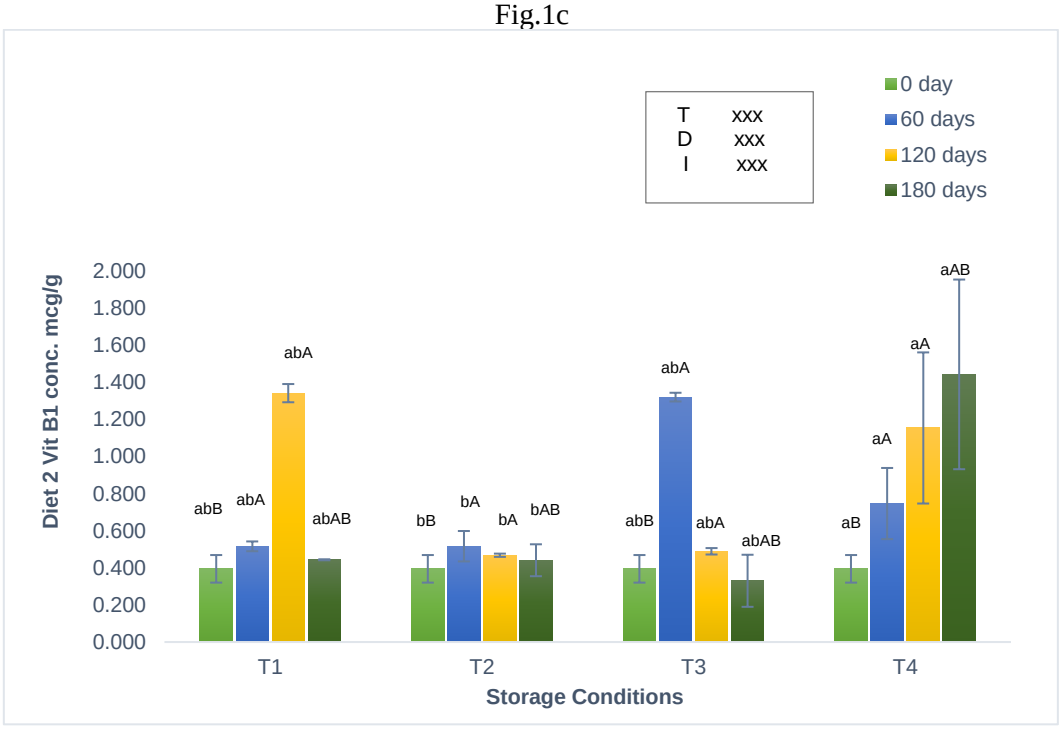


Fig.1d

3.1.3 Vitamin B2 (Riboflavin)

Vit. B2 in D1 show significant (P≤0.05) duration effects with decrease from initial to two-month storage at all temperatures. Effects of temperature and interaction on B2 changes are insignificant for D1; whereas temperature, duration, and not interaction impacts vit. B2 concentrations in D2 significantly (P≤0.05).[36] described higher heat loss of riboflavin compared to thiamin. Increasing temperatures relate to decreases in riboflavin retention [37, 38]. It is noteworthy that vit. B2 decline intensely over storage durations in D1 compared to D2, (fig.1e,1f). In D1 at 60th day, vit. B2 value at T1 decreased

from initial concentration of 0.352 to 0.003 mcgg-1 (with loss of 99.15% that is merely 0.85% retention) and remained undetected beyond two months. At T2, 60-day conc. depicts 15.34% retention (0.054 mcgg-1), at T3 6.57% is retained (0.023 mcgg-1), T4 (7.95% retention at 0.028 mcgg-1). Vit. B2 could not be determined in D1, at any storage temperature during fourth month and at T1, T4 during sixth month. Extrusion loss of riboflavin is large up to 50%. These losses due to extrusion are dependent on increasing shear rates due to increase in screw speed, moisture [19]. Extrusion does not have direct thermal effect on B2 recoveries, rather it increases B2 loss by reducing residence time achieved by increasing screw speed and moisture decrease [18,39]. For non-extruded D2, vit. B2 show significant loss at each storage interval compared to initial concentration. Retention values of vit. B2 at second, fourth and sixth months were determined respectively at T1 as 96.49%, 84.21%, 57.89%; T2 as 96.49%, 78.95%, 71.93%; T3 as 66.67%, 38.60%, 42.11%; and, T4 as 29.82%, 0%, N.D (not defined at end of storage). This is in agreement to [19,39]; elaborating higher retention rates of riboflavin under optimal conditions of processing. Being thermostable, B2 is less prone to temperature effects compared to thiamin; unless affected by processing parameters, composition of food matrix, oxidation, pH, moisture and presence of oxygen. Riboflavin is a light sensitive vitamin. Its stability is affected by oxygen, moisture and water activity [40]. Presence of oxygen during storage can severely accelerate B2 deterioration rate [41]. Photosensitization (visible or UV light exposure) of riboflavin, under oxygenated conditions can form reactive singlet oxygen further capable of oxidative degradation of other vitamins-A,C,D and E along with dietary proteins and lipids[40]. Packaging material plays an important role in B2 photostability. Intense losses of B2 can occur due to improper exposure to light from package material even when other conditions of storage (pH, temperature, moisture) are optimum [42]. Rapid loss of riboflavin from milk packaged in transparent bottle, sachet has been reported, compared to light protected milk packaging in carton or dark bottle [43]. Since transparent packaging was used for storage packaging, following may be the reason for vit. B2 deterioration in our storage diets.

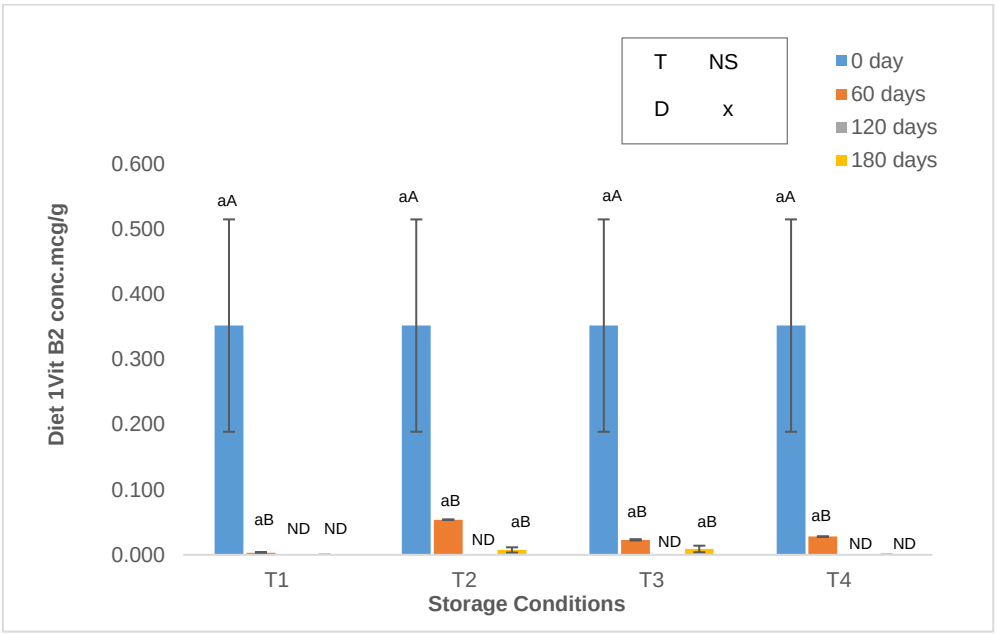


Fig.1e

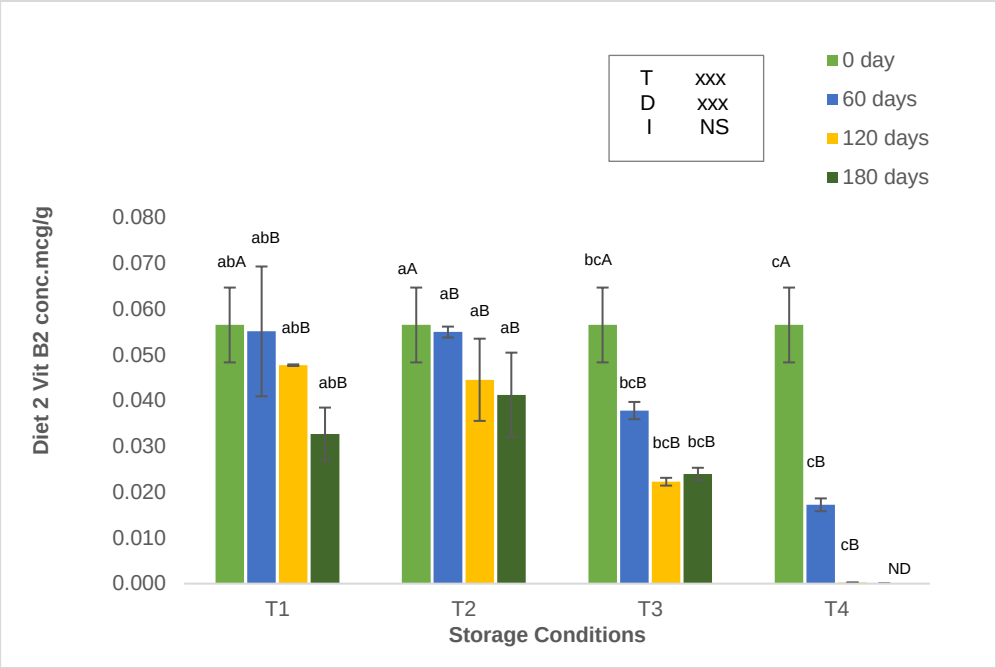


Fig.1f

3.1.4 Vitamin B6 (pyridoxine)

In D1, vit. B6 retention at end of 60,120,180 day followed 69.48%, 61.49%, 56.76% respectively at T1; 88.63%, 72.52%, 83.44% at T2; 98.53%, 75.90%, 74.54% at T3; 90.20%, 73.19%, 46.51% at T4, (fig.1g). D1 results show significant effects ($P \leq 0.05$) of temperature, duration and interaction. D2 depicts significant decrease in retention of B6 at end of storage at all temperatures, with T1,T2,T3 values at 80.37%, 80.74%, 71.48%, except with slight chromatographic overestimation at T4. Overestimations compared to initial storage content are also noteworthy for D2 at 60th day at T2, T3,T4 storage and 120th day T1, T3,T4 conditions (fig.1h). This trend may be due to artifacts of enzymatic activities (here taka-diastase) on release of bound forms vit. B1, B6 and B12 [44] as well as chromatographic overestimation due to co-elutions; a problem in simultaneous estimation of vitamins during single chromatographic run [45]. Vit. B6 exists as six distinct forms of which three vitamers- pyridoxal, pyridoxine and pyridoxamine (PL, PN, PM respectively) are relevant in chromatographic determinations. Presence of several bioactive forms of the vitamins, along with their physicochemical heterogeneity make simultaneous analysis difficult task. Variable forms and properties of B6 vitamers complicate analysis of pyridoxine [46]. Thus, applications reporting food estimations determine B6 concentration as sum of vitamers PL+PM+PN [47], PL and PM being interconvertible forms [48]. In plant-based foods PN dominates; while PM, PL are found in animal origin foods. All these forms (PL, PM, PN) exist in compounded diets consisting of plant and animal-based ingredients. According to [44], enzyme protocol in vit. B6 estimation increased their content due to the release of bound forms from food matrix; this release being under effect of enzyme concentration, purity and differential ability of enzyme combinations used to release specific forms of B6 vitamers.

VITAMINB6

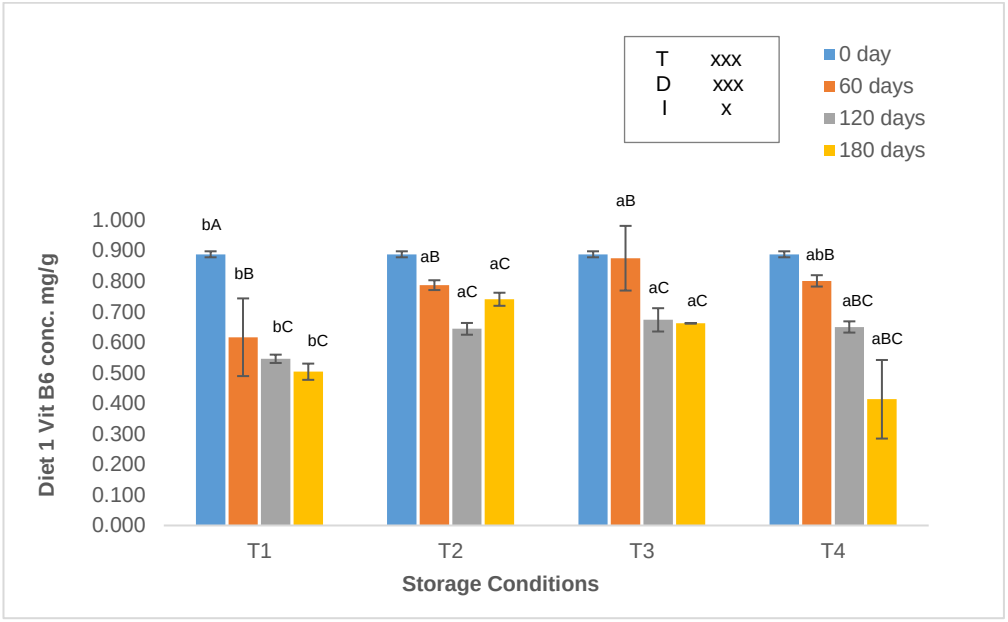


Fig.1g

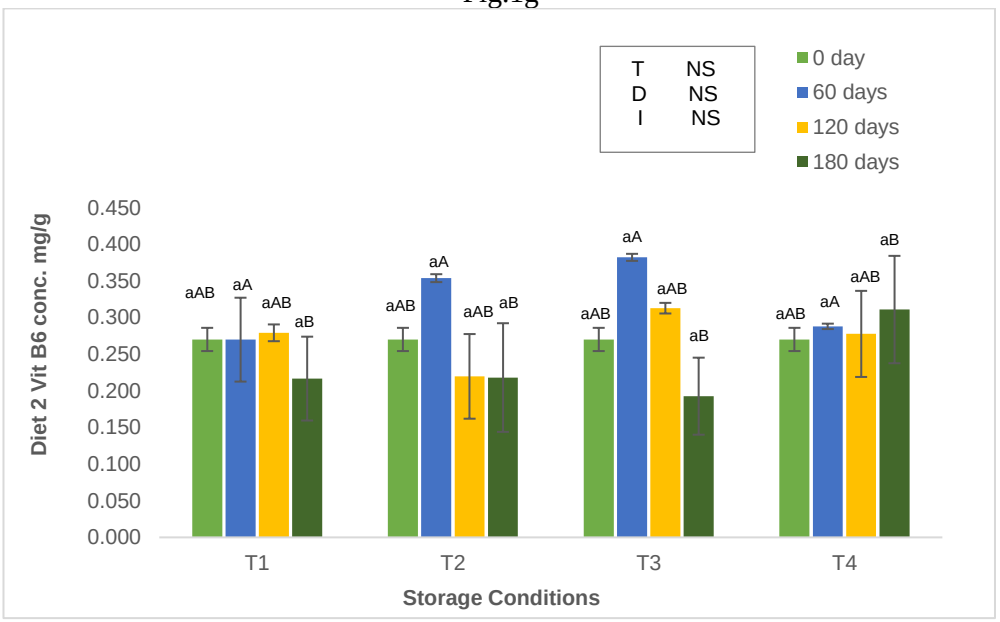


Fig.1h

3.1.5 Vitamin B12 (cobalamin)

In both diets, at each temperature condition variations in vit. B12 content exhibited duration based significant effects ($P \leq 0.05$), between initial and final storage. Temperature and interaction effects are not visible in vit. B12 changes in D1(fig.1i). For D2, duration and interaction effects are highly significant for changes in vit. B12 value, but no noteworthy effect of temperature lies for both diets during storage (fig.1j). T1 retention of cobalamin estimated at 45.27% for D1, 74.07% (D2); T2 (32.92% D1), (55.56% D2); T3 (28.81% D1), (49.38% D2); T4 (65.84% D1), (77.78% D2) clearly with higher vit. B12 retentions in D2 during storage. Although T1, T2 retentions for vit. B12 are higher compared to T3, in both diets at the end of storage; T4 values at 180-day in both diets are highest although this increase at T4 is unexplainable. As per [39] few studies are available on cobalamin; and possibly fewer about its storage behavior in fish feeds limiting our knowledge on the possible effects of variables on cobalamin changes. Present work may fulfill this gap in literature.

VITAMINB12

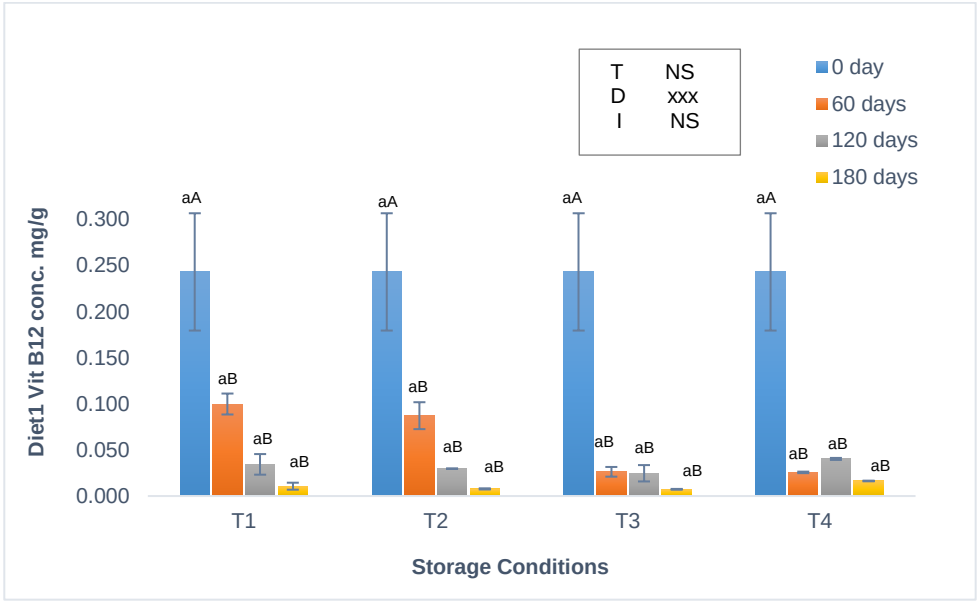


Fig.1i

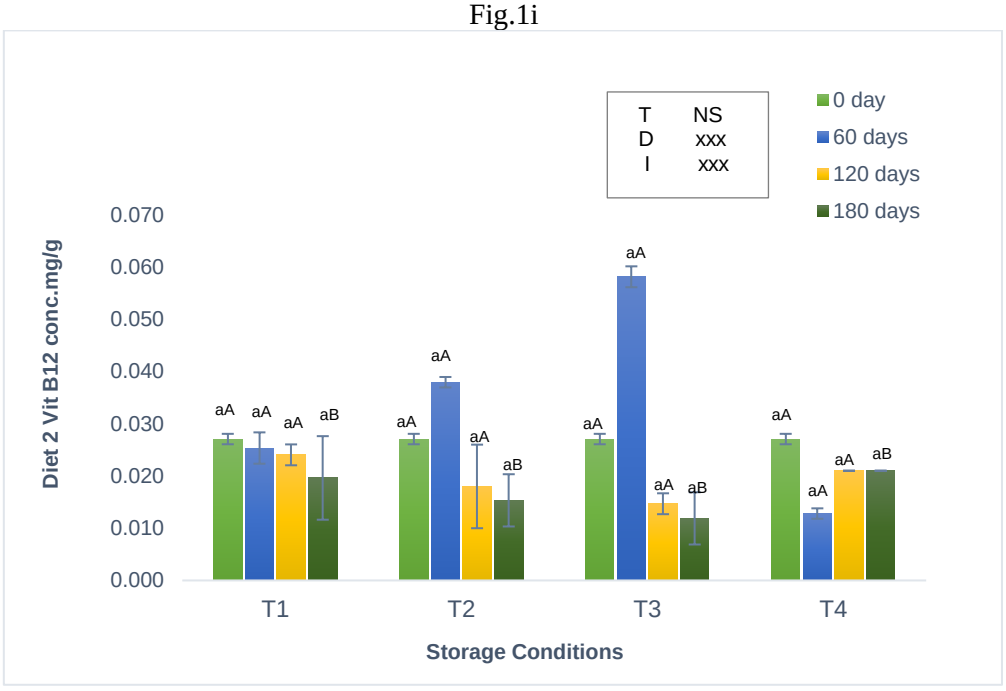


Fig.1j

3.2 Fat-soluble vitamin storage changes

Fat-soluble vitamins are also prone to effects of storage and feed processing variables. Moreover, lipid oxidation in stored diets yield hydroperoxides that reduce availability of fat-soluble vitamins in storage feeds [20].

3.2.1 Vitamin A (retinyl acetate)

At each bimonthly assessment vit. A decreased at all storages temperature in both diets. In D1 end of duration vit. A retentions follow T1(72.52%), >T2 (38.65%), >T3 (20.96%), >T4 (3.26%) showing decline across increments in storage temperatures (fig.2a). Similar trends are observable towards vit. A retention rates in D2, at 71.37% (T1), 68.46% (T2), 24.48% (T3), 4.98% (T4), fig.2b. Under feed storage regime higher vit. A retentions in D2 than D1 explains less severe effects of pelleting compared to extrusion. Temperature affects vit. A values in D2significantly ($P \leq 0.05$), between low temperature (T1,T2) and other temperature conditions (T3,T4), between T3 and T4. Between subject effects of temperature and duration on vit. A storage depletion are highly significant (P=0.000) for both diets, with significant interaction effects only towards D2. High temperatures above 100oC, moisture, extrusion conditions (such as screw speed) have extreme impact on vit. A depletion [13,18,49].Vit. A has essential

function in vision, embryonic development, growth, reproduction, and differentiation [11]. Vit. A is extremely labile to oxidative loss arising from lipid, protein oxidations following free radical attacks.

VITAMIN A

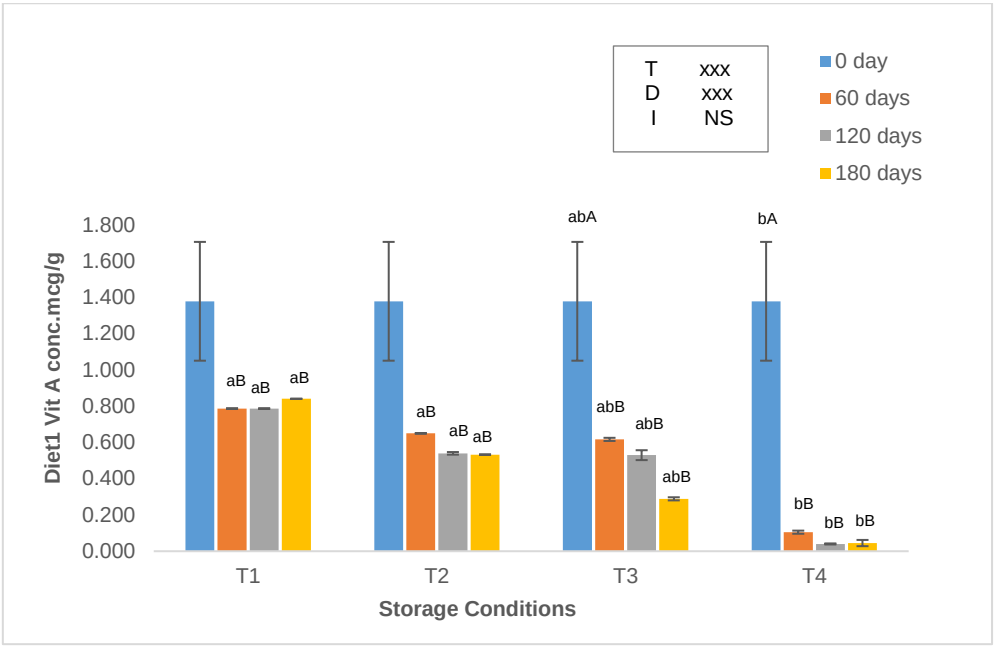


Fig.2a

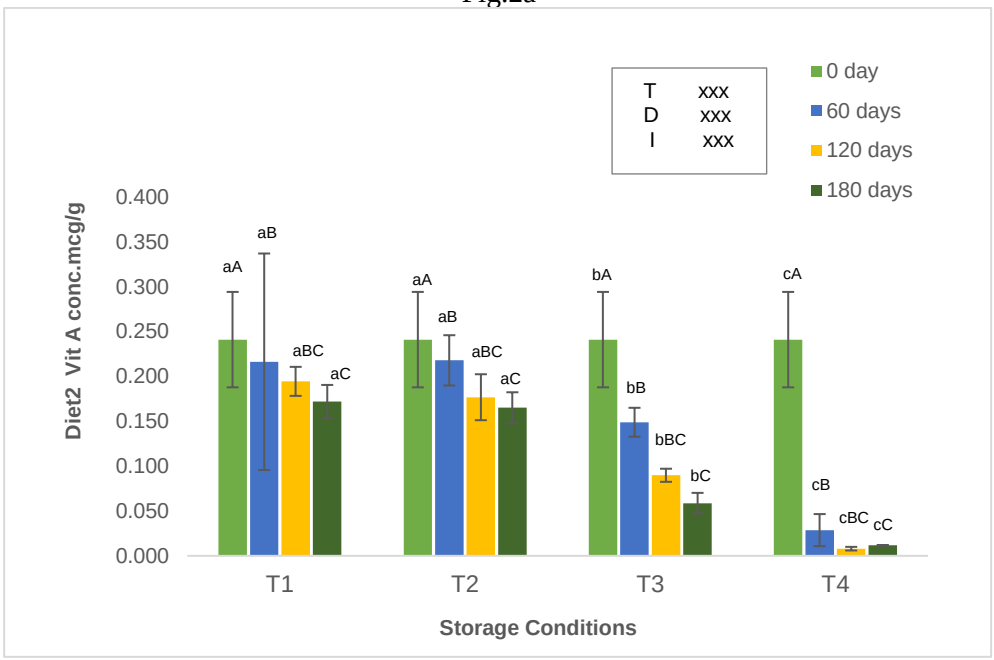


Fig.2b

3.2.2 Vitamin D (vit. D2 ergocalciferol, vit. D3 cholecalciferol)

Detectable vit. D in D1 is plant-based form ergocalciferol (vit. D2). For D2 animal-based form cholecalciferol (vit. D3) was quantifiable. Estimated retention of vit. D2 (in diet D1) at end of storage accounted 99.40% (T1), 50.57% (T2), 30.21% (T3), 39.54% (T4), fig.2c. Severe losses of vit. D3 were encountered in pelleted stored diet with incurred losses accounting 100%, fourth storage month onwards at T4 storage. Even at T1,T2 and T3 accountable vit. D3 retentions were low as 38.46%, 26.92%, 7.69% respectively (fig.2d). These results are in alignment to [36],depicting high pelleting losses of vit. D3 nearing 80-90% with stable extrusion changes. Only duration effects (P=0.000) are noteworthy for vit. D3 changes over storage, with no significant effect due to temperature (P=0.084) or interaction (P=0.442). Alterations in vit. D2 was significantly ($P \leq 0.05$) under impact of temperature, duration and their interactions as well.

VITAMIN D

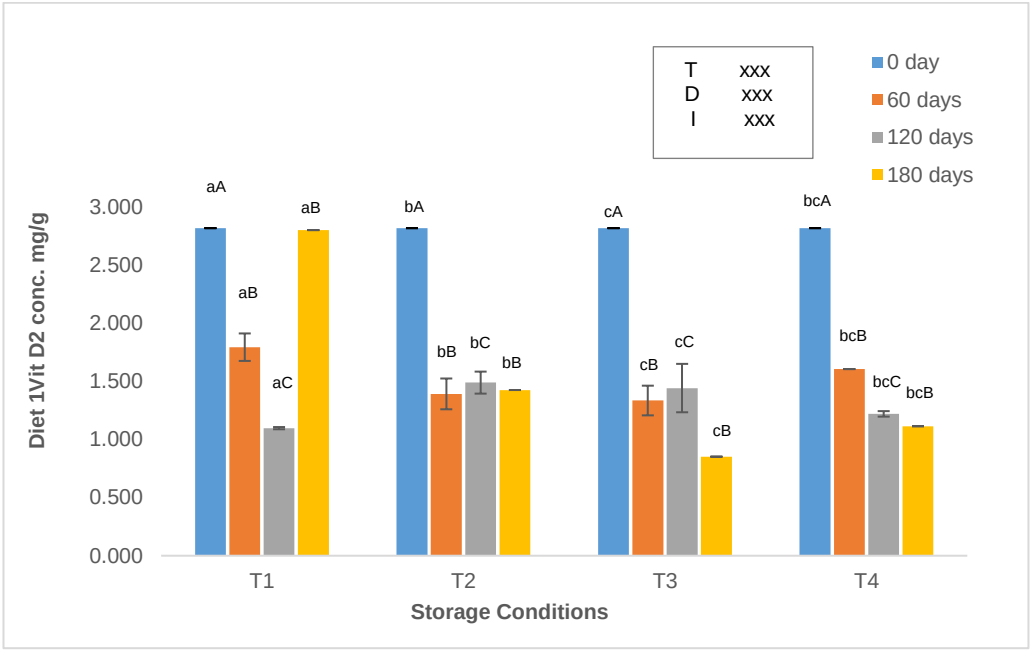


Fig.2c

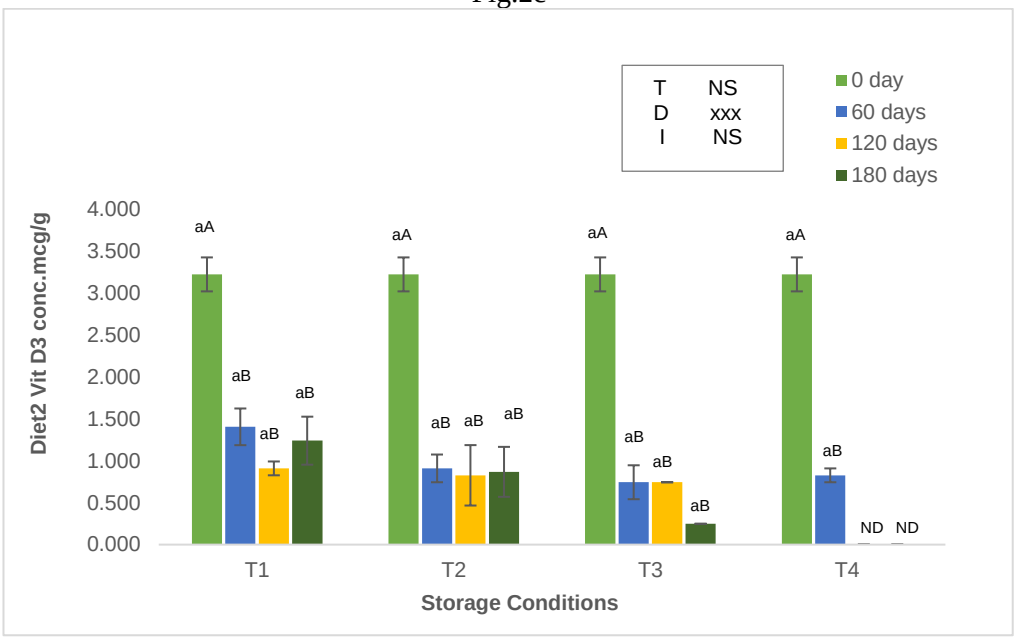


Fig.2d

3.2.3 Vitamin E (tocopherol)

Vit. E loss ranging from 39.41% to 51.36% are estimated for extruded diet D1, at end of storage, (fig.2e).At storage culmination,pelleted D2 loss accounted between 59.78%-80.43% (fig.2f).Variabilities in dietary loss of tocopherols may be with effect to thermal oxidation, degree of lipid rancidity, presence of light, or metal chelators typically zinc, iron, and copper [50]. Vit. E can prevent fragility of erythrocytes and due to its anti-oxidant mechanism prevent free radical oxidation of dietary (PUFA) and membrane lipids [12]. Dietary vit. E requirements for rohu fryis determined to be 131.91 mgkg-1 dry weight of feed by [12] and diet conc. 120 mgkg-1for Atlantic salmon [51]. Vit. E along with vit. A, C have strong antioxidant properties owing to their free radical sequestration potential. Thus, are prone alike to oxidative effects during storage.

VITAMIN E

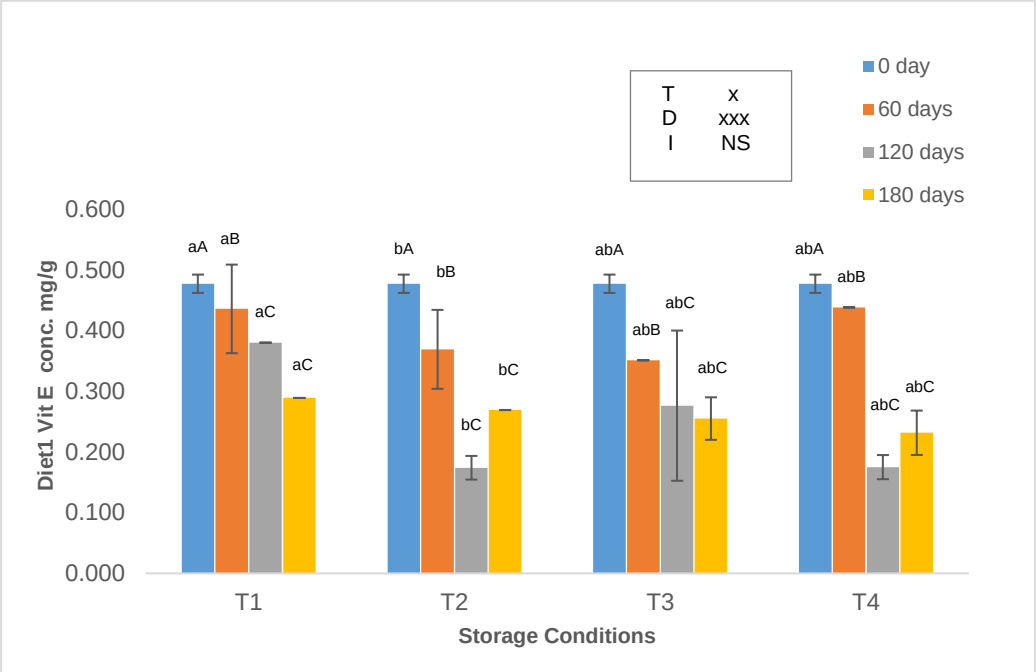


Fig.2e

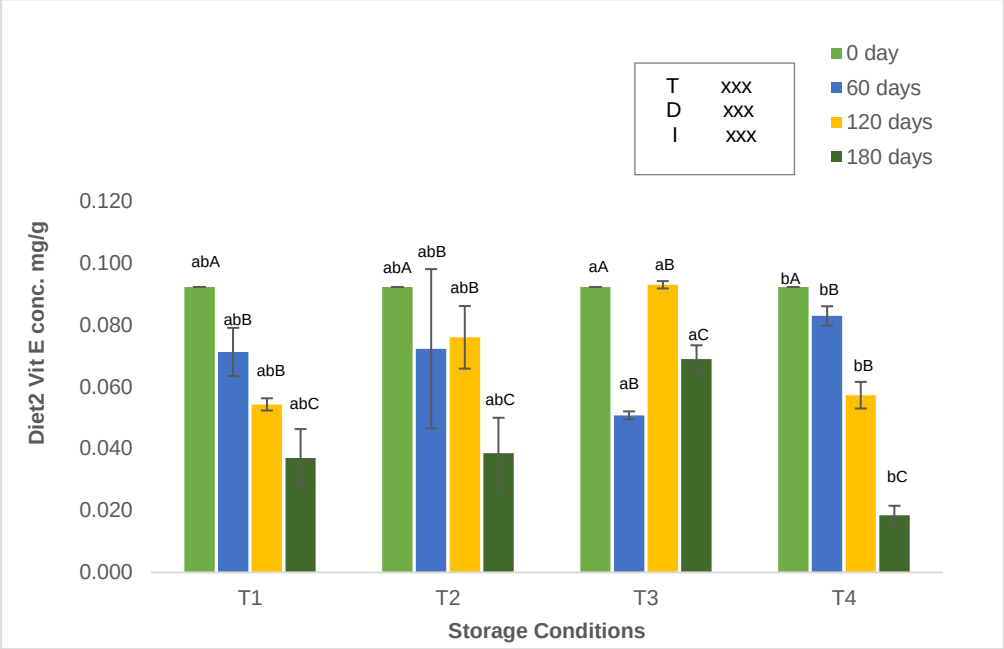


Fig.2f

3.2.4 Vitamin K1 (Phylloquinone)

Vit. K1 in feeds was identified as trans- and cis-isomers (fig.2g-2i) at chromatographic retention times (RTs) 1.09-1.15 for cis-, 1.77-1.807 for trans- isomer. Of these, only trans-form of vitamin K1 is biologically relevant [52]. K1-trans, during storage, in both diets (fig.2g,2h) show significant between-subject-effect for duration; but not for storage temperature and interaction. Incurred losses are higher for extruded diet (between 41.28%to64.41%), than pelleted D2 (4.81%-49.89%).

VITAMIN K

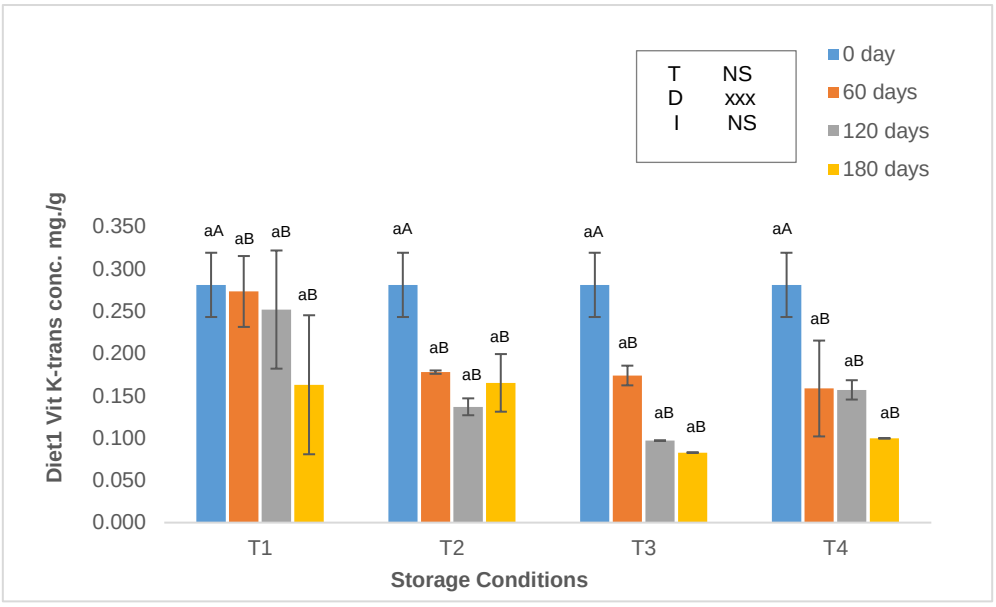


Fig.2g

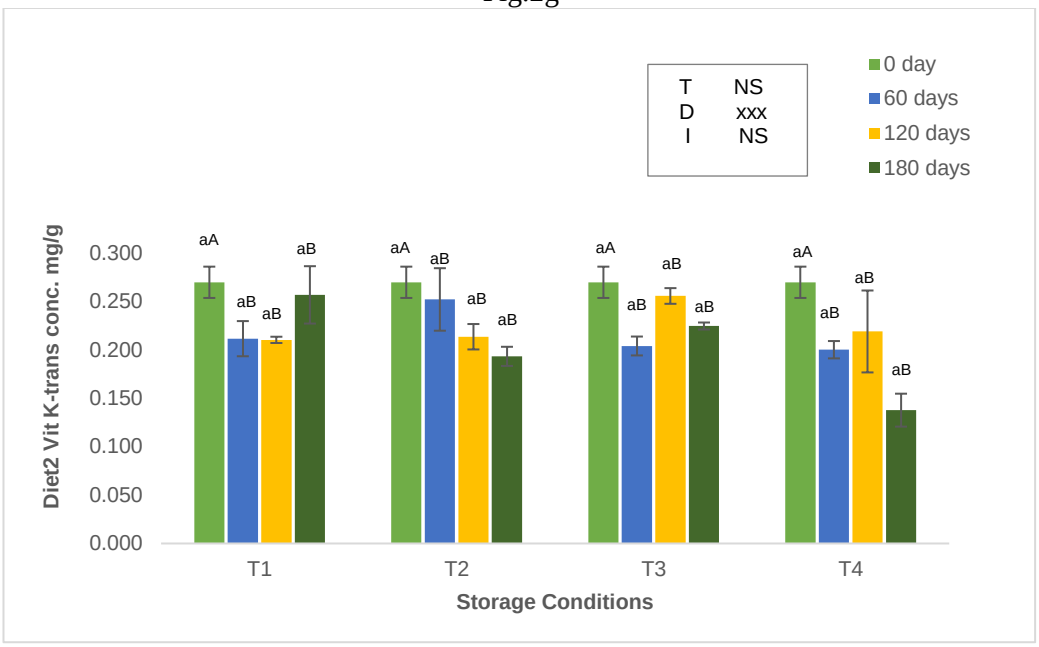


Fig.2h

4. Conclusions

The study evaluates effect of storage duration, temperature and their interaction, on vitamin profile in fish feeds prepared from extrusion and non-extrusion pelletization procedures. For both diets, substantial storage loss in water- miscible (B2, B12 and C), and fat-miscible (A, E, K) vitamins is evidentfrom 60-day onwards, at all storage temperatures. Intense diminutions in vit. D content was observable during four month storage at T4 condition for D2; while storage temperature, duration and interaction effects on dietary vit. D were significant ($P \leq 0.05$) for D1. D2 retentions for thiamin are influenced by interference from antinutrient factors in feed; while co-elutants affect determination of vit. B6 in D2. Through the study, it is evident that depletion in essential micronutrients with regard to feed vitamins incurs severe loss to feed quality. Incurred loss of most vitamins were highest at end of storage under high temperature T4 conditions.These diminutions are caused by oxidativeas well asnon-oxidative changes in vitamins along with interactive consequences of other dietary nutrients (proteins, lipids, elements/metals)on vitamin profileduring storage. It is noteworthy thatvitamin analysis and estimation is strongly influenced by processing technique, co-elutionand antinutrient interferences during chromatographic run.Furthermore, the studyhelps cogniseimpact of temperature and duration on feed

quality at storage inventories, and in farmingenvironments. Storage depletion of vitamins signify need to develop feed formulae taking into consideration the percentage loss of each nutrient from the foodstuffs, specifying at strategic formulation of feedkeeping in account storage impacts.Additionally, these findings offer helpful information for fish farmers in managing storage of formulated feeds, with an aim to prolong their shelf-life safeguarding significant amount of the total production costs of fed-aquaculture and incidental consumer health.

Acknowledgements
Authors are thankful to their institutions Maitreyi College, University of Delhi and Delhi Technological University for constant support towards formalisation of the research.

Author contributions
Parul Puri: experimentation, literature search, conceptualization, writing-original draft preparation, analysed the study, formal analysis. Vaibhav Puri: Formal analysis, result interpretation; Neha Tiwari: standardisation and performance of vitamin determination, formal analysis. Ram Singh, Jai Gopal Sharma, Rina Chakrabarti: supervision, writing-reviewed the overall manuscript, edited, critically revised drafted work. All authors agreed on the final draft.

Statements and Declarations
Competing Interests: Authors declare that there are no financial or non-financial interests directly or indirectly related to the work submitted for publication also there is no conflict of interest in regard to same.

Funding: This work did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability: Data available from corresponding author on reasonable request.

References

1. NRC,Nutrient Requirements of Fish. (Subcommittee on Fish Nutrition, 1993). http://www.nap.edu/catalog/2115.html
2. F.C.P.Silva, J.R.Nicoli, J.L.Zambonino-Infante, M. Le Gall, S.Kaushik, F.J.Gatesoupe, Aquaculture 306,233(2010)
3. S.S. Giri, inFarm-made Aquafeeds: Opportunities, Challenges and Policy Intervention, : ed. By S.S. Giri (SAARC Agriculture Centre, Dhaka, Bangladesh, 2017)p. 1
4. HLPE, Nutrition and food systems. (A report by the high level panel of experts on food security and nutrition of the committee on world food security, Rome, 2017)
5. I.Pípalová, Aquaculture International 11,325(2003)
6. R.Chakrabarti, W.D.Clark, J.G.Sharma, R.K.Goswami, A.K.Shrivastav, D.R. Tocher, Frontiers inChemistry 6, 479(2018)
7. G.M. Srirangam, International Journal of Fisheries and Aquatic Studies 4,133 (2016)
8. NRC, Nutrient requirement of fish and shrimp. (USA: National Academy of Sciences, 2011).
9. S.Afrin, M.M.Rahman,IOSR Journal of Agriculture and Veterinary Science12, 53(2019)
10. NRC, Fat-Soluble Vitamins.in Diet and Health: Implications for Reducing Chronic Disease Risk. National Research Council Committee on Diet and Health. (National Academies Press,USA, 1989). https://www.ncbi.nlm.nih.gov/books/NBK218749/
11. L.H.Hernandez, R.W.Hardy, Aquaculture Research 51, 3061(2020)
12. S.K.Sau, B.N.Paul, K.N.Mohanta, S.N.Mohanty, Aquaculture240, 359(2004)
13. J.M. Harper in Effects of extrusion processing on nutrients, ed. ByE .Karmas, R.S. Harris (Van Nostrand Reinhold Company, New York,1988), p. 365
14. S.G.Solomon, L.O.Tiamiyu, V.T.Okomoda, K. Adaga, Croatian Journal of Fisheries 74, 30(2016)
15. A. Kołakowska, G. Bartosz, in Food oxidants and antioxidants chemical, biological, and functional properties. Chemical and functional properties of food components series. (CRC Press, Taylor and Francis Group,London, 2014), p. 1
16. M. Sorensen, T. Storebakken, K.D. Shearer, Aquaculture Nutrition 11,251(2005)
17. S. Singh, S. Gamlath, L. Wakeling, International Journal of Food Science and Technology 42, 916(2007)
18. M.N.Riaz, M.Asif, R.Ali, CriticalReviews in Food Science and Nutrition 49, 361(2009)
19. P.Yang, H.Wang, M.Zhu, Y.Ma, Animals 10, 894(2020)
20. A.Kop, K.Gamsiz, A.Y.Korkut, H.Saygi, Turkish Journal of Agriculture - Food Science and Technology 7, 43(2019)
21. P.Chakraborty, T.Singh, J.Sharma, R.Chakrabarti, Journal of Aquaculture in the Tropics 35, 149(2020)
22. AOAC, Official methods of analysis of the association of official analytical chemists, 18th edn. (AOAC International,Gaithersburg, MD,2005)
23. R. Sami, Y. Li, B. Qi, S. Wang, Q. Zhang, F. Han, Y. Ma, J. Jing, L. Jiang, Journal of Chemistry831357(2014)
24. H.Qian, M. Sheng, Journal of Chromatography A 825, 127(1998)
25. A.N.Dainton, B.White, L.Lambrakis, C.G.Aldrich, Frontiers in Veterinary Science 9,1090695(2022)
26. D. Granato, V.M.A. Calado, B. Jarvis, Food Research International 55,137(2014)
27. M.C.Giannakourou, P.S.Taukis, Critical Review and Kinetic Modelling Approaches: Food 10, 2630 (2021)
28. A.K.Soliman, K.Jauncey, R.J.Roberts, Aquaculture 60, 73(1987)
29. B.W. Tucker, J.E. Halver, Fish Physiology and Biochemistry 2,151(1986)

30. K. Sandnes, F. Utne, Fisk. Dir. Skr. Ser. Ernaering2,39 (1982)
31. M.F.Nasar, S.Z.H.Shah, K.Aftab, M.Fatima, M., Bilal, M.Hussain,Aquaculture Nutrition(2021)https://doi.org/10.1111/anu.13327
32. S.S.O.Hung, S.J. Slinger, International Journal for Vitamin and Nutrition Research 50, 393 (1980)
33. J.G. Bell, J. McEvoy, D.R. Tocher, J.R. Sargent, Journal of Nutrition 130, 1800(2000)
34. B. DiSabatino, J.M. Parr, G. Kritikos, G. Monteith, A. Verbrugghe, Canadian Veterinary Journal62, 374 (2021)
35. K.Rungruangsak, P.Tosukhowong, B.Panijpan, S.L.Vimokesant, The American Journal of Clinical Nutrition 30, 1680 (1977)
36. F. Kubitza, J.E.P. Cyrino, in Advances in extrusion technology,ed. By Y.K. Chang, S. : (Tecnomic publishing company, Wang Lancaster, 1998),p. 53
37. I. Bjorck, N.-G. Asp,Journal of Food Engineering2, 281(1983)
38. N.Athar, A.Hardacre, G.Taylor, S.Clark, R.Harding, J.McLaughlin,Journal of Food Composition and Analysis 19, 379 (2006)
39. P.Morin, A.Gorman, L.Lambrakis, Animal Feed Science and Technology 277, 114975(2021)
40. E.Choe, R.Huang, D.B.Min, Journal of Food Science 70, R28(2005)
41. D. Dennison, J. Kirk, J. Bach, P. Kokoczka, D. Heldman, Journal of Food Processing and Preservation 1, 43 (1977)
42. M.A.Sheraz, S.H.Kazi, S.Ahmed, Z.A.I.Ahmad,Beilstein Journal of Organic Chemistry10,1999(2014)
43. M. Bekbolet, Journal of Food Protection53,430(1990)
44. S.Ndaw, M.Bergaentzle, D.Aoude-Werner, C. Hasselmann, Food Chemistry 71, 129(2000)
45. A.Mateeva, M.Kondeva-Burdina, L.Peikova, S.Guncheva, A.Zlatkov, M.Georgieva, Heliyon 9, e12706(2023)
46. M.A.Kall, Food Chemistry 82, 315(2003)
47. M.Yaman, Avrupa Bilim ve Teknoloji Dergisi Sayi 15,S 611(2019)
48. E.E. Snell, Journal of the American Chemical Society 67,194 (1945)
49. M.E. Camire,in Extrusion cooking, technology and application, ed. By R. Guy (Woodhead Publishing Limited,Cambridge, 2001), p. 116
50. S.J. Charlton, W.N. Ewing, The vitamin directory. (Context Products Ltd., England,2007)
51. K. Hamre, O. Lie, Comparative Biochemistry and Physiology Part A: Physiology 111, 547(1995)
52. T.A. Berger, B.K. Berger, Chromatographia 76, 109(2013)

Physical and Engineering Sciences in Medicine
https://doi.org/10.1007/s13246-025-01534-0

SCIENTIFIC PAPER

Integrated wearable PPG: a multi-vital sign monitoring based on group sparse mode decomposition framework in remote health care using PPG signal

Pankaj[1] · Pratibha Maan[2] · Manjeet Kumar[3] · Ashish Kumar[4] · Rama Komaragiri[5]

Received: 26 September 2024 / Accepted: 10 March 2025
© Australasian College of Physical Scientists and Engineers in Medicine 2025

Abstract
Monitoring vital signs using a photoplethysmogram (PPG) signal has gained considerable attention, allowing users to monitor anyone, anywhere, and anytime with an objective. In recent years, advances in wearable technology and signal processing techniques have paved the way for accurate and reliable vital sign monitoring using PPG signals. Early detection of cardiovascular diseases can help the physician treat the disease promptly; thus, realtime monitoring of vital signs has emerged. Any deviation in the threshold value of vital signs can indicate potential threats to the cardiovascular system. The need to monitor vital signs in realtime using wearable devices has attracted the interest of the healthcare industry in developing simple and efficient vital sign estimation algorithms. This research introduces a framework to estimate the following important vital signs: heart rate (HR), respiratory rate (RR), blood pressure (BP), and blood oxygen saturation (SpO2), concurrently by overcoming the limitations posed by state-of-the-art techniques that primarily focus on individual or two vital sign estimations. Our proposed approach leverages signal processing techniques to determine the above-mentioned vital signs seamlessly and accurately. This innovation enhances the efficiency of vital sign monitoring and presents a unified solution for comprehensive health assessment. The widespread use of wearable devices for monitoring realtime health status in everyday life manifests in using PPG sensor-enabled wearable devices to perform more complex computational tasks. To date, the algorithms proposed to process an input PPG signal often use multiple processing steps to estimate any vital signs. This can increase the computational complexity of these algorithms, making it challenging to deploy devices with limited computational resources. The proposed work introduces a computationally efficient framework to estimate all four vital signs using the signal framework. The experimental results obtained with the proposed framework demonstrate that the proposed work outperforms the state-of-the-art estimation accuracy and computational complexity.

Keywords Blood pressure (BP) monitoring · Heart rate (HR) · Medical devices · Oxygen saturation level (SpO2) · Photoplethysmography (PPG) · Respiratory rate (RR) · Wearable device

Introduction

Vital signs are measurements of the primary functions of the human body. They are called "vital" because they reflect the essential functions needed to sustain life. The clinical value in measuring vital signs such as heart rate (HR), blood pressure (BP), respiratory rate (RR), and blood oxygen saturation (SpO2) using wearable devices provides a valuable opportunity in the field of digital health [1]. Wearable devices can continuously compute vital signs, thus obtaining measurements made without the conscious knowledge of the individual. The growing interest in using wearable health monitoring devices has naturally expedited PPG sensor-enabled wearable device development. Several algorithms to design wearable devices enabling electrocardiography (ECG) and photoplethysmography (PPG) sensors are discussed in the state of the artwork. Although the state-of-the-art shows significant results, wearable techniques still face research challenges concerning the computational complexity, accuracy, and reliability of data processed by these algorithms [2]. While ECG remains the gold standard for analyzing health status, its complexity, cost, and portability present hurdles to adopting its use in personnel homecare settings. The healthcare industry is trying to find a feasible approach to

Extended author information available on the last page of the article

Published online: 07 April 2025

monitoring and tracking multiple vital signs from a simple, easy-to-use wearable device to tackle the limitations of ECG. Potable wearable sensors are desirable, especially in emerging healthcare [3]. As a result, the healthcare industry's interest in leveraging PPG signals for vital sign estimation is growing due to their affordability, usability, and potential to monitor vital signs continuously. This feature makes PPG techniques more accessible for remote health monitoring and personalized home care.

PPG is an optical method to detect volumetric blood changes during a cardiac cycle. The PPG technique is technically simple to integrate with wearable devices for monitoring realtime vital signs [4].

Over the past few years, PPG sensor-based healthcare devices have been widely used to capture vital signs in clinical settings. PPG signals are produced by detecting blood content and volumetric changes in the microvascular bed of tissue. Pioneering work in wearable healthcare devices has demonstrated the potential of analyzing PPG signals for extracting vital signs in realtime [5]. However, addressing key challenges such as sensor quality, data transfer, and battery life is crucial for improving the effectiveness and practicality of wearable devices for vital sign monitoring. For instance, PPG data acquisition requires an optimal sampling frequency to balance data quality and system constraints. While higher sampling frequencies improve accuracy by capturing more detailed information, they also lead to increased data transfer requirements, higher memory utilization, and high computation requirements, leading to increased power consumption, which impacts the battery life of wearable devices. On the other hand, lower sampling frequency may compromise the accuracy of vital sign estimation, especially in real-time applications where precision is critical. Thus, careful attention must be paid to selecting a suitable sampling frequency to balance these trade-offs while designing a wearable device. The primary cardiac component of the PPG signal typically falls within the 0.5–5 Hz range, while motion artifacts and noise span higher frequencies up to 20 Hz [6]. To accurately capture the vital signs (such as HR, BP, RR, and SpO2) without aliasing, the sampling frequency is well above the minimum required to capture the cardiac behavior accurately. A higher sampling rate ensures accurate signal representation but increases data transfer size, processing power, and battery consumption, which needs to be balanced in wearable devices. Therefore, selecting an optimal sampling frequency is crucial for capturing accurate PPG signals while considering system constraints like power consumption and data transfer.

Further, accurately extracting vital signs using wearable devices poses a significant challenge due to motion artifacts in the acquired PPG signal. Therefore, to achieve reliable and precise vital sign estimation, it is essential to separate the motion artifact component from the vital sign components [7].

To this end, a novel algorithm based on group sparse mode decomposition (GSMD) techniques to separate motion artifacts components and the clean PPG signal components from monitoring all vital signs from a single PPG sensor signal. The proposed framework indicates that the clean PPG signal frequency spectrum consists of a dominant frequency component within the desired frequency range of vital signs [8]. The computation of the dominant frequency component guarded by a low-frequency component indicates the vital sign. In this study, the framework is designed to provide a novel solution to vital sign estimation that can be measured continuously using a wearable PPG sensor. Next, to refine the value of vital signs in real-time, a vital sign tracking algorithm is proposed that enables accurate estimation of vital signs in real-time.

Methodology

Data acquisition and segmentation

This work proposes a GSMD algorithm-based vital sign estimation framework, as shown in Fig. 1. The proposed framework considers HR, SpO2, RR, and BP. No common publicly available dataset provides the PPG signal recording for all vital signs. Thus, a publicly available benchmark dataset for each vital sign is used to analyze the proposed work's performance.

Dataset description

Four publicly available datasets (IEEE SPC [9], BIDMC [10], Capnobase [11], and MIMIC-3 [12]) are employed to check the potential of the proposed vital sign estimation framework. The datasets are recorded using a sampling frequency of 125 Hz, except for the Capnobase dataset, which has a sampling frequency of 300 Hz. Table 1 depicts the details of the dataset used to compute the vital signs.

Window segmentation

Segmentation of PPG signals is a common approach for continuously estimating vital signs. PPG signal segmentation in consecutive windows provides a mechanism for analyzing the PPG signal incrementally in realtime. A PPG signal acquired in realtime may contain physiological information and noise information. Segmentation of the PPG signal in a consecutive window allows the framework for adaptive and continuous analysis [13]. It can track vital signs without waiting to process the complete signal. In realtime processing, segmentation of the signal using a sliding windows

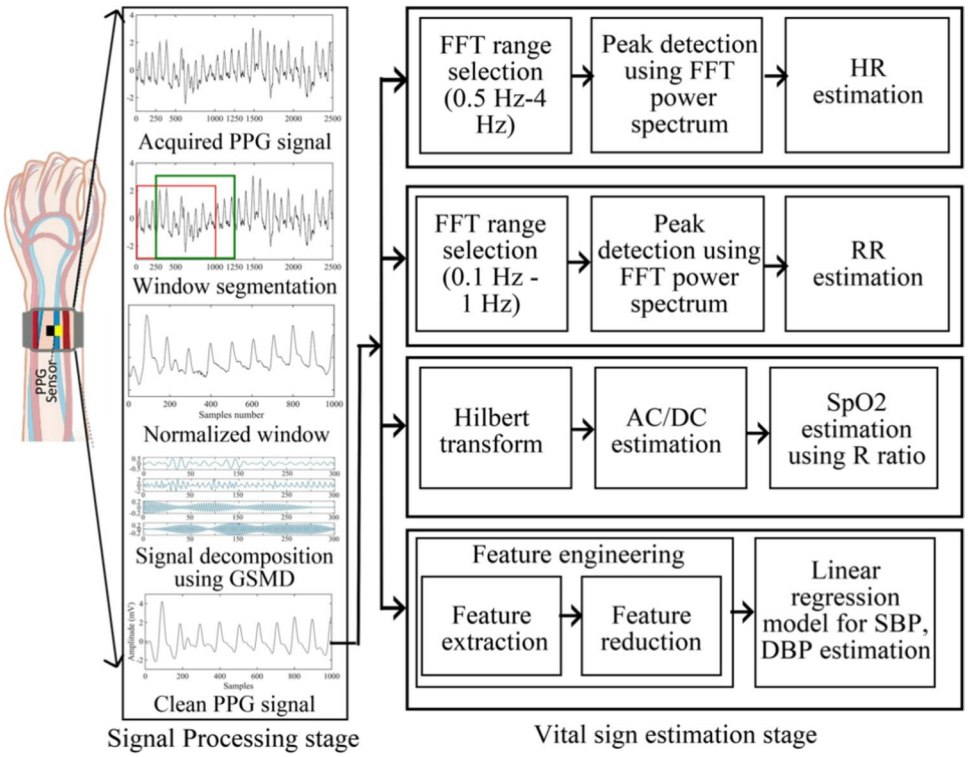


Fig. 1 Functional overview of the proposed vital sign estimation framework

scheme, in which windows advance in time as a new signal arrives, increases accuracy and reliability.

In Table 1, the window size represents the sample duration used to compute the true value of vital signs. In the proposed work, to generalize the model and consider each vital sign with the same scale, the proposed work is evaluated with the same window segmentation protocol used to estimate the true value of the vital sign. These true values of vital signs are used to check the efficacy of the proposed framework. Table 1 shows the total number of windows framed with each dataset employed to compute the true value of vital signs.

Signal decomposition

This work uses the GSMD algorithm to separate the frequency components of the PPG signal and the motion artifacts/noise component. GSMD decomposes the acquired PPG signal, which is non-stationary, into a distinct set of intrinsic mode functions (IMFs). Each IMF has limited bandwidth, and the frequency band of each IMF has no common spectra. The GSMD algorithm uses an ideal filtering technique designed in the frequency domain to extract each individual IMF from the PPG signal [14]. These filters isolate the desired frequency components from the signal. The filter bank f_a is modeled as an on–off group-sparse vector, reflecting that specific frequency components are present in distinct groups.

The coefficients of the ideal filter bank are estimated using a penalized least-squares approach. The penalty term utilized in the estimation process involves a weighted $\ell 0$-norm, and the weighting parameters are determined using techniques that involve energy detection over short windows.

Table 1 Vital sign dataset description and window formation

Dataset	Vital sign estimate	Duration of record (in second)	Window size (in second)	Step size (in second)	Number of windows framed
IEEE SPC [9]	HR	300	8	2	3204
BIDMC [10]	SpO2	480	8	1	4440
Capnobase [11]	RR	480	32	3	5102
MIMIC-III [12]	SBP, DBP	600	8	2	3610

The penalty term encourages both on–off and group-sparsity properties in an ideal filter bank, which promotes the presence of certain groups of frequencies and keeps other frequencies zero. A regularization parameter λ is also used, and its optimal value depends on factors like noise variance and window size. The decomposed levels are adaptive. The GSMD algorithm adjusts the window size based on the bandwidth of different IMFs and estimates the noise variance from the measured data, thus providing adaptivity. After convergence, the estimated filter bank f_a is decomposed into nonoverlapping filters f_{ai}. Each filter f_{ai} contains a block of adjacent nonzero entries in f_a. These filters are then used to extract individual IMFs. Thus, GSMD reduces user intervention by adaptively decomposing an input PPG signal into several IMFs. The utilization of the weighted l0-norm and the group sparsity algorithm converges the complete framework in a relatively short time compared to similar algorithms.

The key highlights of the proposed GSMD-based vital sign estimation framework are as follows.

- This work introduces a computationally efficient framework to estimate all four vital signs using the signal framework that solely requires a PPG signal, unlike existing frameworks that concentrate on single or paired vital sign estimations.
- Our approach proposes a comprehensive solution capable of simultaneously assessing all vital signs using a single wearable device, offering a holistic picture of an individual's health status in realtime.
- The proposed GSMD-based algorithm utilizes the concept of group sparsity to identify specific frequency components that are grouped.
- GSMD converges relatively quickly and is suitable for decomposing long signals. This efficiency is advantageous when dealing with extended data sequences.

Additionally, the GSMD algorithm automatically extracts the IMFs related to the clean PPG signal, and the sum of these IMFs provides a clean PPG signal. Algorithm 1 depicts the stepwise process of the proposed GSMD-based vital sign estimation framework.

Algorithm 1: Proposed GSMD-based estimation of vital signsusing corrupted PPG signal

Input: Acquired PPG signal. (Vital sign estimation dataset)
Result: HR, SpO_2, RR, SBP, DBP
1: Segregate S into a consecutive window containing N samples.

$$S^b = (S((b-1)\times w)+1 \ldots . S((b-1)\times w)+N)$$

Where w is the step size equal to two seconds, b is the window number
2: for b = 1 to Z do; Z is the total number of eight-second window.
3: Obtain normalized S^b using the relation:

$$S^b = \frac{S^b - \mu\left(S^b\right)}{\sigma\left(S^b\right)}$$

Here, μ and σ represent the mean and standard deviation, respectively.
4: Decompose S^b into J basic functions called intrinsic mode functions (IMFs)

$$S^b = \sum_{j=1}^{J} h_j^b$$

5: h_j is optimally calculated by using a relation

$$h_j = \mathrm{T}^{-1}(IF_j \quad C_S), j = 1,\ldots..,J,$$

C_S and IF_J are the discrete cosine transform (DCT) coefficient vectors of the input signal and an ideal filter. T^{-1} is the inverse DCT
6: Construct the matrix of the obtained IMF
$M \in R^{J\times N}$: $M = [h_1; h_2; \cdots h_j]$
7: Obtain the denoised signal $X^b = \sum_{j=1}^{J} h_j^b$
8: Detect the peak of the dominant frequency component corresponding to the vital sign
9: Estimate the vital the using the detected dominant peak frequency F_d
10: end for
12: return estimated vital sign

Figure 2 depicts the corrupted PPG signal acquired during physical exercise, and Fig. 3 indicates the clean PPG signal obtained using the GSMD algorithm.

Vital sign estimation

HR estimation

Estimating HR using a PPG signal is a standard non-invasive method for realtime monitoring of an individual's HR. The number of heartbeats per minute (BPM) is termed HR. HR is an essential vital sign that reflects the rate at which the heart pumps blood throughout the body. Generally, the average HR for adults lies between 60 and 100 BPM, although the HR range varies based on individual factors such as age, fitness level, and medical conditions [15]. Daily physical activity, stress, emotions, temperature, and health can

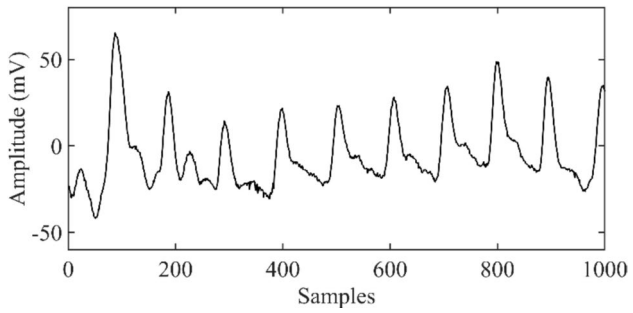


Fig. 2 Motion artifact corrupted PPG signal of subject 15 taken from IEEE SPC dataset

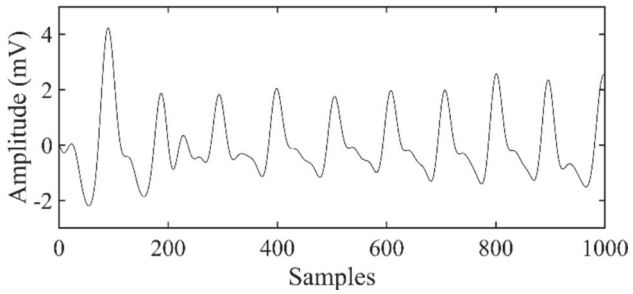


Fig. 3 The representation of a clean PPG signal was obtained using the GSMD algorithm

influence HR. Continuous monitoring of HR plays a pivotal role in healthcare, sports, and fitness as it provides valuable information about the cardiovascular system and indicates an individual's health status. HR can be monitored by manually counting pulse beats at a certain arterial point. Recently, electronic devices embedded with PPG sensors have been commonly used for realtime continuous monitoring of HR. Tracking changes in HR in realtime can help assess cardiac health, track fitness progress, and identify abnormalities in cardiac function. The clean PPG signal obtained from the GSMD algorithm is used to estimate the HR. Estimating HR from a clean PPG signal using the Fast Fourier Transform (FFT) to identify the dominant peak frequency is a common and effective method [16]. FFT transforms the time-domain signal into its frequency-domain representation. The output is a complex spectrum that contains information about the signal's frequency content. Signal transformation helps identify the signal's different frequency components, including the HR frequency. Further, to identify peaks, the magnitude spectrum of the complex spectrum is computed, which provides the distribution of signal energy across different frequencies. The obtained magnitude spectrum is examined to locate the potential peaks. Examine the magnitude spectrum to locate local maxima. These represent potential peaks in the signal. One common approach is to iterate through

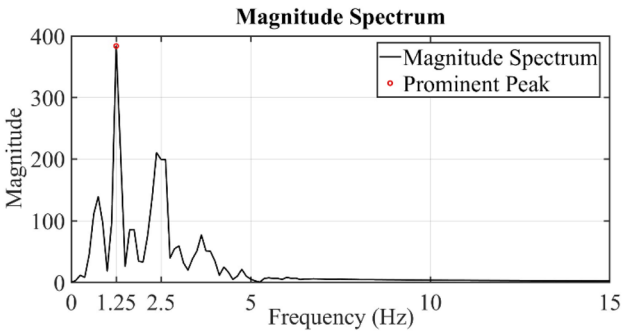


Fig. 4 Magnitude spectrum representation of clean PPG signal to extract dominant peak frequency

the magnitude spectrum and compare each point with its neighboring points. If the magnitude of the point is greater than its neighbors, it can be considered a local maximum. Figure 4 shows the magnitude spectrum of the clean PPG signal with a.

red marker indicating the frequency of the prominent peak corresponding to HR. After identifying potential peaks, the frequency bin of the peak with the highest magnitude in the FFT spectrum is determined. The frequency of the prominent peak is used to compute HR in BPM using the relation. Since FFT output is in Hz, thus converting the peak frequency from Hz to HR in BPM, Eq. (1) is used.

$$HR\,(BPM) = prominent\,peak\,frequency \times 60 \tag{1}$$

RR estimation

RR indicates the number of breaths a person completes in one minute. RR is commonly used in clinical assessments to assess the individual's respiratory health status. For adults with good health status, the value of RR typically lies between 12 and 20 breaths per minute (BPM). However, factors like age, physical condition, and health conditions cause RR to deviate from its normal value. Table 2 depicts the variation in RR with age.

Realtime monitoring of RR is vital in healthcare settings, as it provides valuable information about individuals' lung and cardiovascular functions [17]. RR deviation from its standard threshold value, or abnormal RR, can indicate forthcoming illness. Monitoring changes in blood volume during inhalation and exhalation can provide valuable information about RR. During inhalation, the pressure within the thoracic cavity decreases, increasing intrathoracic volume [18]. This results in an influx of blood into the thoracic cavity, increasing the venous blood return to the heart.

Conversely, during exhalation, the pressure within the thoracic cavity increases, leading to a decrease in intrathoracic volume. The increase in thoracic pressure impedes the return

Table 2 Variation in RR with change in age

Age	Rate (BPM)
Birth to 6 months	30–60
6 months to 1 year	30–50
1–3 years	24–40
3–5 years	22–34
5–12 years	16–30
12–70 years	12–20
> 70 years	15–20

of venous blood to the heart. This concept is often used to estimate RR using a PPG signal. Respiratory rate estimation using a PPG signal is common in healthcare and wearable devices. The PPG waveform reflects these variations as blood volume changes due to inhalation and exhalation [19].

PPG signal captures the volumetric change in the blood due to respiration; thus, the acquired PPG signal consists of periodic variations associated with RR. Thus, to identify the dominant peak frequency corresponding to the RR, the proposed work used FFT-based dominant peak detection techniques to estimate RR using the PPG signal. Figure 5 shows the magnitude spectrum of the clean PPG signal with a red marker indicating the frequency of the prominent peak corresponding to RR.

After identifying potential peaks, the frequency bin of the peak with the highest magnitude in the FFT spectrum within the RR frequency range is determined. The frequency of the prominent peak is used to compute the RR in breath per minute (BPM) using the relation. Since FFT output is in Hz, Eq. (2) converts the peak frequency from Hz to RR in BPM.

$$RR\,(BPM) = prominent\,peak\,frequency \times 60 \tag{2}$$

BP estimation

The force of blood exerted on the artery wall during the systolic and diastolic phases of the heart is termed BP. BP is typically expressed in pressure values, namely systolic BP (SBP) and diastolic BP (DBP). The pressure in the arteries when the heart beats and pumps blood into the circulation is noted as SBP, which is usually the first and highest number in a BP measurement. Conversely, the pressure in the arteries when the heart is at rest between beats is termed DBP. DBP reflects the constant pressure in the circulatory system. BP is measured in millimeters of mercury (mmHg) and is an essential indicator of cardiovascular health. Any deviation in the threshold BP levels can indicate abnormal conditions of individual health. The potential of realtime monitoring of BP using PPG signals has attracted significant interest in the healthcare industry [20].

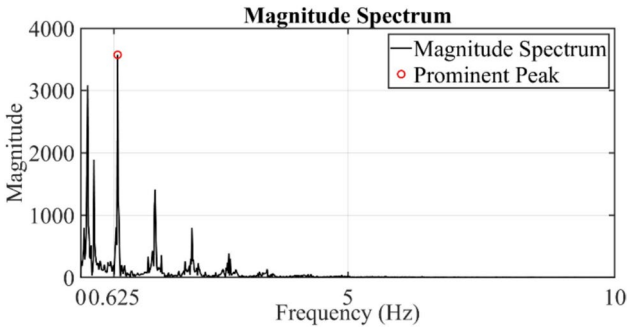


Fig. 5 Magnitude spectrum of clean PPG signal for extracting dominant peak frequency

This interest is driven by the need for continuous monitoring of BP in realtime, which is essential in diagnosing and managing cardiovascular diseases. The gold standard for monitoring BP is the arterial BP (ABP) technique. ABP is considered typical for its accuracy and reliability. However, the invasive nature limits the ABP use in the hospital. Traditionally, cuff-based mercury sphygmomanometers have been the standard for BP measurement; however, they are not suitable for continuous monitoring due to discomfort and the time it takes to obtain a reading [21].

In contrast, a cuffless, non-invasive monitoring system has emerged to resolve the above issue, allowing individuals to monitor their BP continuously. PPG and ECG are promising techniques when used with artificial intelligence (AI) for monitoring and diagnosing cardiovascular diseases (CVD). PPG is particularly favored due to its affordability, ease of use, portability, and suitability for wearables. As a result, there is growing interest in leveraging PPG signals and AI algorithms to develop innovative solutions for continuous BP monitoring and early detection of cardiovascular issues, ultimately improving the management of cardiovascular health.

Recent research has been focused on developing single-sensor-based continuous BP monitoring solutions to overcome the limitations of multiple sensors. This approach aims to provide convenience and comfort to users while maintaining accuracy. There is no direct measure to monitor SBP and DBP solely from a PPG signal. PPG signals offer valuable information about blood volume changes and pulse waveforms, but they require careful calibration and predictive models to estimate BP accurately [22].

The proposed work combines the potential of the GSMD algorithm with a linear regression-based machine learning model to create a realtime BP monitoring device. Linear regression is a relatively lightweight machine-learning algorithm that can provide reasonable results while being computationally efficient. The clean PPG signal obtained from the GSMD algorithm is processed to extract features for model training. The extraction and selection of highly

correlated features is crucial for extracting SBP and DBP information from the PPG signal. Initially, a subset of features, namely absolute deviation (MAD), entropy, power spectral density (PSD), dominant peak frequency (Freq), the index of the maximum value (*midx*), mean and standard deviation (std) within each window is extracted from the clean PPG signal obtained using the GSMD algorithm.

The clean PPG signal makes extracting reliable features accurately representing SBP and DBP monitoring easier. These features help train a model to learn the relationships between extracted features and ground-truth BP labels, leading to better model performance. Moreover, machine learning models trained on features extracted with a clean PPG signal improve their generalized capability. However, a model trained on a corrupted PPG signal may learn to adapt to noise rather than the underlying physiological patterns related to BP.

Further, the filter method selects the best feature with a high correlation with the target SBP and DBP values. Filter methods are feature selection methods that choose features based on some criteria before building the model. The filter method is computationally efficient and does not require multiple-model training. Also, the filter method is less prone to overfitting in the feature selection phase. The proposed work uses a correlation coefficient and the chi-square-based feature filtering approach. The correlation coefficient filter method filters out irrelevant features from the features set by evaluating the correlation between the input feature and the target BP value. This step retains the highly relevant feature strongly correlated with the target value. Based on the analysis with different features, a threshold value of 0.6 is fixed in the proposed work—the feature whose correlation value is above 0.6 is considered relevant. Additionally, the chi-square method is performed to filter the feature set. A contingency table is designed between the individual feature and the target BP value to obtain the *p*-value using the chi-square test for feature filtering. This Table reflects the frequency of each combination of features and the target BP value. The chi-square test is implemented on the contingency table to obtain the *chi*-square statistic and the *p*-value. The *p*-value represents the significance between the feature set and the target BP value. The feature that obtains a *p*-value less than or equal to 0.05 indicates that the feature is likely to be informative in predicting the BP value. Based on the filtering method, *midx* is identified as the best feature, which correlates highly with the target BP value (Fig. 6). This feature obtains the optimized linear regression equation for predicting SBP and DBP.

Initially, a grid search optimization algorithm is used to find the optimized value of the linear regression model's hyperparameters a and b, where *a* is the intercept and *b* is the slope of the PPG feature. The optimization algorithm likely controls the coefficients of the linear equation used to predict

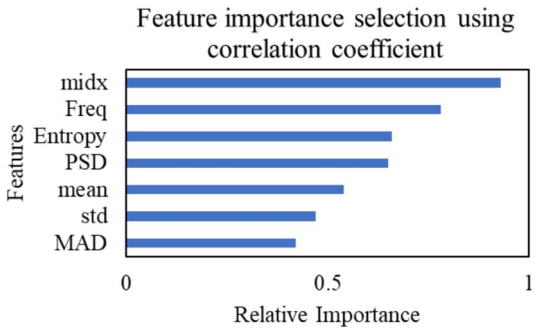


Fig. 6 Feature and its importance in estimating BP

SBP and DBP from PPG features. Further, to improve the generalized capability of the model, a fivefold cross-validation approach is used to split the input data into five subsets.

To evaluate the model efficiency with different values of hyperparameters *a* and *b*, mean absolute error (MAE) is used as a performance metric. During each fold, the model iterates through different combinations of hyperparameters *a* and *b* within the defined search range. For each combination of hypermeters, the proposed framework trains the linear regression model using the training data. After training, the trained model predicts the SBP and DBP values for the testing data. The MAE parameter is computed based on the predicted and actual values of SBP and DBP. The value of *a* and *b* at which the model obtains the minimum MAE value is filtered as an optimized hypermeter for that combination of hyperparameters and fold. The obtained value of MAE for each combination of fold and the optimized hyperparameters is kept in a matrix for comparison. After completing fivefold cross-validation, the framework finds the best combination of hyperparameters that provides the lowest MAE across all folds. This combination of a and b is considered an optimized hyperparameter setting for the model. Once the framework obtains the optimized hyperparameters, the framework trains the final linear regression model using the entire dataset. The obtained final linear regression is based on an optimized hyperparameter using a fivefold cross-validation approach intended to be the best-performing model. After training with the optimized hyperparameters, the linear regression model provides a linear equation. This equation represents the correlation between PPG features and the predicted SBP and the DBP value.

This study uses linear and nonlinear regression as data-driven models to predict the PPG signal's SBP and DBP values. The linear regression model is a lightweight machine learning model form of regression used to predict the SBP and DBP values. Linear regression is a supervised machine learning algorithm that predicts SBP and DBP values based on the extracted features. Linear regression assumes a linear relationship between the input features and the true SBP

and DBP values. The coefficient of linear regression is optimized, and the best-fit line equation that minimizes the sum of squared differences between the estimated and true values is computed. Conversely, nonlinear regression can capture a more complex relationship between the estimated and the true value. The complexity of the nonlinear regression model is higher than the linear model, as nonlinear regression requires more parameter tuning.

SpO2 estimation

Blood oxygen saturation (SpO2) is an important physiological parameter that measures the percentage of oxygen-bound hemoglobin in the blood. SpO2 is an important parameter that checks an individual's respiratory and circulatory status; thus, it is commonly used in healthcare for patient monitoring. SpO2 measurement helps assess how well oxygen is transported to the body tissues. Change in SpO2 value about its threshold can be an early indicator of abnormality in the respiratory, cardiovascular, and circulation systems [23]. A pulse oximeter is a device that captures the interaction between light and the volume of blood. It is commonly used in households and clinics to measure SpO2. The pulse oximeter consists of two light-emitting diodes (LED) of different wavelengths. Red and infrared LEDs are commonly used in pulse oximeters to measure the light absorbed by oxygenated (HbO2) and deoxyhemoglobin (Hb) hemoglobin (*Hb*). A pulse oximeter computes the SpO2 using the relation shown in Eq. (3) by analyzing the concentration of light absorbed by oxygenated and deoxygenated hemoglobin.

$$SpO2(\%) = \frac{C(HbO_2)}{C(HbO_2) + C(Hb)} \times 100 \tag{3}$$

In a healthy body, the typical value of SpO2 typically lies in the 95% to 100% range. A SpO2 value lower than 95% indicates the body has a low oxygen level in the blood, termed hypoxemia. Untreated hypoxemia can result in respiratory diseases and heart problems. Moreover, if not treated promptly, a deficient value of SpO2 can endanger life. Thus, continuous monitoring of SpO2 in realtime can help the physician to diagnose patients with chronic respiratory conditions. The proposed work possesses a single PPG sensor to monitor SpO2. A Hilbert transform is applied to the clean PPG signal to compute its envelope. The envelope represents the variation in the PPG signal amplitude due to a change in blood flow. Further, the envelope's alternating (AC) and direct (DC) components are obtained to compute the *R*-ratio. The AC component indicates the variation in amplitude due to pulsatile blood flow and is calculated as the difference between the envelope's maximum and minimum values. The mean value of the envelope depicts the DC component, representing the baseline or non-pulsatile part of the signal. The *R*-ratio depicts the ratio of AC components to DC components.

The state-of-the-art techniques indicate that the R-value-based computation of SpO2 does not always provide reliable outputs. Thus, classical signal processing approaches fail to provide accurate and reliable solutions. The advancement in the machine learning model shows the potential to overcome the drawbacks of traditional signal processing methods to estimate SpO2 [24]This work estimates the SpO2 value using three simple machine learning-based regression models: linear, nonlinear, and polynomial. Equation (4) represents the optimized calibration curve formula obtained using polynomial regression to estimate the SpO2 value. Equation (4) relates the SpO2 value with the *R*-ratio.

$$SpO_2 = (-45.060 \times R^2 + 30.354 \times R + 94.845) \tag{4}$$

The derived formula is validated on a separate dataset to assess its accuracy and generalizability.

Post-processing stage

A post-processing stage is introduced in this work to enhance the accuracy and reliability of the algorithm. This stage leverages the inherent temporal variation of vital signs. Specifically, when subjects engage in intense physical exercise or when the frequency of motion artifact components is closely aligned with that of the true vital sign component, the post-processing stage aids in tracking the vital sign within the current window. This improvement enhances the overall performance of the proposed algorithm, making it more robust in challenging situations where motion artifacts or vital sign variations might otherwise pose difficulties. The clean PPG signal extracted from the based algorithm extracts the vital signs despite directly considering the obtained vital sign values computed using a clean PPG signal provided

by the GSMD algorithm. A vital sign tracking algorithm is used to refine the value of the vital sign. This step involves the proposed framework to improve the reliability of the proposed algorithm in realtime. A signal quality index (SQI) parameter qualifies the PPG signal windows as clean/corrupted to classify the correct/incorrect vital sign prediction. The proposed work initially uses the GSMD to denoise the PPG signal. The denoise PPG signal is passed to the vital sign estimation algorithm to estimate the vital sign. Four SQIs, namely, entropy, signal-to-noise ratio (SNR), skewness, and zero-crossing parameters, are analyzed to check the reliability of the estimated vital signs.

A unified SQI parameter consolidates individual SQI parameters into a singular value to enhance signal quality analysis. Before consolidation, all SQI parameters undergo normalization to unify them onto a common scale. The resultant combined SQI index offers a comprehensive perspective of the PPG signal derived from the collective evaluation of various SQI indexes within the signal. Different weights are assigned to each normalized SQI parameter. A higher weight is assigned to that SQI parameter, which provides significant information regarding motion artifacts in a PPG signal. After weight assignment, all the SQI parameters are combined to give a single composite SQI metric. For the algorithm to flag the signal quality as acceptable or not acceptable, an optimum threshold value for the composite SQI metric is assigned. The selection of the threshold value is fixed based on the performance of the proposed algorithm on different threshold values. Suppose a PPG signal window SQI index parameter is less than the specified threshold value. In that case, it indicates that the PPG signal window is affected by strong motion artifacts, and the vital sign value obtained with such a window is erroneous. A vital sign smoothing algorithm is used to refine the vital sign value of such corrupted windows. The PPG signal window flag is not acceptable. The vital sign value of that window is replaced using a not-valid value (*NaN*). The *NaN* value is considered an outlier in the vital sign, which is replaced by the smoothing algorithm for accurate vital sign estimation. A linear interpolation protocol is used to replace the *NaN* value. The steps involved in the smoothing algorithm are depicted in Algorithm 2.

Algorithm 2: Proposed SQI index-based Smoothing Algorithmusing History Tracking of estimated vital sign

Input: GSMD-based motion artifact-free PPG signal window of length N samples.

$$X^b = \left[X^1, X^2, X^3, \ldots\ldots X^Z\right]$$

Result: Refine estimated vital sign value using the smoothing algorithm.

1: Compute the composite SQI index parameter

$$SQI_{comp} = f(x)$$

Where, $f(x)$ is a function that computes the SQI of the PPG signal

2: Define the optimum threshold value (τ)

3: If composite SQI ≥ τ, $vitalsign_{estimated} \leftarrow vitalsign_{refined}$ Flag estimated vital signs acceptable and jump step 8; otherwise, follow steps 3 to 7.

4: Flag the obtained vital sign value as not valid value (*NaN*)

If $SQI_{comp} \leq \tau$, the estimated vital sign is flagged as not valid: $vitalsign_{estimated} \leftarrow NaN$

Use the interpolation to replace the *NaN* value

5: Retrieve the non-NaN value from the original vital sign data.

$vitalsign_{valid} = \{\text{vitalsign}_i \in vitalsign_{estimated} | vitalsign_i \neq NaN\}$

6: Omit *NaN* values by applying the *NaN* handling function:

$$vitalsign_{valid} = omitNaN(vitalsign_{estimated})$$

7: Specify the window size to implement the moving average function to smooth the signal.

$$vitalsign_{refined} = \frac{1}{q}\sum_{i=k}^{k+q-1} vitalsign_{estimated}, \quad \text{for } k = 1, 2, \ldots. N\text{-}q+1$$

Where *NaN* values are omitted

8. Process the signal as a refined estimated vital sign.

9: end for

10: return vital signs estimated

In those windows where *NaN* represents the corrupted estimated vital sign value, the *NaN* value is replaced by the mean of the estimated vital sign value of the previous three windows. Each window is checked for its SQI value. If the SQI value of a particular window is below a specified threshold value, then the estimated value is discarded, and vital sign tracking is used.

This work analyzes the PPG signal using the GSMD approach using four SQI indices to determine accurate vital sign prediction based on the SQI index. If a window SQI index is reported lower than the specified threshold, then that window is passed to the tracking algorithm to refine the vital sign value.

Results analysis

This section is divided into five subsections. No work has been reported to estimate all vital signs using a single framework. Thus, to analyze the proposed framework's performance, subsequent sections are designed to compare the results of an individual vital sign with state-of-the-art work.

Performance analysis of vital sign estimation

The performance analysis of the proposed vital sign estimation framework is visualized and analyzed using a violin plot. In a violin plot, the width of the plot at a location indicates the density of data points. The violin plot contains both mean and median lines. The difference in the position of mean and median values between the two groups (vital sign estimated and vital sign true) provides insights into the average performance of the proposed framework. Figure 7 shows the violin plot comparing the estimated and true value of various vital signs. The shape, mean, and median values of true and estimated vital signs are almost identical, thus indicating the potential of the proposed framework that estimates multiple vital signs using a single algorithm. The proposed framework is a computationally efficient algorithm that can be embedded in a wearable device for realtime monitoring of vital signs.

The proposed work is the first framework in the literature to estimate four vital sign values simultaneously. The proposed framework allows individuals to select the vital sign parameter. The primary concern in designing a wearable health monitoring device is the accuracy of results in realtime. Monitoring vital signs in realtime faces the hurdle of motion artifacts, which affect the accuracy of the vital signs; thus, analysis becomes erroneous. In the proposed work, the motion artifact components are separated by using a computationally efficient GSMD algorithm. GSMD provides a clean PPG signal component. To critically analyze the performance of the proposed framework, besides the violin plot, the Pearson correlation coefficient (r) and the mean absolute error (MAE) are computed for each vital sign, represented by Eqs. (5) and (6), respectively.

$$r = \frac{n\left(\sum xy\right) - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right]\left[n\sum y^2 - \left(\sum y\right)^2\right]}} \tag{5}$$

$x \rightarrow vitalsign_{true}\ y \rightarrow vitalsign_{estimated}$

$$MAE = \frac{1}{W}\sum_{b=1}^{W} \left|vitalsign_{true}(b) - vitalsign_{estimated}(b)\right| \tag{6}$$

Result analysis of HR estimation

Continuous HR monitoring via PPG-equipped wearables presents challenges regarding precision in HR estimation.

The proposed algorithm reliably attains superior accuracy when the subject indulges in various physical activities. The proposed algorithm is trained and tested using the IEEE SPC dataset recorded during the subject undergoes physical activity closer to realtime daily routine activities. The algorithm's performance is assessed by measuring the MAE (BPM), violin plot, and the Pearson correlation plot. The proposed framework has shown an MAE of 1.28 BPM for IEEE SPC datasets. Table 3 compares the results of the proposed method with those of recently published algorithms. Notably, the MAE value outperforms those of other state-of-the-art methods.

The violin plot in Fig. 8 shows the difference between $HR_{estimated}$ and HR_{true} values. As observed from Fig. 8, the maximum density of error point (estimated—true) lies within the range of − 2.5 BPM to 2.5 BPM.

Additionally, Fig. 8 shows that most data points are within the small error range, indicating that the difference in errors within the HR range is minimal and remains within an acceptable threshold. The experimental outcomes demonstrate that the proposed algorithm effectively performs when tested with datasets collected while subjects are engaged in vigorous physical activities, such as weight lifting and boxing.

Result analysis of BP estimation

The mean error (ME) and the standard deviation (SD) of the error are computed for both SBP and DBP values to check the performance of the BP estimation algorithm. The obtained values of ME and the SD are checked to find if the values of ME and SD suit the criterion provided by the Association for the Advancement of Medical Instrumentation (AAMI) and the British Hypertension Society (BHS). The AAMI standard assesses compliance based on whether the ME in the performance of both SBP and DBP is below five mmHg, along with the requirement that the SD of the error remains under eight mmHg. The BHS standard criterion to check the performance of the framework is summarized in Table 4. A BP estimation device must achieve B-grade for SBP and DBP per BHS standards. The performance of the proposed framework is validated using data from 1557 subjects with 3610 windows. The results are compared against the BHS standard. The optimized linear model equation for which the model shows the lowest MAE is considered. The equation of the linear regression model performed in MATLAB® software and the structure of the SBP and DBP model is as follows:

Linear regression equation:

$$\mathrm{SBP} = 113.0048 + 3.5783 \times \mathrm{PPG_feature}$$

$$\mathrm{DBP} = 74.1955 \pm 0.13513 \times \mathrm{PPG_feature}$$

Using the proposed algorithm, 81.29% of windows pose an MAE below 5 mmHg for SBP, while 86.81% pose an MAE below 5 mmHg for DBP. Comparing the above values with the BHS standards from Table 5 confirms that the requirements for Grade A in the BHS standard are met. Further, the framework resulted in MAE values of 2.56 and 3.25 mmHg for SBP and DBP, respectively, thus meeting the requirements of the AAMI standard. These MAE values for both SBP and DBP are within the range of AAMI standards. These findings highlight the effectiveness and reliability of the proposed model for BP estimation. The performance of the proposed framework to estimate BP is evaluated using different performance assessment measures such as MAE, Bland–Altman plot, and violin plot. The MAE values of 2.46 and 2.56 on training and testing databases, respectively, are obtained. Further, for DBP, MAE values of 3.19 and 3.25 on training and testing databases, respectively achieved using the linear regression model, demonstrate the superiority of the proposed model over the other data-driven models. The results show that using GSMD, the performance of the model training is improved using features extracted from the clean PPG signal for BP estimation.

Table 6 compares the performance of the proposed work with related state-of-the-art work. For a fair comparison, Table 6 consists of the same performance metrics and database. The proposed study, which used the combination of GSMD and linear regression to predict SBP and DBP, achieved better performance than the related state-of-the-art studies found in the literature.

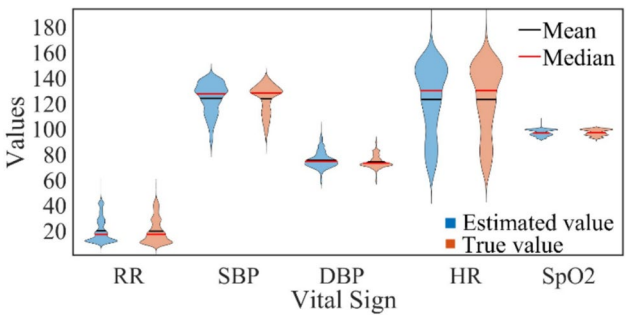


Fig. 7 Violin plot representation of the estimated vital sign with true vital sign

Result analysis of RR estimation

The MAE in breaths per minute (BPM) of the estimated RR ($RR_{estimated}$) is calculated to assess the performance of the proposed GSMD-based RR estimation method. Table 7 compares the MAE in BPM with various techniques from

Table 3 Comparison of MAE values estimated using the proposed technique with other existing methods

Method	Dataset	Sensor used	Algorithm	MAE (BPM)
Kong 2019 [25]	IEEE SPC training	PPG + 3-axis Accelerometer	VFCDM + Spline Regression	2.91
	Chon Lab			2.97
PPGNet 2020 [26]	IEEE SPC training	PPG	Inception + LSTM	3.36
	IEEE SPC Test			12.48
BinCor Net 2020 [27]	IEEE SPC training	PPG	Bin. CNN + LSTM	6.78
	IEEE SPC Test			7.32
Wilkosz 2021 [28]	PPG-Dalia	PPG + 3-axis Accelerometer	CNN + LSTM	6.28
Deep Heart 2021 [29]	IEEE SPC training	ECG + PPG	DnCNN	1.61
FDM 2022 [16]	IEEE SPC	PPG	FDM	1.87
Deep Pulse 2022 [30]	IEEE SPC training	PPG + 3-axis Accelerometer	TCNBest	2.76
	IEEE SPC Test			5.05
	BAMI-II			2.38
Ismail 2022 [31]	IEEE SPC Training + Test	PPG	Convolutional recurrent regression	2.41
Hankel 2023 [32]	IEEE SPC Training + Test	PPG + ACC	Hankel decomposition	1.86
FrFT [5]	IEEE SPC Training + Test	PPG	FrFT	1.88
VADAF [33]	IEEE Dataset	PPG + ACC	Adaptive filter	1.48 bpm
ASC [34]	IEEE SPC Training	PPG + ACC	Adaptive Spectrum Correction	0.88 BPM
Aditta [21]	PPG-BP dataset	PPG	Linear regressor	3.17
Proposed work	IEEE SPC	PPG	GSMD	**1.28 BPM**

Bold value indicate the result obtained with the proposed work

the literature. Additionally, the performance of the GSMD-based RR estimation method from the PPG signal is assessed using the Bland–Altman plot and Pearson correlation coefficient. Table 7 reveals that MAE values obtained with the proposed approach are lower than in previous works. The correlation between RR_{true} and $RR_{estimated}$ obtained with the proposed framework is visualized in the violin plot, as shown in Fig. 8.

The LOA for the signals from the Capnobase dataset ranges from − 2.1 to 2.3 BPM. As evident from Fig. 8, the error differences are minimal and within acceptable limits, suggesting that the proposed RR estimation method exhibits good agreement with the true data and can be used for real-time RR monitoring.

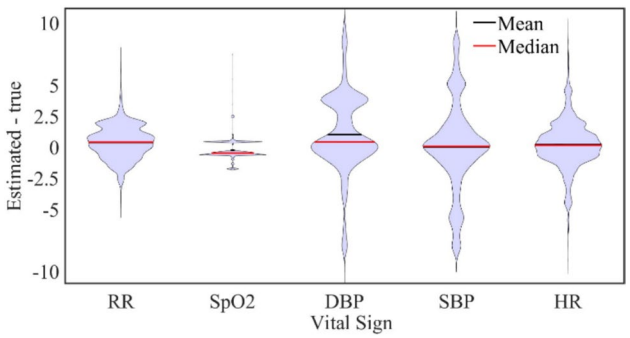


Fig. 8 Violin plot representation for vital sign estimation

Result analysis of SpO2 estimation

Reflectance PPG data collected using the SpO2 data acquisition platform used as the true value is synchronized with the acquired PPG signal. Three regression machine learning models, namely, linear, nonlinear, and polynomial regression models, are exclusively tested on the test data. Nonlinear and polynomial regression can capture nonlinear relationships between the feature and the true BP value. However, nonlinear regression models are not restricted to polynomial functions and can employ a wider range of nonlinear forms to describe the relation.

Polynomial regression, an extension of linear regression, specifically uses polynomial functions to model the relationship between the PPG feature and the true value of BP. The proposed work's polynomial regression demonstrated commendable performance over the remaining two. Table 8 indicates that the proposed work outperforms similar state-of-the-art work. From Fig. 8, it can be observed that most correctly predicted observations have a lower error rate in estimating the SpO2 value.

Discussion

Monitoring continuous vital signs using PPG-enabled wearable devices presents ongoing challenges regarding accuracy. The proposed novel framework consistently delivers

Table 4 Performance comparison of the proposed framework with BHS standard

	The cumulative absolute difference between the true and estimated percentage		
	≤ 5 mmHg (%)	≤ 10 mmHg (%)	≤ 15 mmHg (%)
BHS standard			
Grade A	60	85	95
Grade B	50	75	90
Grade C	40	65	85

high accuracy for vital sign tracking, even during intense physical activity. The framework performance is evaluated through the dataset available publicly for individual vital signs. Its effectiveness is assessed by calculating the MAE and analyzing the violin plot. Tables 3, 6, 7, and 8 compare the results of the proposed work with the state-of-the-art work for individual vital signs such as HR, BP, RR, and SpO2, respectively. The last row of Tables 3, 6, 7, and 8 highlights the results of the proposed algorithm. Among the previously reported algorithms, the proposed algorithm for estimating vital signs outperforms most of the previously reported methods. Based on the results obtained, the proposed framework is a promising solution for predicting vital signs. In Table 3, the last column represents the MAE values achieved with the proposed approach. Table 3 reveals that the MAE value obtained with the proposed approach is small compared to previous works. However, it's important to note that the works in references [25, 26] necessitate an additional 3-axis accelerometer sensor to provide a reference motion artifact signal. This requirement increases the overall system complexity and energy consumption. The performance of the approach presented in reference [29] relies on the simultaneous use of an ECG signal alongside the PPG signal. The inclusion of an additional ECG sensor increases the system's complexity. In contrast, the proposed PPG signal denoising technique, GSMD, stands out as it doesn't require an additional reference signal. Thus, the GSMD-based algorithm reduces the overall system complexity and energy consumption, making it a more practical and efficient choice for real-time HR estimation. In Table 6, the work proposed in [20, 35, 37, 39, 41, 42] extracted features from ECG and PPG signals. These methods show notable MAE for SBP and DBP estimation, but the need for additional

Table 5 Performance comparison of SBP and DBP estimation for training and test dataset

Performance metrics	Training dataset				Test dataset			
	Linear regression		Non-linear regression		Linear regression		Nonlinear regression	
	SBP	DBP	SBP	DBP	SBP	DBP	SBP	DBP
Pearson Correlation (*r*)	0.9823	0.9911	0.9675	0.9691	0.9826	0.9881	0.9612	0.9646
MAE	2.46	3.19	3.26	3.92	2.56	3.25	3.64	4.12

Table 6 Details of performance evaluation metrics for the data-driven models

Model	Number of subjects	Engineered feature	MAE (mmHg)	
			SBP	DBP
Classical ML [35] 2017	MIMIC-II, 1000 subjects	Yes, PTT-based features	11.17	5.35
Regression tree [36] 2018	Queensland 32 subjects	Yes, three PPG features, less subjects	4.82	3.25
ELM [37] 2018	MIMIC-II 4254 subject	Yes, ECG and PPG signal features	4.25	3.95
ResNet [38] 2019	MIMIC-III 510 subjects	Yes, but it does not show the BHS standard	9.43	6.88
GA-SVR [39] 2019	MIMIC-III 772 subject	Yes, fourteen features from ECG and PPG signals	3.27	1.16
Harfiya [40] 2021	MIMIC 5289 subject	PPG, PPG', PPG"	4.05	2.41
Sakib [41] 2022	MIMIC 1710 subject	ECG, PPG, PPG', PPG"	2.33	2.73
Rastegar [20] 2023	MIMIC	ECG, PPG features	1.23	3.08
Qiu [42] 2023	MIMIC 53	Feature extracted from ECG and PPG signals	3.65	4.56
Nour [43] 2023	MIMIC	24 feature statistics from PPG	3.069	1.72
Qin [44] 2023	MIMIC	Multi-scale feature extraction	5.98	3.24
Aditta [21] 2024	PPG-BP	Linear Regression	4.75	3.34
OVAR BP-NET [22] 2024	MIMIC	Pulse wave deep learning	7.54	5.31
Proposed work	MIMIC	Dominant peak index	2.56	3.25

Table 7 Comparison of MAE of the proposed method with existing methods

Work	Signal used	Algorithm	MAE (BPM)
Zhang [45] 2017	PPG	M-FOCUSS	0.82
Zhu [46] 2018	ECG and PPG	Bayesian fusion model	1.89
Motion [47] 2019	PPG	EEMD	2.4
Zhu [48] 2021	PPG	Bayesian approach	1.18
Stankoski [49] 2022	PPG	ML model	1.38
Zhao [50] 2023	ECG and PPG	TransRR	1.20
Yang [51] 2024	PPG	Deep learning	1.15
Chi [52] 2024	PPG	Deep learning	1.9
This work	PPG	GSMD	0.59

ECG sensors increases the computation complexity of the device. The work proposed in [36] shows an acceptable performance through the AAMI and BHS standards. However, further improvements in MAE and the model's reliability can be improved by including more data. The work proposed in [38] does not satisfy the BP estimation standard. The work proposed in [40] uses three PPG signal variants to extract features for model training. Maintaining a correlation between the PPG signal and its first and second derivatives is challenging in realtime. Violin plots of SBP and DBP are shown in Fig. 8, indicating that the difference in errors within the BP estimation is minimal and remains within an acceptable threshold. The work proposed in [53] segregates the complete range of SpO2 values into twenty classes. Instead of making predictions, the proposed model does the classification task. The MAE is calculated based on the prediction of the classifier. The work proposed in [54] shows an acceptable performance. However, statistical research has indicated that further improvements can be made to the model by increasing the data and enhancing the selectivity of the variables. The work proposed in [55] uses red, IR, and green LEDs. SpO2 is determined by processing the red and IR PPG signals captured from the wrist, and the green PPG signal is used as a reference to determine the window size for the moving average filter. The green color aids in eliminating noises present in the signal. In [24], 107 features are extracted to train the different machine learning regression models, thus having a large computational complexity. The performance of the work proposed in [54] relies on the additional reference signal, increasing the model's computational complexity. HRV estimation using PPG signals is critical for assessing autonomic nervous system health but is often impeded by inaccuracies in interbeat interval detection and motion artifacts. The proposed framework enhances HRV analysis through GSMD. GSMD-based denoising refines the PPG signal, allowing accurate detection of IBIs even in real-time environments. Unlike traditional methods, which require additional reference signals, the proposed framework uses only PPG signals, reducing device complexity while achieving competitive accuracy. RR estimation from PPG signals poses challenges due to the weak respiratory components often masked by noise and motion artifacts. The proposed framework addresses these challenges by leveraging the GSMD algorithm to decompose the PPG signal and isolate low-frequency respiratory components. The GSMD's adaptive decomposition capability automatically tunes the bandwidth, ensuring optimal isolation of respiratory signals and eliminating the need for manual parameter adjustments.

Table 8 Result comparison of SpO_2 obtained using the proposed framework

Work	Signal used	Algorithm	MAE (BPM)
Venkat [53] 2019	Red and IR signals PPG 95 subjects	ML model	0.50
Vijayarangan [54] 2020	Red and IR signals PPG 28 subjects	ML model	1.81
Berwal [55] 2022	Red, IR, and Green PPG signal	LED peak detection	1.0
Zhong [23] 2023	Red and IR signals PPG 33 subject	Cascade of unsupervised and supervised regression	1.7
Shuzan [24] 2023	PPG signal	Gaussian process	0.57
Koteska [56] 2024	PPG signal	Deep neural network	1.13
Zhang [57] 2024	PPG signal + reference signal	Adaptive filter	0.57
Proposed work	PPG signal	EEMD + polynomial regression	0.45

The GSMD-based vital sign estimation results obtained from corrupted PPG signals show that estimated vital signs are highly correlated with the true value of vital signs in the presence of noises and motion artifacts. A post-processing stage is employed to enhance the accuracy and reliability of the proposed framework. Based on the natural variation property of the vital sign concerning the time, the post-processing stage is used to refine the value of the vital sign.

The proposed framework's evaluation is primarily based on publicly available datasets, which may not encompass all possible variations in user demographics (e.g., age, skin tone, and physiological conditions). These datasets are recorded under specific demographic conditions. The proposed framework primarily evaluates performance during moderate physical activities. Its applicability to scenarios involving vigorous exercises or outdoor environments, where motion artifacts and environmental noise are more significant, remains to be explored. Further, the proposed framework validation on diverse and larger datasets is needed to generalize the framework's applicability. Further, collaborations with medical institutions will also be pursued to validate the framework clinically on patients with cardiovascular diseases, establishing its reliability as a comprehensive tool for healthcare monitoring in diverse and challenging real-world conditions. To address the limitations of the study and further advance the study, future work will focus on evaluating the framework performance using diverse and larger realtime acquired datasets encompassing a wide range of user demographics, including variations in age, skin tone, and physiological conditions. Additionally, experiments should be conducted in realtime scenarios involving vigorous physical activities and outdoor environments to assess the framework's robustness against motion artifacts and environmental noise. Collaboration with clinical experts to validate the framework on clinical-grade data and exploring its integration with wearable devices for realtime monitoring would significantly expand its applicability and practical use.

Conclusion

This work introduces a health monitoring framework design that estimates HR, BP (SBP, DBP), RR, and SpO2. The framework presents the GSMD algorithm to clean the motion artifacts corrupted PPG signal in this work. This framework leverages a single PPG signal and a computationally efficient signal processing algorithm for continuous vital health monitoring. In the proposed work, a few highly correlated features are extracted from the PPG signals, and computationally efficient, simple linear regression-based machine learning models are employed to map these features to SBP, DBP, and SpO2 values. The results successfully demonstrated that the proposed framework can accurately estimate vital sign values compared to true measurements.

This work introduces signal pre-processing, feature extraction, and feature filtering steps to ensure that meaningful features are provided in the model. A total of 7 features were initially extracted, and then feature selection algorithms were applied to reduce the computational complexity for BP estimation. For SpO2, a single R-value is used as a feature for model training. Three models are trained for BP and SpO2 estimation, with linear and polynomial regression delivering the best performance, respectively. A post-processing stage is introduced in the proposed framework to improve the accuracy of the vital signs in realtime. When subjects engage in intense physical exercise or when the frequency of motion artifact components is closely aligned with that of the true vital sign component, the post-processing stage aids in tracking the vital sign within the current window. This improvement serves to enhance the overall performance of the proposed framework. The MAE value obtained for all the vital signs shows a better performance of the proposed framework. It surpasses the performance of those reported in existing literature, establishing them as state-of-the-art models. The proposed framework used a computationally efficient GSMD algorithm and signal processing block, which holds significant potential for developing wearable, non-invasive, realtime devices for vital sign estimation and offers a valuable contribution to healthcare technology. Thus, in the future, the proposed framework will be deployed in the computationally efficient STM 32 processing device for realtime estimation of vital signs.

Funding No funding is available.

Declarations

Conflict of Interest All Authors of this work declare no conflict of interest.

Ethical Approval This article contains no studies with human participants or animals performed by authors.

References

1. Nwibor C et al (2023) Remote health monitoring system for the estimation of blood pressure, heart rate, and blood oxygen saturation level. IEEE Sens J 23(5):5401–5411. https://doi.org/10.1109/JSEN.2023.3235977
2. Majumder S, Mondal T, Deen MJ (2017) Wearable sensors for remote health monitoring. Sensors (Switzerland). https://doi.org/10.3390/s17010130
3. Shirmohammadi S, Barbe K, Grimaldi D, Rapuano S, Grassini S (2016) Instrumentation and measurement in medical, biomedical, and healthcare systems. IEEE Instrum Meas Mag 19(5):6–12. https://doi.org/10.1109/MIM.2016.7579063

4. Pankaj, Kumar A, Komaragiri R, Kumar M (2022) A review on computation methods used in photoplethysmography signal analysis for heart rate estimation. Springer.https://doi.org/10.1007/s11831-021-09597-4
5. Pankaj, Kumar A, Ashdhir A, Komaragiri R, Kumar M (2023) Analysis of photoplethysmogram signal to estimate heart rate during physical activity using fractional Fourier transform—a sampling frequency independent and reference signal-less method. Comput Methods Programs Biomed. https://doi.org/10.1016/j.cmpb.2022.107294
6. Pankaj, Kumar A, Komaragiri R, Kumar M (2023) A novel CS-NET architecture based on the unification of CNN, SVM and super-resolution spectrogram to monitor and classify blood pressure using photoplethysmography. Comput Methods Programs Biomed. https://doi.org/10.1016/j.cmpb.2023.107716
7. Pankaj, Kumar A, Kumar M, Komaragiri R (2023) Optimized deep neural network models for blood pressure classification using Fourier analysis-based time-frequency spectrogram of photoplethysmography signal. Biomed Eng Lett. https://doi.org/10.1007/s13534-023-00296-6
8. Mourad N (2022) Group-sparse mode decomposition: a signal decomposition algorithm based on group-sparsity in the frequency domain. Dig Signal Process Rev J. https://doi.org/10.1016/j.dsp.2021.103375
9. Zhang Z, Pi Z, Liu B (2015) TROIKA: a general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise. IEEE Trans Biomed Eng 62(2):522–531. https://doi.org/10.1109/TBME.2014.2359372
10. Shuzan MNI et al (2024) PPG2RespNet: a deep learning model for respirational signal synthesis and monitoring from photoplethysmography (PPG) signal. Phys Eng Sci Med. https://doi.org/10.1007/s13246-024-01482-1
11. Karlen W, Raman S, Ansermino JM, Dumont GA (2013) Multiparameter respiratory rate estimation from the photoplethysmogram. IEEE Trans Biomed Eng 60(7):1946–1953. https://doi.org/10.1109/TBME.2013.2246160
12. Kumar A, Komaragiri R, Kumar M (2023) Blood pressure estimation and classification using a reference signal-less photoplethysmography signal: a deep learning framework. Phys Eng Sci Med 46(4):1589–1605
13. Pankaj, Kumar A, Kumar M, Komaragiri R (2022) STSR: spectro-temporal super-resolution analysis of a reference signal less photoplethysmogram for heart rate estimation during physical activity. IEEE Trans Instrum Meas. https://doi.org/10.1109/TIM.2022.3192831
14. Maan P, Kumar M, Kumar A, Komaragiri R (2024) A cascaded approach of group sparse mode decomposition and deep neural network for heart rate estimation using reference signal-less PPG signal. Measurement 246:116546
15. Shirzadfar H (2018) Heart beat rate monitoring using optical sensors. Int J Biosens Bioelectron 4(2):45–51. https://doi.org/10.15406/ijbsbe.2018.04.00097
16. Pankaj, Kumar A, Komaragiri R, Kumar M (2022) Reference signal less Fourier analysis based motion artifact removal algorithm for wearable photoplethysmography devices to estimate heart rate during physical exercises. Comput Biol Med. https://doi.org/10.1016/j.compbiomed.2021.105081
17. Kumar A, Komaragiri R, Kumar M (2023) Estimation of respiratory rate from a corrupted PPG signal using time-frequency spectrogram. In: Proceedings of the 10th international conference on signal processing and integrated networks, SPIN 2023. Institute of Electrical and Electronics Engineers Inc., pp 259–264. https://doi.org/10.1109/SPIN57001.2023.10117416
18. Sharma H (2019) Extraction of respiration from PPG signals using Hilbert vibration decomposition. In: Proceedings of the 2019 9th International Conference on Bioscience, Biochemistry and Bioinformatics, pp 48–52. https://doi.org/10.1145/3314367.3314369
19. Shuzan MNI, et al (2021) A novel non-invasive estimation of respiration rate from photoplethysmograph signal using machine learning model [Online]. http://arxiv.org/abs/2102.09483
20. Rastegar S, Gholam Hosseini H, Lowe A (2023) Hybrid CNN-SVR blood pressure estimation model Using ECG and PPG signals. Sensors. https://doi.org/10.3390/s23031259
21. Chowdhury A, Das D, Eldaly ABM, Cheung RCC, Chowdhury MH (2024) Photoplethysmogram-based heart rate and blood pressure estimation with hypertension classification. IPEM-Translation 9:100024. https://doi.org/10.1016/j.ipemt.2024.100024
22. Cen Y et al (2024) OVAR-BPnet: a general pulse wave deep learning approach for cuffless blood pressure measurement. IEEE J Biomed Health Inform. https://doi.org/10.1109/JBHI.2024.3423461
23. Zhong Y, Jatav A, Afrin K, Shivaram T, Bukkapatnam STS. Explainable machine learning-based SpO 2 estimation using photoplethysmograms measured from a neck wearable [Online]. https://ssrn.com/abstract=4333613
24. Shuzan MNI et al (2023) Machine learning-based respiration rate and blood oxygen saturation estimation using photoplethysmogram signals. Bioengineering. https://doi.org/10.3390/bioengineering10020167
25. Kong Y, Chon KH (2019) Heart rate tracking using a wearable photoplethysmographic sensor during treadmill exercise. IEEE Access 7:152421–152428. https://doi.org/10.1109/ACCESS.2019.2948107
26. Shyam A, Ravichandran V, Preejith SP, Joseph J, Sivaprakasam M (2019) PPGnet: deep network for device independent heart rate estimation from photoplethysmogram. Comput Sci. 1899–1902. arXiv:1903.08912v1 [cs.LG]
27. Rocha LG et al (2020) Binary CorNET: accelerator for HR estimation from Wrist-PPG. IEEE Trans Biomed Circuits Syst 14(4):715–726. https://doi.org/10.1109/TBCAS.2020.3001675
28. Wilkosz M, Szczęsna A (2021) Multi-headed conv-lstm network for heart rate estimation during daily living activities. Sensors. https://doi.org/10.3390/s21155212
29. Chang X, Li G, Xing G, Zhu K, Tu L (2021) DeepHeart: a deep learning approach for accurate heart rate estimation from PPG signals. ACM Trans Sens Netw 17(2):1–18. https://doi.org/10.1145/3441626
30. Ray D, Collins T, Ponnapalli PV (2022) DeepPulse: an uncertainty-aware deep neural network for heart rate estimations from wrist-worn photoplethysmography. https://doi.org/10.36227/techrxiv.19368140.v2
31. Ismail S, Siddiqi I, Akram U (2022) Heart rate estimation in PPG signals using Convolutional-Recurrent Regressor. Comput Biol Med 145:105470. https://doi.org/10.1016/j.compbiomed.2022.105470
32. Thakur S, Chao PCP, Tsai CH (2023) Precision heart rate estimation using a PPG sensor patch equipped with new algorithms of pre-quality checking and hankel decomposition. Sensors. https://doi.org/10.3390/s23136180
33. Hasan K, Chowdhury MH, Pathan NS, Hossain QD (2023) Heart rate estimation from wrist PPG signal during intense physical exercise. SN Comput Sci 4(5):684
34. Tan Y, Xie L, Yang S, Zhang S, Xia Z, Ma P (2024) Real-time heart rate estimation algorithm based on adaptive spectrum correction and peak localization. IEEE Sens J. https://doi.org/10.1109/JSEN.2024.3444038
35. Kachuee M, Kiani MM, Mohammadzade H, Shabany M (2017) Cuffless blood pressure estimation algorithms for continuous health-care monitoring. IEEE Trans Biomed Eng 64(4):859–869. https://doi.org/10.1109/TBME.2016.2580904

36. Khalid SG, Zhang J, Chen F, Zheng D (2018) Blood pressure estimation using photoplethysmography only: comparison between different machine learning approaches. J Healthc Eng. https://doi.org/10.1155/2018/1548647
37. Ertuğrul ÖF, Sezgin N (2018) A noninvasive time-frequency-based approach to estimate cuffless arterial blood pressure. Turk J Electr Eng Comput Sci 26(5):2260–2274. https://doi.org/10.3906/elk-1712-215
38. Slapni G, Ar Č, Mlakar N, Luštrek M (2019) Blood pressure estimation from photoplethysmogram using a spectro-temporal deep neural network. Sensors (Switzerland). https://doi.org/10.3390/s19153420
39. Chen S, Ji Z, Wu H, Xu Y (2019) A non-invasive continuous blood pressure estimation approach based on machine learning. Sensors (Switzerland). https://doi.org/10.3390/s19112585
40. Harfiya LN, Chang CC, Li YH (2021) Continuous blood pressure estimation using exclusively photopletysmography by lstm-based signal-to-signal translation. Sensors. https://doi.org/10.3390/s21092952
41. Mahmud S et al (2022) A shallow U-Net architecture for reliably predicting blood pressure (BP) from photoplethysmogram (PPG) and electrocardiogram (ECG) signals. Sensors 22(3):919. https://doi.org/10.3390/s22030919
42. Qiu S, Zhang YT, Lau SK, Zhao N (2022) Scenario adaptive cuffless blood pressure estimation by integrating cardiovascular coupling effects. IEEE J Biomed Health Inform. https://doi.org/10.1109/JBHI.2022.3227235
43. Nour M, Polat K, Şentürk Ü, Arıcan M (2023) A novel cuffless blood pressure prediction: uncovering new features and new hybrid ML models. Diagnostics. https://doi.org/10.3390/diagnostics13071278
44. Qin C, Li Y, Liu C, Ma X (2023) Cuff-less blood pressure prediction based on photoplethysmography and modified ResNet. Bioengineering 10(4):400. https://doi.org/10.3390/bioengineering10040400
45. Zhang X, Ding Q (2017) Biomedical Signal Processing and Control Respiratory rate estimation from the photoplethysmogram via joint sparse signal reconstruction and spectra fusion. Biomed Signal Process Control 35:1–7. https://doi.org/10.1016/j.bspc.2017.02.003
46. Zhu T, Johnson AE, Yang Y, Clifford GD, Clifton DA (2018) Bayesian fusion of physiological measurements using a signal quality extension Bayesian fusion of physiological measurements using a signal quality extension. Physiol Meas 39(6):065008. https://doi.org/10.1088/1361-6579/aac856
47. Motin MA, Karmakar CK, Palaniswami M (2019) Selection of empirical mode decomposition techniques for extracting breathing rate from PPG. IEEE Signal Process Lett 26(4):592–596. https://doi.org/10.1109/LSP.2019.2900923
48. Zhu T, Clifton DA (2021) Comparison of parametric and non-parametric Bayesian inference for fusing sensory estimates in physiological time-series analysis. Healthc Technol Lett 8(2):25–30. https://doi.org/10.1049/htl2.12003
49. Stankoski S, Kiprijanovska I, Mavridou I, Nduka C, Gjoreski H, Gjoreski M (2022) Breathing Rate Estimation from Head-Worn Photoplethysmography Sensor Data Using Machine Learning. Sensors. https://doi.org/10.3390/s22062079
50. Zhao Q et al (2023) Predicting respiratory rate from electrocardiogram and photoplethysmogram using a transformer-based model. Bioengineering 10(9):1024. https://doi.org/10.3390/bioengineering10091024
51. Yang G et al (2024) Energy-efficient ppg-based respiratory rate estimation using spiking neural networks. Sensors. https://doi.org/10.3390/s24123980
52. Chin WJ et al (2024) A novel respiratory rate estimation algorithm from photoplethysmogram using deep learning model. Diagnostics. https://doi.org/10.3390/diagnostics14030284
53. Venkat S, et al (2019) Machine learning based spo 2 computation using reflectance pulse oximetry. In: 2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC). IEEE, pp 482–485
54. Vijayarangan S, Suresh P, Preejith SP, Joseph J, Sivaprakasam M (2020) Robust modelling of reflectance pulse oximetry for spo 2 estimation. In: 2020 42nd annual international conference of the IEEE engineering in medicine & biology society (EMBC). IEEE, pp 374–377
55. Berwal D, Kuruba A, Shaikh AM, Udupa A, Baghini MS (2022) SpO 2 measurement: non-idealities and ways to improve estimation accuracy in wearable pulse oximeters. IEEE Sens J 22(12):11653–11664
56. Koteska B, Bogdanova A, Vićentić T, Ilić S, Tomić M, Spasenović M (2024) Prediction of oxygen saturation from graphene respiratory signals with PPG trained DNN. INSTICC. https://doi.org/10.5220/0012354100003657
57. Zhang L, Yu X, Xie H, Lin J, Wang Z (2024) Removal of motion artifacts in PPG signals based on the CEEMDAN-MPE and VS_LMS adaptive filter. In: IEEE Transactions on Instrumentation and Measurement, vol. 73, pp. 1–10, , Art no. 4010610. https://doi.org/10.1109/TIM.2024.3453341

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Pankaj[1] · Pratibha Maan[2] · Manjeet Kumar[3] · Ashish Kumar[4] · Rama Komaragiri[5]

✉ Manjeet Kumar
manjeetchhillar@dtu.ac.in

Pankaj
pankajbatra.ece@piet.co.in

Pratibha Maan
maanpratibha@gmail.com

Ashish Kumar
ashish.kumar4@bennett.edu.in

Rama Komaragiri
rama.komaragiri@galgotiasuniversity.edu.in

[1] Department of Electronics and Communication Engineering, Panipat Institute of Technology, Panipat, India

[2] CSIR-National Institute of Science Communication and Policy Research (CSIR-NIScPR), Council of Scientific & Industrial Research (CSIR), New Delhi, India

[3] Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India

[4] School of Computer Science Engineering and Technology, Bennett University, Greater Noida, India

[5] Department of Electronics and Communication Engineering, Galgotias University, Greater Noida, India

Latch and Flip-Flop Design Using Memristors and Transmission Gates with Timer Counter Implementation

Aditya Rayer
Dept. of ECE
Delhi Technological University
Delhi, India
adityarayer_23vls24@dtu.ac.in

Mavulluru Ganesh
Dept. of ECE
Delhi Technological University
Delhi, India
mganesh@dtu.ac.in

Lavi Tanwar
Dept. of ECE
Delhi Technological University
Delhi, India
lavi.tanwar@dtu.ac.in

2025 10th International Conference on Signal Processing and Communication (ICSC) | 979-8-3315-1189-0/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICSC64553.2025.10967986

***Abstract*—Memristors are a combination of the words Memory and Resistor, it never forgets. It is a passive element which remembers the amount of current that flowed through it and is small in size. All these qualities of a memristor make it highly advantageous and useful in the upcoming times for chip manufacturing and designing. Latches are sequential circuits using them in master-slave configuration a flip flop can be generated which is highly useful in applications such as counters, registers, etc. In this paper we implement a Modified version of D-latch and flip flop using memristors and Transmission gates to improve the switching speed and to obtain full voltage swing. Further we show a comparative analysis of proposed modified D-latch and flip-flop with an earlier method along with a timer counter as an application of the D-flip flop circuit. Paper also explains the challenges of using MRL gates in complex circuits. TEAM model of memristor and 180nm technology node files are used in simulation of all the circuits using Cadence Virtuoso.**

***Index Terms*—TEAM, Pass Transistor, Transmission Gate, MRL, Latch, Flip-Flop, Timer Counter.**

I. Introduction

Until 1971, the resistor, capacitor, and inductor—the three fundamental circuit elements—dominated circuit theory. The resistor establishes a relationship between voltage and current, the capacitor connects voltage to charge, and the inductor links current to magnetic flux. The link between the Flux and Charge was missing until Prof. Leon Chua [1] postulated the missing link on a basis of logical symmetry. He named this as a Memristor which is a combination of two words: Resistor and Memory.

Memristor relates charge and magnetic flux by:

$$d\phi = M(q)dq \tag{1}$$

The M value is known as the memristance of the memristor or the ability to charge the resistance based on the amount of charge that has passed through the memristor in the past. Memristors were only theoretical devices until the year 2008 when Stanley Williams implemented the first practical memristor at HP labs. It was a Pt/TiO2 based bipolar device. It was termed ReRAM (resistive switching random access memory) and was called a subset of Memristor. It is a two-terminal nonlinear resistor, a passive element which it does not stores any energy, it remembers the last amount of current passed through it or the last value of voltage applied to it.

Memristor-based logic circuits provide us with lots of benefits such as nonvolatility, faster switching rates, etc. There are many memristor-based logic designs, some of them are IMPLY, Memristor Aided Logic (MAGIC), Memristor Ratioed Logic (MRL), Modified Gate Diffusion Input (M-GDI), etc. This paper mainly focuses on the MRL Logic Circuits and briefly discusses the reasons why it is not used for implementation of D-latch and D-flip flop. Flip-flops and latches are fundamental sequential logic elements used for data storage. A D-latch (data latch) allows data to pass through only when the enable signal is active, storing it otherwise. The D flip-flop, triggered by clock edges, ensures synchronized data storage and transfer, making it vital for timing-critical applications.

II. Literature Review

There are various memristor models that are available for use. One of the most used and accurate is ThrEshold Adaptive Memristor Model (TEAM) of memristor presented in [2]. Earlier models such as linear ion drift model and non linear ion drift models had some limitations under threshold dependent applications. TEAM model takes these models forward by taking the threshold voltage applications in account. These threshold dynamics are highly useful in precise circuits and critical applications where control over the activation and transition of states is required. The paper [2] reviews the situations where the earlier models failed or did not provide acceptable results and compares TEAM model with others with an aim to demonstrate the improvements, accuracy and efficiency.

Many researchers have explored the use of memristor models in creation of logic gates and using those logic circuits for creation of latches and flip-flops, one such method of creation of D-flip flop is shown in [17], the paper presents the limitaions of using the CMOS D flip flops such as power consumption, and overall circuit size which is big due to presence of large number of transistors. Since the memristors use low power and have higher switching speed as compared to traditional

979-8-3315-1189-0/25/$31.00 ©2025 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:33:48 UTC from IEEE Xplore. Restrictions apply.

CMOS gates. The paper used another memristor model named Knowm memristor model and TSMC 180nm technology node files to implement the NOR gate and using that gate a D latch using cross coupled orientation was created and further using master slave configuration D flip flop is created. Similar methods of creation of logic gates and other circuits using those circuits have been presented in papers [3] - [12] . Other significant memristor model has been presented in [13].

A. Contributions and organization

In this paper, we implement a modified D-latch circuit, explain its functionality, and compare its performance with a previous D flip-flop design [3] using a single NMOS pass transistor. The proposed design replaces the pass transistor with a transmission gate, improving speed, power consumption, and robustness.

In Section III, we discuss the MRL-based design of basic logic circuits. In Section IV, we introduce the modified D-latch circuit. Section V presents circuit outputs, observations, and comparisons. In Section VI, we highlight a Timer Counter application using cascaded flip-flops based on the modified design. Finally, concluding remarks are provided in Section VII.

III. Circuit Designing using Memristors

Logic Gates are the most fundamental components of all the digital systems and circuits. They are used to execute many logical operations that might be required by a circuit, logic gate output is dependent on the input combinations and the logical operation that it is designed to perform. The logic gates are made by a combination of transistors, resistors, capacitors etc. Logic gates designed using transistors are CMOS logic gates. This paper presents a study of memristor based logic gates mainly MRL (Memristor Ratioed Logic) and MAGIC.

MAGIC orientation focuses on using only the memristors to create the logic gates, whereas the MRL used both memristors and CMOS Inverters and hence it is also known as the Hybrid CMOS Memristor Logic Design.

A. Design and Functioning of Basic Gates

In order to create a flip flop or a latch we need the basic logic gates mainly AND and OR gates which can be later converted to NAND and NOR using the MRL logic according to which the output of two gates is passed through a CMOS inverter to achieve the desired outputs of NAND and NOR.

In AND gate "Fig. 1", two memristors are connected in anti-serial manner and its functioning can be easily understood as follows:

Case 1: When both the the inputs A and B are equal to zero. The memristor resistance from N-P is "high" and P-N resistance is "low", so no current flows through the memristors and the output received is zero.

Case2: When both the inputs A and B are equal to one. The memristor resistance from N-P is "low" and P-N resistance is "high", so the current through both the memristors flows towards the output and a logic 1 output is received.

Case 3: When the inputs are A=0 (gnd) and B=1(Vdd), or A=1(Vdd) and B=0(gnd), the memristor with input 0 has high resistance from N-P and low resistance from P-N, and the memristor with input 1 has low resistance from N-P and high resistance from P-N. Due to this the current flows from the higher value towards the lower value and no current flows towards the output resulting a logic 0 as output.

In a similar way the functioning of OR gate can be explained and the truth tables of both the gates can be verified.

Simply adding a CMOS inverter at the output of the AND and OR gate circuits gives us the required NAND and NOR gates this method of using CMOS inverter at output is MRL logic.

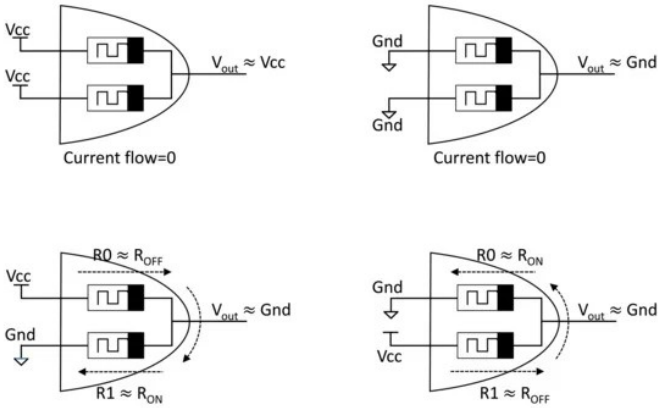


Fig. 1. MRL AND Gate Logical Operation [4]

A comparative analysis of the memristor based and conventional CMOS NAND and NOR gates on basis of power consumption and number of transistors is shown in the Table I. The power consumption is calculated using the simulation software which uses the average power consumption formula:

$$P_{av} = \frac{V_{DD}}{T} \int_0^T i_{DD}(t)\, dt \tag{2}$$

where P_{av} is the average power,
V_{DD} is the supply voltage,
$i_{DD}(t)$ is the supply current as a function of time,
T is the time period.

TABLE I
Comparison of CMOS and Memristor Based Gates

Component	Power Consumption	Transistor Count
CMOS NAND Gate	2.267 μW	4
Memristor Based NAND Gate	1.514 μW	2
CMOS NOR Gate	1.5181 μW	4
Memristor Based NOR Gate	0.846 μW	2

B. ThrEshold Adaptive Memristor (TEAM) Model

The TEAM model is a general memristor model [2]. The TEAM model is simple to use and very flexible. Its simplicity improves the efficiency of the simulation and thus the model

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:33:48 UTC from IEEE Xplore. Restrictions apply.

has acceptable accuracy. The model is suitable for memristive circuit designs and has been implemented in VerilogA for use in simulations [2].

In this project TEAM model of memristor is used with two additional parameters added to the model VSET and VRESET to tackle the 01 and 10 states of logic gates.

- VSET: Sets the memristor to value 1 when the output of memristor is above a certain threshold value.
- VRESET: Resets the memristor to value 0 when the output of memristor is below a certain threshold value.

C. Flip Flop and Latch using the Basic Gates

The creation of basic logic gates makes the creation of flip flops and latches easier theoretically but using the NAND and NOR gates created using MRL creates a problem. As explained in the previous section at states 01 and 10 the current flows from higher potential node to lower potential node, this creates a problem for circuits with feedback from output to input, example JK flip flop.

In such a circuit when the input logic gate is in a state where both inputs are different there will be a reverse flow of current due to which the overall output of the circuit may get disturbed, this causes a lot of reliability issues in using the basic logic gates created using memristors. Cascading several of these gates may lead to a floating output (between logic ”0” and logic ”1”) due to voltage degradation. Since memristors are passive devices, they cannot amplify signals.

This can be solved at the cost of overall power consumption and transistor count by adding buffers to the circuit. Other methods to make D-latch and D-flip flop is shown in [3]. The method presented in [3] includes the use of memristors, buffers and NMOS pass transistor to make the D-latch and flip flop. This paper takes that circuit and presents a comparative analysis of this circuit with the one which replaces the single pass transistor with the transmission gate.

D. Real Time Applications of Memristor Based Gates

Memristor although postulated in 1971 and later implemented by HP Labs in 2008 is still not used in real time applications due to issues and claims about its reliability and practical functionality. Multiple memristor models are present to theoretically explain the working of the memristors but using these models device modeling is still a big challenge. Using these models and different memristor orientations various applications are implemented using software based simulations, one such application is Half and Full Adder.

IV. Proposed Methodology

This section will show a Modified D-Latch and D-Flip Flop Circuit. In the paper [3] the authors have proposed a design of a positive level D-Latch, which consists of a single nmos pass transistor, memristors and a few inverters ”Fig. 2”. The pass transistor acts as a switch, when the clock signal is given at the gate of the transistor, controlling the switching function.

When the clock signal is high the pass transistor acts as a closed switch which allows the D signal to pass through it,

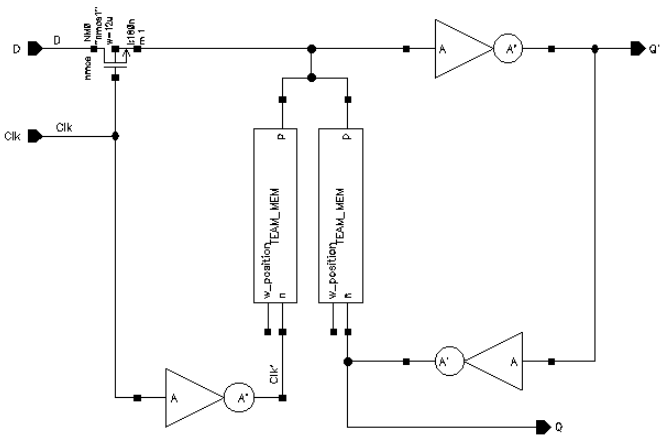


Fig. 2. D-Latch Circuit using NMOS Pass Transistor [3]

and the pass transistor acts as a open switch or it is in high impedance state when clock signal low is passed through it.

When the clock signal is high, the input D flows through the transistor to the output. When the clock signal becomes low, the current value of D is stored at the output, thus preventing changes until the clock signal goes high again. The latching mechanism is implemented using two memristors: one receives the inverted clock signal, and the other operates in a feedback loop. When the clock is high, feedback is disabled, but when it is low, the feedback activates, enabling the input to be latched securely.

This circuit does not get stuck even though it has a feedback loop as when the feedback is active both the memristors which are essentially in the AND gate configuration have same values at input because of presence of two inverters or a buffer in the feedback loop.

A. Modified Circuit Using Transmission Gate

A transmission gate is a simple yet effective switch in digital circuits, consisting of a PMOS and an NMOS transistor connected in parallel. It uses a control signal to determine its state:

When the control is active, the gate conducts, allowing signals to flow freely in either direction. When the control signal is inactive, it blocks the signal. This design minimizes voltage drops and ensures efficient signal transfer, making it ideal for applications like multiplexers, latches, and flip-flops where precise and low-resistance switching is crucial.

Disadvantage of using a single NMOS pass transistor is that it passes a ’Strong 0’ and ’Weak 1’. Whereas in transmission gates ”Fig. 3” since both nmos and pmos are present in parallel so the issue of passing a weak signals is solved.

When the input signal A is low, both nmos and pmos are cut off and circuit acts as an open switch. When A is high the transistors are biased to conduct and circuit acts as a closed switch, so we can say transmission gates are voltage controlled switches which passes either a strong 0 or a strong 1 based on the control signal value. Because of this the circuit

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:33:48 UTC from IEEE Xplore. Restrictions apply.

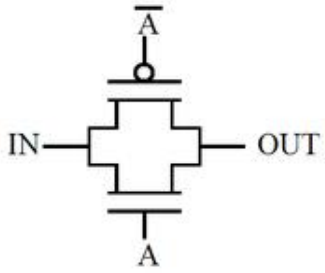


Fig. 3. Transmission Gate

is able to achieve full voltage swing, the node capacitance is reduced, lower switching energy required to charge up the node, switching rate is also improved. The modified D-latch circuit can be seen in "Fig. 4",

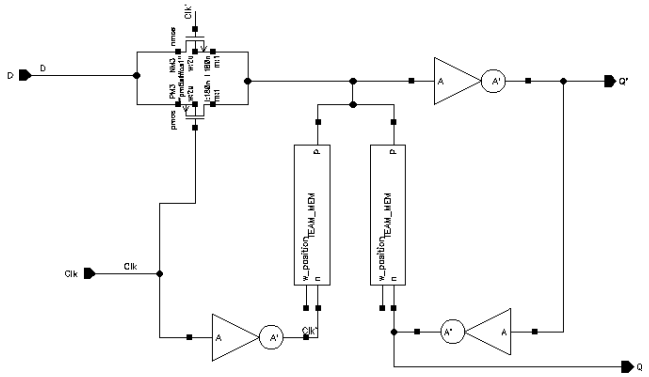


Fig. 4. Modified D-latch using Transmission Gate

a D-flip flop created using the master slave configuration connecting two D-latches is also shown in "Fig. 5" The outputs of both the circuits is shown in "Fig. 6" and "Fig. 7" respectively.

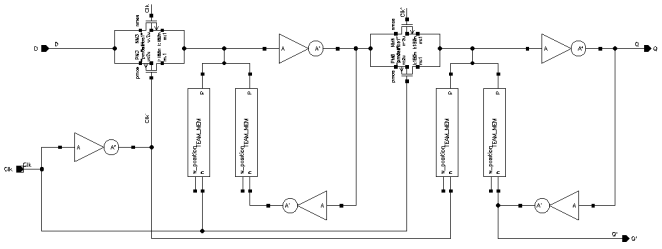


Fig. 5. Modified D-Flip Flop using Transmission Gate

V. Results and Analysis

A comparison on the basis of number of transistors, memristors, power consumption, and circuit delay is presented in Table II. The outputs of circuits "Fig. 4", "Fig. 5" are shown in "Fig. 6" and "Fig. 7" respectively. Both the outputs of latch and flip flop can be verified, it can be observed that the outputs have good switching and full voltage swing.

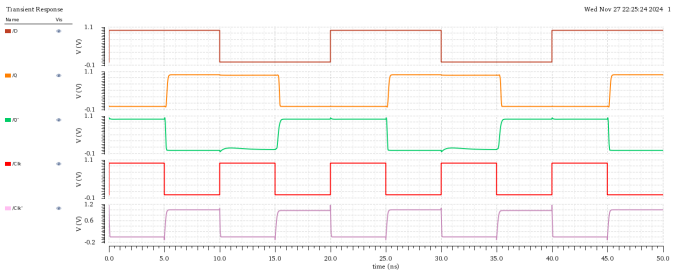


Fig. 6. Modified D-latch Output

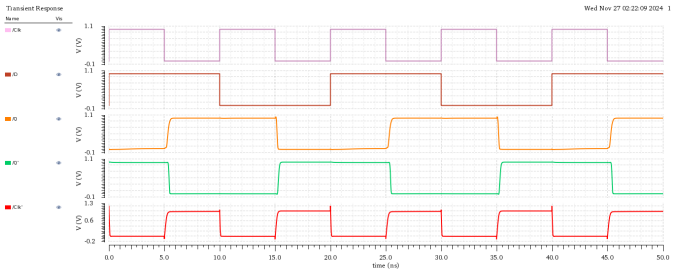


Fig. 7. Modified D-Flip Flop Output

TABLE II
Comparison of Implemented Circuits

Comparison Parameters	**Circuits**			
	D-Latch using Pass Transistor	***D-Latch Using Trans-mission Gates***	***D-Flip Flop using Pass Transistor***	***D-Flip Flop using Trans-mission Gates***
No. of Transistors	7	8	14	16
No. of Memristors	2	2	4	4
Avg. Power (Watt)	5.04μ	5.42μ	7.54μ	8.05μ
D to Q Delay (Seconds)	10.185μ	5.324μ	5.287μ	5.16μ

Delay and Power consumption was calculated using the virtuoso software calculator tool. The comparison table shows the slight increase in power consumption of the circuit which is mainly due to increased transistor, and significantly low delay in D-Latch due to use of transmission gates.

VI. Application

A timer counter "Fig. 8" using the modified D-flip flop circuit presented in "Fig. 5" is made using the symbol of flip flop, the output of the circuit "Fig. 9" is ideal and verifies that the latch and flip flop circuit works correctly.

The output shows slight degradation of the outputs which is due to cascading of 4 memristor based D-flip flops and as discussed earlier memristors are non- amplifying circuits cascading them leads to output degradation.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:33:48 UTC from IEEE Xplore. Restrictions apply.

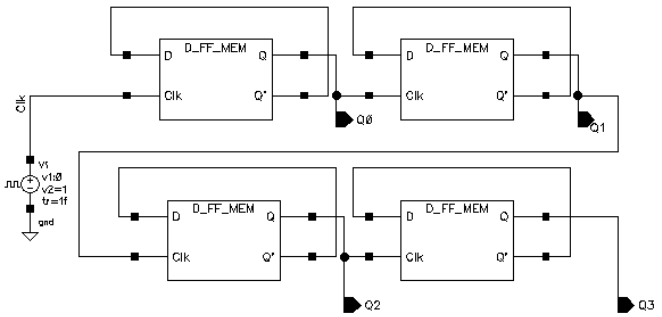


Fig. 8. Timer Counter using Modified D-flip flop symbol

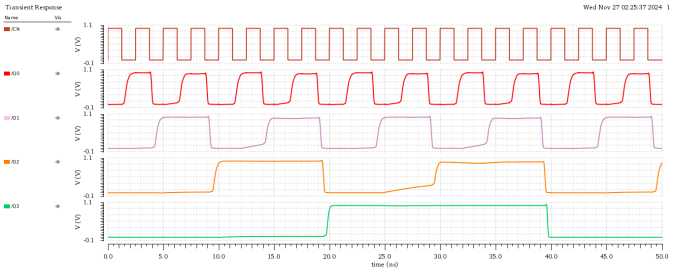


Fig. 9. Timer Counter Output

VII. Conclusion

Based on the comparison it is observed that the circuit using the transmission gate has slightly higher power consumption due to increased transistor count. The delay is lesser due to better switching capability of transmission gate. A trade-off between power and performance is observed. The power usage is slightly increased while other important parameters take benefit of the transmission gates. Using the Modified D-Flip Flop a timer counter as an application is also shown which verifies the functioning of the modified circuit with good results.

References

[1] L. Chua, "Memristor-The missing circuit element," in IEEE Transactions on Circuit Theory, vol. 18, no. 5, pp. 507-519, September 1971, doi: 10.1109/TCT.1971.1083337.

[2] S. Kvatinsky, E. G. Friedman, A. Kolodny and U. C. Weiser, "TEAM: ThrEshold Adaptive Memristor Model," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 1, pp. 211-221, Jan. 2013, doi: 10.1109/TCSI.2012.2215714.

[3] P. Dhongade, K. Singh and Shalini, "Design and Analysis of Memristor-CMOS Based Hybrid D Latch," *2022 2nd International Conference on Intelligent Technologies (CONIT)*, Hubli, India, pp. 1-4, 2022, doi: 10.1109/CONIT55038.2022.9847980.

[4] Ali, K.A.; Rizk, M.; Baghdadi, A.; Diguet, J.-P.; Jomaah, J. Hybrid Memristor–CMOS Implementation of Combinational Logic Based on X-MRL. *Electronics*, vol. 10, pp. 1018, 2021. https://doi.org/10.3390/electronics10091018.

[5] Hussain, S.A., Prasad V, P.N.S.B.S.V. & Sanki, P.K. Synergistic m_GDI-Based ALU Design Using CMOS and VTEAM Memristor Model for Low-Power High-Speed Applications. *J. Electron. Mater.* vol. 53, pp. 3626–3641, 2024. https://doi.org/10.1007/s11664-024-11125-6.

[6] J. Zheng, Z. Zeng and Y. Zhu, "Memristor-based nonvolatile synchronous flip-flop circuits," *2017 Seventh International Conference on Information Science and Technology (ICIST)*, Da Nang, Vietnam, pp. 504-508, 2017, doi: 10.1109/ICIST.2017.7926812.

[7] Z. Chang, A. Cui, Z. Wang and G. Qu, "Novel Memristor-based Nonvolatile D Latch and Flip-flop Designs," *2021 22nd International Symposium on Quality Electronic Design (ISQED)*, Santa Clara, CA, USA, pp. 244-250, 2021, doi: 10.1109/ISQED51717.2021.9424269.

[8] P. W. C. Ho, H. A. F. Almurib and T. N. Kumar, "Non-volatile D-latch for sequential logic circuits using memristors," *TENCON 2015 - 2015 IEEE Region 10 Conference*, Macao, China, pp. 1-4, 2015, doi: 10.1109/TENCON.2015.7372849.

[9] Rajesh, R. Arundeepakvel, Jatin, P. Khatter and N. Pandey, "Realization of Memristor based D-Latch," *2019 International Conference on Computing, Power and Communication Technologies (GUCON)*, New Delhi, India, pp. 764-768, 2019.

[10] N. Kaushik and S. Bodapati, "IMPLY-Based High-Speed Conditional Carry and Carry Select Adders for In-Memory Computing," in *IEEE Transactions on Nanotechnology*, vol. 22, pp. 280-290, 2023, doi: 10.1109/TNANO.2023.3284845.

[11] N. Kaushik and B. Srinivasu, "Memristor-based High Speed and Area Efficient Comparators in IMPLY Logic," *2023 36th International Conference on VLSI Design and 2023 22nd International Conference on Embedded Systems (VLSID)*, Hyderabad, India, pp. 139-144, 2023, doi: 10.1109/VLSID57277.2023.00040.

[12] S. Kvatinsky, K. Talisveyberg, D. Fliter, A. Kolodny, U. C. Weiser and E. G. Friedman, "Models of memristors for SPICE simulations," *2012 IEEE 27th Convention of Electrical and Electronics Engineers in Israel*, Eilat, Israel, pp. 1-5, 2012, doi: 10.1109/EEEI.2012.6377081.

[13] S. Kvatinsky, M. Ramadan, E. G. Friedman and A. Kolodny, "VTEAM: A General Model for Voltage-Controlled Memristors," in *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 62, no. 8, pp. 786-790, Aug. 2015, doi: 10.1109/TCSII.2015.2433536.

[14] L. Xie, H. A. D. Nguyen, M. Taouil, S. Hamdioui and K. Bertels, "A Mapping Methodology of Boolean Logic Circuits on Memristor Crossbar," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 2, pp. 311-323, Feb. 2018, doi: 10.1109/TCAD.2017.2695880.

[15] M. Gavilánez and S. Guitarra, "A Calibrated Memristor Model Implementation for an SR-Latch Based on CMOS-Memristor Technology," *2024 IEEE 15th Latin America Symposium on Circuits and Systems (LASCAS)*, Punta del Este, Uruguay, pp. 1-4, 2024, doi: 10.1109/LASCAS60203.2024.10506178.

[16] M. D. Marco, M. Forti, L. Pancioni, G. Innocenti, A. Tesi and F. Corinto, "Oscillatory Circuits With a Real Non-Volatile Stanford Memristor Model," in *IEEE Access*, vol. 10, pp. 13650-13662, 2022, doi: 10.1109/ACCESS.2022.3146419.

[17] S. Panda and C. S. Dash, "Hybrid CMOS Memristor Based Frequency Divider Using D Flip Flop," *2024 Parul International Conference on Engineering and Technology (PICET)*, Vadodara, India, pp. 1-5, 2024, doi: 10.1109/PICET60765.2024.10716182.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:33:48 UTC from IEEE Xplore. Restrictions apply.

Journal of Physics D:
Applied Physics

PURPOSE-LED
PUBLISHING™

ACCEPTED MANUSCRIPT

Linear and nonlinear photoluminescence in Ho^{3+}, Yb^{3+} doped $CaBi_2Nb_2O_9$ phosphor for photonic applications

To cite this article before publication: Sheetal Kumari *et al* 2025 *J. Phys. D: Appl. Phys.* in press https://doi.org/10.1088/1361-6463/adc9bc

Manuscript version: Accepted Manuscript

Accepted Manuscript is "the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an 'Accepted Manuscript' watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors"

This Accepted Manuscript is **© 2025 IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.**.

During the embargo period (the 12 month period from the publication of the Version of Record of this article), the Accepted Manuscript is fully protected by copyright and cannot be reused or reposted elsewhere.

As the Version of Record of this article is going to be / has been published on a subscription basis, this Accepted Manuscript will be available for reuse under a CC BY-NC-ND 4.0 licence after the 12 month embargo period.

After the embargo period, everyone is permitted to use copy and redistribute this article for non-commercial purposes only, provided that they adhere to all the terms of the licence https://creativecommons.org/licences/by-nc-nd/4.0

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions may be required. All third party content is fully copyright protected, unless specifically stated otherwise in the figure caption in the Version of Record.

View the article online for updates and enhancements.

This content was downloaded from IP address 14.139.251.98 on 15/04/2025 at 05:52

Linear and Nonlinear Photoluminescence in Ho^{3+}, Yb^{3+} Doped $CaBi_2Nb_2O_9$ Phosphor for Photonic Applications

Sheetal Kumari[1], Pooja Rohilla[1], Anu[1], Srinibas Barik[2], Bhanupriya Yadav[3], Kaushal Kumar[2], Runit Isaac[4], A.S.Rao[1], Aman Prasad[5]

[1]Department of Applied Physics, Delhi Technological University, Bawana Road, Delhi-110042, India

[2]Optical Materials and Bio-imaging Research Laboratory, Department of Physics, IIT (ISM) Dhanbad-826004, India

[3]Department of Chemistry, Dayalbagh Educational Institute (DEI), Deemed University, Dayalbagh, Agra-282005, India

[4]Department of Chemistry, Sam Higginbottom University of Agriculture, Technology and Sciences, Prayagraj-211007, India

[5]Department of Physics and Computer Science, Dayalbagh Educational Institute (DEI), Deemed University, Dayalbagh, Agra-282005, India

Abstract:

$CaBi_2Nb_2O_9$ (CBNO) phosphors doped with Ho^{3+} and Yb^{3+} ions were synthesized through a conventional solid-state reaction technique. The microstructure of the prepared samples was characterized using scanning electron microscopy (SEM). The energy bandgap was determined using diffuse reflectance spectroscopy (DRS). Down- and Up- converted photoluminescence analyses were conducted to examine the excitation and emission behaviours of the as prepared phosphors, along with the decay lifetimes measurement under 980 nm excitation. Up-conversion (UC) emission spectra revealed emission bands at 545 nm, and 658 nm, which correspond to the $^5F_4/^5S_2$ to $^5I_8$, and $^5F_5$ to $^5I_8$, transitions of Ho^{3+} ions, respectively. Analysis of UC emission intensity variation with pump power of 980 nm laser source confirmed the involvement of a two-photon absorption mechanism. The chromaticity coordinates derived from the emission spectra positioned the light emission in the green region. Studies of temperature-dependent photoluminescence (TD-PL) highlighted the phosphor's high thermal stability, supported by its notable thermal activation energy. This study shows encouraging results thereby demonstrating the potential of CBNO phosphors for applications in solid-state

Accepted Manuscript

lighting, display systems, security printing, phototherapy, and other up-conversion-based technologies.

Corresponding author: E-mail : amanprasad2891@gmail.com
Tel:+91-8800514652,

Keywords: Phosphor; Photoluminescence; Up-Conversion; PL Decay, TD-PL

Introduction:

Due to their special optical characteristics generated from 4f–4f electronic transitions, oxide materials doped with rare earth (RE) ions have garnered interest for their use in photonic devices [1–3]. These materials are vital to many industries, such as commercial products, everyday applications, and cutting-edge technologies. White light-emitting diodes (w-LEDs), lasers, display technology and solar cells are some of the key areas where RE doped materials find immense uses.. InGaN blue LEDs with yellow phosphor (YAG: Ce^{3+}) are a popular choice for producing white light because of their high efficiency, small size, light weight, improved durability, and long operating lives [3–6]. However, because of inadequate red light output, this method has problems including a low colour rendering index (CRI) and a high correlated colour temperature (CCT). Utilising trichromatic phosphors in conjunction with ultraviolet (UV) chips is an alternate method of enhancing CRI. Researchers are encouraged to further optimise phosphor characteristics due to the growing demand for sophisticated illuminating materials. Low light efficiency and unstable correlated colour temperature (CCT), two intrinsic problems of the commonly used commercial red phosphor (Y_2O_2S: Eu^{3+}) , are caused by colour reabsorption and degradation [7–9]. Oxide materials have become extremely flexible for a variety of applications due to their strong chemical stability and moderate synthesis capabilities. Single-phase white light-emitting phosphors, which are effectively activated by near-UV LED chips, have emerged as the go-to remedy for overcoming these difficulties [10,11]. Over the past few decades, scholars have given this technique a lot of attention. Due to its remarkable precision, high sensitivity, and quick response, optical thermometry has been increasingly popular recently.

Accepted Manuscript

Various host materials, such as phosphates, silicates, zincates, vanadates, niobates, and tungstates, can be doped with rare earth (RE) ions. Among these, niobates stand out as dependable host materials due to their outstanding chemical and thermal stability and their eco-friendly characteristics [12–14]. Niobium oxides, specifically, possess a low phonon frequency, which effectively reduces non-radiative transitions and enhances the fluorescence intensity of phosphors. Aurivillius-type bismuth-layered structure ferroelectrics (BLSFs) have attracted interest due to their general formula $(Bi_2O_2)^{2+}$ $(A_{m-1}B_mO_{3m+1})^{2-}$,, and their relatively high Curie temperature (T_c), which contributes to their superior thermal stability [12,15]. CBNO, a member of this family with m=2, incorporates Ca^{2+} at the A site and Nb^{5+} at the B site within its pseudo-perovskite $(CaNb_2O_7)^{2-}$ blocks. Its structure alternates between $(Bi_2O_2)^{2+}$ layers and $(CaNb_2O_7)^{2-}$ units, which contain double NbO_6 octahedral layers [16] . CBNO exhibits low phonon energy, which helps minimize non-radiative relaxation processes. This feature significantly boosts photoluminescence efficiency by reducing multi-phonon relaxation. CBNO also exhibits relatively thermal stability, due to its high Curie temperature and chemically robust layered structure. Unlike sulfide-based phosphors, which tend to deteriorate under high temperatures and humid conditions, CBNO maintains its luminescent performance even in demanding environments. CBNO provides a stable environment for controlled doping and enhanced energy transfer between sensitizer (Yb^{3+}) and activator (Ho^{3+}) ions. This improvement in up-conversion emission efficiency makes CBNO a promising candidate for applications in anti-counterfeiting, security printing, and biomedical imaging. Although the perovskite blocks can accommodate various cations, modifying the bismuth oxide layers remains challenging. To date, there has been no research on CBNO doped with Ho^{3+}/Yb^{3+} ions for fabricating white LEDs.

Doping is a fundamental technique in materials science, involving the incorporation of specific atoms into a host material to develop hybrid materials with customized properties. RE ions are highly regarded for their ability to emit a wide range of colors from the UV to IR regions. Their remarkable green emission, long-lived excited states, and numerous ladder-like energy levels give them this ability, making them ideal for use in lighting and display technologies [17,18]. Despite these advantages, up-conversion (UC) phosphors doped solely with Ho^{3+} have low excitation efficiency and weak emission under near-infrared (NIR) stimulation, primarily due to limited absorption. Yb^{3+} ions, with their strong absorption cross-section near 980 nm, are effective at transferring excitation energy to Ho^{3+} ions, making them excellent sensitizers [19–22]. By co-doping Ho^{3+} with Yb^{3+} the emission efficiency of Ho^{3+} can be significantly

Accepted Manuscript

improved. The distance between donor ions (Yb^{3+}) and acceptor ions (Ho^{3+}) often changes with different doping concentrations, affecting energy transfer efficiency [23]. Therefore, optimizing the Ho^{3+}/Yb^{3+} doping ratio is essential to enhance energy transfer and overall performance.

In this study, we synthesized a series of new CBNO: xHo^{3+}/yYb^{3+} phosphors using a solid-state reaction method at high temperatures. Systematic analysis is done to check the influence of test temperatures, Ho^{3+}/Yb^{3+} concentrations, and laser pump powers on UC luminescence. We have given a thorough demonstration of the transfer of energy mechanism of the Ho^{3+}/Yb^{3+} co-doped system. The suitability of the synthesised phosphor for uses in solid-state lighting and optical thermometry was finally evaluated.

2. Synthesis and instrumentations:

In this work, a high-temperature solid-state reaction method was utilized to synthesize a series of CBNO: xHo^{3+} (x = 1.0, 2.0, 3.0, 4.0, and 5.0 mol%) and CBNO: xHo^{3+}/yYb^{3+} (x = 3.0 mol%, y = 1.0, 2.0, 3.0, 4.0, and 5.0 mol%) doped phosphors. The raw materials used, $CaCO_3$ (Thermo Fisher 98%), Bi_2CO_3 (Loba Chemie 99%), Nb_2O_5 (Loba Chemie 99.9%), Ho_2O_3 (Central Drug House 99.9%), and Yb_2O_3 (Loba Chemie 99%), were weighed according to the stoichiometric ratios. The materials were transferred in an agate mortar, ground for 1 hour to achieve a uniform powder, and then transferred into an alumina crucible. The mixture was heated to 1250°C using a programmable furnace and allowed to cool to ambient temperature. After cooling, the samples were again ground to get the final powder form required for experimental characterisations. The complete synthesis process is shown in Fig. 1.

XRD data was recorded using a Bruker D8 setup using Cu (K_α=1.54 Å) wavelength. Morphological, elemental, and energy dispersive X-ray (EDX) analyses were conducted using a JEOL 7610F Plus microscope. A Jasco V-770 UV-Vis spectrophotometer was used to acquire the materials' diffuse reflectance spectra at room temperature. Jasco FP-8300 Spectrofluorometer, fitted with a Xenon lamp, was used to record the excitation and emission spectra. The up-conversion (UC) emission spectra were recorded using a 980 nm continuous-wave (CW) diode laser coupled with a fast CCD spectrometer (ULS2048x64, Avantes, USA). For temperature-dependent UC emission studies, the spectra were captured with a CCD spectrometer. The UC decay times were determined using the Horiba PTI Fluorescence Master System in conjunction with FelixGX software.

Accepted Manuscript

3. Result and discussion

3.1 XRD and FE-SEM:

The phase formation of the as-prepared phosphor was initially examined utilizing XRD analysis. The XRD patterns of the undoped CBNO phosphor align with the standard reference data for CBNO (card no. 49-0608) reported in our previous published work [24]. The XRD patterns of the undoped, singly Ho, and codoped Ho/Yb CBNO phosphors are shown in Fig.2. The absence of additional phases confirms that the samples are of single-phase. CBNO is characterized by an orthorhombic crystal structure belonging to the A21am space group, with lattice parameters a = 5.4827 Å, b = 5.4415 Å, c = 24.8974 Å, and $\alpha = \beta = \gamma = 90°$.

The morphology and particle size of the synthesized phosphors were examined using FE-SEM technique. Fig. 3(a, c) presents SEM images of undoped, single Ho, and Ho^{3+}/Yb^{3+} doped CBNO phosphors revealing their irregular shape. Agglomeration of particles is observed, likely due to the high-temperature synthesis and grinding processes. The particle size of the prepared samples falls within the range of 3-8 μm, aligning with the typical size of 2–10 μm for commercial phosphors, making them suitable for pc w-LED applications. Fig.3(d-f) represents the EDX images of undoped, single Ho, and Ho^{3+}/Yb^{3+} doped CBNO phosphors. The EDX images shows the presence of all the precursors used in the host lattice.

3.2 DRS and band gap calculations:

The DRS spectra of Ho^{3+} and Ho^{3+}/Yb^{3+} ion-doped CBNO phosphors, illustrated in Fig. 4, span the wavelength range of 200–1000 nm. Two distinct bands are seen in the visible region at 543 nm and 649 nm, corresponding to electronic transitions from the $^5I_8$ ground state of Ho^{3+} ions to the $F_4/^5S_2$ and $^5F_5$ excited states, respectively [25]. The absorption band was observed in NIR region owing to Yb^{3+} ions at wavelength 978 nm corresponding to transition $^2F_{5/2} \rightarrow {}^2F_{7/2}$. Given the strong absorption cross-section of Yb^{3+} in the NIR region, a 980 nm laser was employed to record the up-conversion photoluminescence (PL) spectra and associated lifetime curves. The optical band gap for both phosphor series was calculated using the Kubelka-Munk method [26,27].

$$F(R)h\nu = C(h\nu - E_g)^n \qquad (1)$$

Here Planck's constant (h), the incident light frequency (v), a proportionality constant (C), and the optical band gap (E_g) are key parameters in this analysis. The value of n, which determines the nature of the electronic transition, depends on whether it is direct or indirect, as well as

Accepted Manuscript

allowed or forbidden. In this study, the direct band gap was calculated using n=1/2. Fig. 5 depicts the $(\alpha h\nu)^2$ versus hν plots for both singly Ho^{3+} doped and Ho^{3+}/Yb^{3+} doped phosphors. The calculated optical band gap was 4.71 eV for the Ho^{3+}-doped sample and 4.79 for the Ho^{3+}/Yb^{3+} doped, based on the extrapolated slope of the Tauc plot. These values are consistent with previously reported findings. A smaller band gap increases the number of ions reaching the excited state, thereby boosting photoluminescence (PL) intensity.

3.3 PL characteristics:

Fig. 6(a) depicts the photoluminescence (PL) excitation spectra for CBNO: xHo^{3+} phosphors (x= 1.0, 2.0, 3.0, 4.0 & 5.0 mol%) measured at the green emission wavelength of 545 nm. The spectra reveal several bands corresponding to transitions from the, ground state to various excited states, including , $^3H_5$ (362 nm), $^5G_4$ (388 nm), $^5G_5$ (419 nm), $^5G_6$ (456 nm), $^3K_8$ (477 nm), and $^5F_2$ (487 nm) [28,29]. These transitions are consistent with data previously reported by Carnall et al. Among them, the $^5I_8 \rightarrow {}^5G_6$ transition at 456 nm exhibits the highest intensity, making it the preferred wavelength for analyzing the visible photoluminescence spectra of Ho^{3+}-doped CBNO phosphors at varying concentrations.

Fig. 6(b) depicts the PL emission spectra of these phosphors under 456 nm excitation. A prominent emission band is observed at 545 nm, ascribed to the $^5F_4/{}^5S_2 \rightarrow {}^5I_8$ transition of Ho^{3+} ions [30–32]. Additionally, a weaker band is present at 658 nm, attributed to the $^5F_5 \rightarrow {}^5I_8$ transition. The 545 nm emission band is significantly more intense compared to the 658 nm band in the presented spectra.

The emission spectra of CBNO : xHo^{3+}/yYb^{3+} phosphors (x=3.0 mol%, & y = 1.0, 2.0, 3.0, 4.0 & 5.0 mol%) recorded at 456 nm excitation, are depicted in Fig. 7. When exposed to blue light, the co-doped CBNO phosphors exhibit visible emissions. Three prominent peaks are observed at wavelengths of 485 nm, 545 nm, and 658 nm, which correspond to the transitions $^5F_3 \rightarrow {}^5I_8$, $^5F_4/{}^5S_2 \rightarrow {}^5I_8$, and $^5F_5 \rightarrow {}^5I_8$ respectively [33,34]. As shown in Fig. 7, the emission intensity increases with higher concentrations of Yb^{3+}, which is attributed to energy transfer from Yb^{3+} to Ho^{3+} ions and cross-relaxation between them [25]. Additionally, up-conversion (UC) properties of Ho^{3+}/Yb^{3+} co-doped CBNO phosphors were investigated, given that Ho^{3+} ions alone show minimal emission under 980 nm excitation.

Accepted Manuscript

Fig. 8 represents the up-conversion (UC) spectra of the synthesized CBNO: Ho^{3+}/Yb^{3+} phosphors under 980 nm laser excitation, spanning the 475–750 nm range. The spectra show two distinct emission bands in the green, and red regions at 545 nm, and 658 nm, respectively. These bands correspond to transitions of Ho^{3+}ions from the $^5F_4/^5S_2$, and $^5F_5$ excited states to the $^5I_8$ ground state [34,35]. The intensity of the UC emission increases with rising Yb^{3+} ion concentration. Since Ho^{3+} ions have a low absorption cross-section at 980 nm, Yb^{3+} ions act as sensitizers, transferring energy to Ho^{3+} ions, which then act as activators. The spectra indicate a steady increase in UC emission as the Yb^{3+} concentration rises to 5.0 mol%, attributed to the effective energy transfer from Yb^{3+} to Ho^{3+} ions, enhancing luminescent efficiency and leading to stronger red and green emissions. Power-dependent photoluminescence studies were also carried out to further confirm the up-conversion process.

3.4 Pump power dependence analysis:

A CBNO sample co-doped with 3.0 mol% Ho^{3+} and 5.0 mol% Yb^{3+} was subjected to a 980 nm laser with adjustable power to explore energy transfer mechanisms between Ho^{3+} and Yb^{3+} ions. Figure 9(a) illustrates the variations in green and red emission intensities as a function of pump power. The up-conversion (UC) spectra reveal that the green emission is more prominent than the red emission. At lower pump powers, the UC intensity increases significantly but begins to level off at higher powers, likely due to heat generation within the material. The relationship between emission intensity and pump power generally follows a power-law trend [36,37]:

$$I_{up} = (P_{in})^n \qquad (2)$$

The parameters I_{up}, P_{in}, and n correspond to the intensity of up-conversion (UC) luminescence, the power of the pump laser, and the number of photons involved in the process, respectively. Figure 9(b) illustrates a log relationship between pump power and emission intensity for the sample. The slopes (n) determined for the 545 nm and 658 nm emissions were 1.53 and 1.97, respectively. These values are consistent with the expected value of approximately 2, characteristic of two-photon processes. This indicates that the green and red emissions in the phosphor result from two-photon excitation mechanisms.

3.5 Energy transfer mechanism:

Accepted Manuscript

The energy transfer mechanism is depicted in the energy level diagram for Ho^{3+} and Yb^{3+} ions in Fig. 10. Yb^{3+} ions transition from the $^2F_{7/2}$ ground state to the $^2F_{5/2}$ excited state when excited by a 980 nm laser. In this process, Yb^{3+} ions act as sensitizers, transferring energy to Ho^{3+} ions via phonon-assisted mechanisms, which excite the Ho^{3+} ions to the $^5I_6$ state. Once in the $^5I_6$ state, some Ho^{3+} ions are further excited to higher energy states, such as $^5F_4$ and $^5F_5$, through either excited-state absorption (ESA) or energy transfer (ET) from Yb^{3+} ions [19,38]. These ions in higher states relax non-radiatively to lower energy levels. From there, radiative transitions occur from the $^5S_2/^5F_4$, and $^5F_5$ states to the $^5I_8$ state, leading to green, and red emissions, respectively [20]. These emissions can involve two -photon processes. The two photon emission arises from the accumulation of energy in the $^5S_2/^5F_4$, and $^5F_5$ levels through sequential ESA, and ET, processes, as shown in Fig. 10.

3.6 Colorimetric coordinates:

Figure 11(a) and (b) display the colorimetric properties of the as-prepared phosphors, derived from DC and UC photoluminescence emission data respectively. The CIE coordinates for both Ho^{3+}-doped and Ho^{3+}/Yb^{3+} co-doped samples under 456 nm and 980 nm excitations are summarized in Table 1. As depicted in the figures, all the CIE coordinates fall within the green region. The correlated color temperature (CCT) values were calculated using McCamy's formula [39]

$$CCT = -449n^3 + 3525\,n^2 - 6823.2\,n + 5520.3 \quad (3)$$

The parameter n is calculated using the formula n = (y−0.1858)/(x−0.3320), where x and y represent the calculated CIE coordinates. The corresponding CCT values are provided in Table 1. These findings indicate that the CIE coordinates and CCT values confirm the potential of the synthesized phosphors for solid-state lighting applications.

Color purity is a critical parameter for w-LED applications. For the synthesized phosphor samples, the color purity was estimated from the equation given below [24,40]:

$$\text{Color purity} = \frac{\sqrt{(x-x_{ee})^2+(y-y_{ee})^2}}{\sqrt{(x_d-x_{ee})^2+(y_d-y_{ee})^2}} \quad (4)$$

Accepted Manuscript

In this formula, (x,y) denotes the CIE chromaticity coordinates of the samples, the CIE coordinates for traditional white light are represented by (x_{ee},y_{ee}), while the CIE coordinates of the dominant wavelength are shown by (x_d,y_d). For the calculations, (x_{ee},y_{ee}) was set to (0.33, 0.33). The calculated color purity values for the optimized sample CBNO: xHo^{3+}/yYb^{3+} (x =3.0 mol%, y = 5.0 mol%) phosphor is 97.2%. Such high levels of color purity highlight the strong potential of these sample for use in w-LEDs. Therefore, CBNO: Ho^{3+}/Yb^{3+} stands out as a viable candidate for applications in lighting and multicolor display systems.

3.7 PL Decay curves:

The up-conversion (UC) decay experiments were performed to study the UC mechanism and the energy transfer mechanisms between Ho^{3+} & Yb^{3+}. Fig. 12 displays the corresponding decay curves that were captured at 545 nm under pulsed 980 nm laser stimulation. The double-exponential function was used to fit the curves [2,41]:

$$I = I_0 + A_1 * \exp\left(-\frac{t}{\tau_1}\right) + A_2 * \exp(-\frac{t}{\tau_2}) \qquad (5)$$

Here, t stands for time, τ_1 & τ_2 are the corresponding decay durations, and A_1 and A_2 are constants. The up-conversion emission intensities at time t and the beginning time (t = 0) are denoted by I and I_0, respectively. The following formula was used to determine the average decay lifetimes for each sample [42,43]:

$$\tau_{exp} = \frac{A_1\tau_1^2 + A_2\tau_2^2}{A_1\tau_1 + A_2\tau_2} \qquad (6)$$

The Table 1 shows the average decay durations at 545 nm emission for the CBNO: 3.0 mol% Ho^{3+}/yYb^{3+} phosphors generated by a 980 nm laser with y = 1.0, 2.0, 3.0, 4.0, and 5.0 mol%). Higher Yb^{3+} concentrations result in lower decay lifetimes for 545 nm emission wavelength, suggesting that energy transfer from Yb^{3+} to Ho^{3+} becomes more effective with increasing Yb^{3+} concentrations. The up-conversion process may contain additional energy transfer mechanisms or migration paths, which could account for the single-exponential function's incomplete fit to the decay curves. Fig. 12 shows the trend of decay lifespan values with different Yb^{3+} concentrations.

3.8 Temperature Dependent PL(TD-PL):

The thermal stability of phosphors is a critical aspect of their performance. Fig. 13(a) presents the temperature-dependent photoluminescence (TD-PL) spectra of CBNO: xHo^{3+}/yYb^{3+} (x

Accepted Manuscript

=3.0 mol%, y = 5.0 mol%) phosphor, recorded under 980 nm laser source over the temperature range of 300 to 410 K. The spectra demonstrate a gradual decrease in PL intensity as the temperature increases, without any observable shift in the emission peaks. This reduction in PL intensity is attributed to non-radiative transitions caused by thermal quenching. As the temperature increases, the lattice undergoes slight expansion and structural disorder, which alters the local crystal field surrounding the luminescent centers, thereby diminishing emission intensity [44,45]. The activation energy (ΔE) is determined using the Arrhenius equation, expressed as follows [46,47]:

$$I_T = \frac{I_0}{1 + C\exp\left(-\frac{\Delta E}{K_B T}\right)} \quad (7)$$

Here, I_0 and I_T represent the emission intensities at 300 K and temperature T (in Kelvin), respectively, C is a constant, and k_B denotes the Boltzmann constant (8.629×10^{-5}eV/K).The plot of $\ln(I_T/I_0-1)$ against $1/k_BT$ is shown in Fig. 13(b). The activation energy, calculated from the slope of this graph, is determined to be 0.293 eV. This value is higher than those reported in the literature, indicating superior thermal stability of the phosphor [25,48,49]. These results, derived from the TD-PL spectra of CBNO: x Ho^{3+}/y Yb^{3+} (x =3.0 mol%, y = 5.0 mol%) doped phosphor, confirm its excellent thermal stability.

4. Conclusion:

A series of CBNO phosphors doped with Ho^{3+} ions and co-doped with Ho^{3+}/Yb^{3+} ions were synthesized using a solid-state reaction technique. Morphological studies indicated that the phosphor particles are in the micrometer size range, and elemental mapping confirmed the even distribution of all the constituent elements across the phosphor surface. Both up-conversion (UC) and down-conversion (DC) photoluminescence (PL) studies demonstrated intense green light emission at 545 nm, with UC emission excited by a 980 nm laser and DC emission excited by a 456 nm xenon lamp. The relationship between pump power and UC intensity provided details about the photon interaction mechanisms in the UC process. Decay measurements under 980 nm excitation revealed a steady decrease in lifetime at 545 nm. Chromaticity coordinates analysis placed the phosphor emissions in the green region. These results highlight the material's potential for various applications, including green light-emitting diodes (LEDs), display systems, security printing, phototherapy, and solid-state lighting technologies.

Acknowledgements:

Accepted Manuscript

One of the authors, Ms. Sheetal Kumari is very much thankful to CSIR/UGC for the award of a Senior Research Fellowship (SRF) (UGC Ref. 1431/CSIR-UGC NET JUNE 2019). Dr. Aman Prasad is grateful to Dayalbagh Educational Institute (DEI), Deemed University, Agra for providing a Research Seed Money Grant (DEI/GBMF(3052022)/518). He also graciously acknowledges Prof. Rohit Srivastava, Head, Department of Chemistry, Dayalbagh Educational Institute (DEI), Deemed University, Agra for his continued support to our research work. We are really grateful to Dr. Manju Srivastava and Dr. Shyama Prasad, Department of Chemistry, Dayalbagh Educational Institute (DEI), Deemed University, Agra for making the FE-SEM facility available to us.

Author Contributions:

Ms. Sheetal Kumari: work plan, experiments, calculation and writing the manuscript

Dr. Pooja Rohilla: Data analysis

Dr. Anu: Data analysis

Mr. Srinibas Barik: Characterisation, data collation

Mrs. Bhanupriya Yadav: Characterisation

Dr. Kaushal Kumar: Work Plan, review

Dr. Runit Isaac: Methodology, Analysis

Prof. A.S. Rao: Work plan, correction of manuscript.

Dr. Aman Prasad: Supervision, validation, review

Data Availability:

The data that support the findings of this study are available on request from the corresponding author Dr. Aman Prasad. The data are not publicly available due to restrictions, e.g. their containing information that could compromise the privacy of research participants.

Declaration:

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References:

Accepted Manuscript

[1] D. Ramachari, L. Rama Moorthy, C.K. Jayasankar, Spectral investigations of Sm^{3+} -doped oxyfluorosilicate glasses, Mater. Res. Bull. 48 (2013) 3607–3613. https://doi.org/10.1016/j.materresbull.2013.05.081.

[2] D.V.K. Reddy, T.S. Rao, S. Taherunnisa, A. Suchocki, Y. Zhydachevskyy, M. Piasecki, M.R. Reddy, Tunable white light by varying excitations in yttrium alumino bismuth borosilicate glasses co-doped with Dy^{3+} -Eu^{3+} for cool WLED applications, J. Non. Cryst. Solids. 513 (2019) 167–182. https://doi.org/10.1016/j.jnoncrysol.2019.03.011.

[3] Y. Zhong, P. Sun, X. Gao, Q. Liu, S. Huang, B. Liu, B. Deng, R. Yu, Synthesis and optical properties of new red-emitting $SrBi_2Ta_2O_9$: Eu^{3+} phosphor application for w-LEDs commercially based on InGaN, J. Lumin. 212 (2019) 45–51. https://doi.org/10.1016/j.jlumin.2019.03.057.

[4] X. Geng, Y. Xie, S. Chen, J. Luo, S. Li, T. Wang, S. Zhao, H. Wang, B. Deng, R. Yu, W. Zhou, Enhanced local symmetry achieved zero-thermal-quenching luminescence characteristic in the Ca_2InSbO_6:Sm^{3+} phosphors for w-LEDs, Chem. Eng. J. 410 (2021) 128396. https://doi.org/10.1016/j.cej.2020.128396.

[5] Q. Yang, G. Li, Y. Wei, H. Chai, Synthesis and photoluminescence properties of red-emitting $NaLaMgWO_6$:Sm^{3+},Eu^{3+} phosphors for white LED applications, J. Lumin. 199 (2018) 323–330. https://doi.org/10.1016/j.jlumin.2018.03.011.

[6] Zijun Chen, Huiyi Xu, Chunyan Cao , Xiaoting Chen, Min Zhang, Minkun Jian, Yuechan Li and An Xie, Synthesis, Luminescent Properties and White LED Fabrication of Sm^{3+} Doped Lu_2WMoO_9, (2021). 10.3390/coatings11040403

[7] Y. Wang, Y. Ke, S. Chen, J. Luo, S. Shu, J. Gao, B. Deng, R. Yu, Luminescence investigation of red-emitting Sr_2MgMoO_6:Eu^{3+} phosphor for visualization of latent fingerprint, J. Colloid Interface Sci. 583 (2021) 89–99. https://doi.org/10.1016/j.jcis.2020.09.024.

[8] M.A. Hernández-Rodríguez, P. Tadge, A.D. Lozano-Gorrín, U.R. Rodríguez-Mendoza, S. Ray, V. Lavín, Upconversion and luminescence temperature sensitivity of Er^{3+} ions in yttrium oxysulfate nanophosphor, Opt. Mater. (Amst). 95 (2019) 109197. https://doi.org/10.1016/j.optmat.2019.109197.

Accepted Manuscript

[9] S. Dong, S. Ye, L. Wang, X. Chen, S. Yang, Y. Zhao, J. Wang, X. Jing, Q. Zhang, $Gd_3B(W,Mo)O_9$: Eu^{3+} red phosphor: From structure design to photoluminescence behavior and near-UV white-LEDs performance, J. Alloys Compd. 610 (2014) 402–408. https://doi.org/10.1016/j.jallcom.2014.05.011.

[10] S. Shu, Y. Wang, Y. Ke, B. Deng, R. Liu, Q. Song, J. Wang, R. Yu, $NaCaTiTaO_6$:Sm^{3+}: A novel orange-red-emitting tantalate phosphor with excellent thermal stability and high color purity for white LEDs, J. Alloys Compd. 848 (2020) 156359. https://doi.org/10.1016/j.jallcom.2020.156359.

[11] C.R. Kesavulu, H.J. Kim, S.W. Lee, J. Kaewkhao, N. Chanthima, Y. Tariwong, Physical, vibrational, optical and luminescence investigations of Dy^{3+}-doped yttrium calcium silicoborate glasses for cool white LED applications, J. Alloys Compd. 726 (2017) 1062–1071. https://doi.org/10.1016/j.jallcom.2017.08.091.

[12] S.R. Alharbi, M. Alhassan, O. Jalled, S. Wageh, A. Saeed, Structural characterizations and electrical conduction mechanism of $CaBi2Nb_2O_9$ single-phase nanocrystallites synthesized via sucrose-assisted sol–gel combustion method, J. Mater. Sci. 53 (2018) 11584–11594. https://doi.org/10.1007/s10853-018-2458-2.

[13] O. Jalled, Z. Alharbi, S.R. Alharbi, A. Saeed, M. Alhassan, S. Al-Heniti, H.Y. Mohammed, Y. Al-Hadeethi, F. Al-Marzouki, A. Al-Mujtaba, Synthesis and dielectric properties of nanocrystalline strontium bismuth niobate, J. Nanosci. Nanotechnol. 17 (2017) 594–600. https://doi.org/10.1166/jnn.2017.12459.

[14] T.S. Sreena, P.P. Rao, A.K.V. Raj, T.R.A. Thara, Narrow-band red-emitting phosphor, $Gd_3Zn_2Nb_3O_{14}$:Eu^{3+} with high color purity for phosphor-converted white light emitting diodes, J. Alloys Compd. 751 (2018) 148–158. https://doi.org/10.1016/j.jallcom.2018.04.135.

[15] H.C. Gupta, Archana, V. Luthra, Lattice vibrations of $ABi_2Nb_2O_9$ crystals (A = Ca, Sr, Ba), Vib. Spectrosc. 56 (2011) 235–240. https://doi.org/10.1016/j.vibspec.2011.03.002.

[16] K. Miura, Electronic properties of ferroelectric $SrBi_2Ta_2O_9$, $SrBi_2Nb_2O_9$, and $PbBi_2Nb_2O_9$ with optimized structures, Appl. Phys. Lett. 80 (2002) 2967–2969. https://doi.org/10.1063/1.1474607.

Accepted Manuscript

[17] V.A. Online, A.A. Saboury, F. Davar, A. Semnani, RSC Advances, (2016). https://doi.org/10.1039/C6RA08116A.This.

[18] V. Singh, M. Seshadri, D. Taikar, S.J. Dhoble, R.S. Yadav, Optical transitions and radiative properties of green emitting Ho^{3+}:YVO_4 phosphor, RSC Adv. 13 (2023) 3592–3601. https://doi.org/10.1039/d2ra06287a.

[19] S. Liu, S. Liu, H. Ming, D. Hou, W. You, Investigation on the upconversion luminescence in Ho^{3+} / Yb^{3+} co-doped, Mater. Res. Bull. 98 (2018) 187–193. https://doi.org/10.1016/j.materresbull.2017.10.025.

[20] K. Dharendra, A. Roy, A. Dwivedi, A. Kumar, S. Bahadur, D. Kumar, Two steps enhancement of dual mode (UC and DS) behaviour of Ho^{3+} / Yb^{3+} and Tm^{3+}/ Yb^{3+} co-doped $GdVO_4$ phosphors : Improvement in spectral and color purity, Mater. Res. Bull. 139 (2021) 111280. https://doi.org/10.1016/j.materresbull.2021.111280.

[21] P. Tadge, I.R. Martín, S.B. Rai, S. Sapra, T.M. Chen, V. Lavín, R.S. Yadav, S. Ray, Intense green photoluminescence and upconversion-based intrinsic optical bistability in Ho^{3+} / Yb^{3+} codoped $CaSc2O_4$ and upconversion in Ho^{3+} / Yb^{3+} codoped CaY_2O_4 phosphors, J. Lumin. 252 (2022) 119261. https://doi.org/10.1016/j.jlumin.2022.119261.

[22] P. Tadge, R.S. Yadav, P.K. Vishwakarma, S.B. Rai, T.M. Chen, S. Sapra, S. Ray, Enhanced photovoltaic performance of Y_2O_3: Ho^{3+} / Yb^{3+} upconversion nanophosphor based DSSC and investigation of color tunability in Ho3+/Tm3+/Yb3+ tridoped Y2O3, J. Alloys Compd. 821 (2020) 153230. https://doi.org/10.1016/j.jallcom.2019.153230.

[23] M. Vega, I.R. Martin, J. Llanos, Near-infrared to visible upconversion and second harmonic generation in $BaTiO_3$: Ho^{3+} and $BaTiO_3$: Ho^{3+} / Yb^{3+} phosphors, J. Alloys Compd. (2019). https://doi.org/10.1016/j.jallcom.2019.07.311.

[24] S. Kumari, P. Rohilla, S. Diwakar, R.A. Talewar, A. Shandilya, Y. Tayal, K. Swapna, A. Prasad, A.S. Rao, Synthesis, analysis and characterizations of Dy^{3+} ions-doped $CaBi_2Nb_2O_9$ phosphors for optoelectronic device applications, J. Mater. Sci. Mater. Electron. 35 (2024). https://doi.org/10.1007/s10854-024-12640-2.

[25] S. Kumari, A. Prasad, A.S.R.R.K. Sinha, Up - Conversion Luminescence in -

Accepted Manuscript

$Sr_9Y_2W_4O_{24}$: Ho^{3+} / Yb^{3+} Phosphors for applications in Temperature Sensing and w - LEDs, (2024).

[26] B.C. Jamalaiah, P.S. Khan, $Sr_3Gd(PO_4)_3$: Dy^{3+} phosphors for lighting applications, J. Sol-Gel Sci. Technol. 105 (2023) 266–277. https://doi.org/10.1007/s10971-022-05995-7.

[27] K.S. Rao, B.V.N. Kumar, Y.N. Rajeev, K. Venkatarao, S. Cole, Structural , morphological and photoluminescence properties of Eu^{3+} doped $Cd_2Sr(PO_4)_2$ nanopowder, J. Mol. Struct. 1284 (2023) 135406. https://doi.org/10.1016/j.molstruc.2023.135406.

[28] V. Singh, K. V Dabre, S.J. Dhoble, G. Lakshminarayana, Optik Green emitting holmium (Ho) doped yttrium oxide (Y_2O_3) phosphor for solid state lighting, Opt. - Int. J. Light Electron Opt. 206 (2020) 164339. https://doi.org/10.1016/j.ijleo.2020.164339.

[29] P. Kaur, A. Kaur, S. Singh, L. Singh, Investigation on structural and thermoluminescence properties of Ho^{3+} doped SrB_4O_7 phosphor for dosimetry applications, J. Mol. Struct. 1248 (2022) 131500. https://doi.org/10.1016/j.molstruc.2021.131500.

[30] V. Singh, C.B. Annapurna, S. Kaur, A.S. Rao, Optik Optical properties of $Sr_2La_8(SiO_4)_6O_2$ doped with Ho^{3+} phosphor, Optik (Stuttg). 242 (2021) 167268. https://doi.org/10.1016/j.ijleo.2021.167268.

[31] A. Marimuthu, R.N. Perumal, Investigations on structural, optical and electrical properties of holmium doped rubidium titanyl phosphate single crystals, J. Alloys Compd. 895 (2022) 162568. https://doi.org/10.1016/j.jallcom.2021.162568.

[32] V.R. Prasad, S. Damodaraiah, S.N. Devara, Y.C. Ratnakaram, Photoluminescence studies on holmium (III) and praseodymium (III) doped calcium borophosphate (CBP) phosphors, J. Mol. Struct. 1160 (2018) 383–392. https://doi.org/10.1016/j.molstruc.2018.02.034.

[33] P.K. Vishwakarma, S.B. Rai, A. Bahadur, Intense red and green emissions from Ho^{3+} / Yb^{3+} co-doped Sodium Gadolinium Molybdate Nano-phosphor : Effect of calcination temperature and Intrinsic optical bistability, Mater. Res. Bull. 133 (2021) 111041.

Accepted Manuscript

https://doi.org/10.1016/j.materresbull.2020.111041.

[34] T. Zhang, W. Yang, H. Li, X. Wang, X. Jiang, Z. Leng, Enhanced red emission of Yb , Ho : NaYGd $(WO_4)_2$ phosphors by codoping Ce^{3+}, 240 (2021). https://doi.org/10.1016/j.jlumin.2021.118432.

[35] G.B. Nair, A. Kumar, H.C. Swart, S.J. Dhoble, Facile precipitation synthesis of green-emitting BaY_2F_8 : Ho^{3+} / Yb^{3+} upconverting phosphor, Ceram. Int. 45 (2019) 14205–14213. https://doi.org/10.1016/j.ceramint.2019.04.127.

[36] X. Cheng, X. Dong, K. Peng, H. Zhang, Y. Su, L. Jiang, Upconversion Luminescence and Optical Temperature-Sensing Properties of $LaNbO_4$:Yb^{3+}/Er^{3+} Phosphors, J. Electron. Mater. 49 (2020) 518–523. https://doi.org/10.1007/s11664-019-07776-5.

[37] H. Suo, C. Guo, W. Wang, T. Li, C. Duan, M. Yin, Mechanism and stability of spectrally pure green up-conversion emission in Yb^{3+}/Ho^{3+} co-doped $Ba_5Gd_8Zn_4O_{21}$ phosphors, Dalt. Trans. 45 (2016) 2629–2636. https://doi.org/10.1039/c5dt02985f.

[38] Y. Zhou, H. Wang, B. Xu, L. Cao, Concentration-dependent up-converted luminescence of Ho^{3+} / Yb^{3+} co-doped yttrium-stabilized zirconia phosphors, J. Lumin. 161 (2015) 288–292. https://doi.org/10.1016/j.jlumin.2015.01.019.

[39] H. George, N. Deopa, S. Kaur, A. Prasad, M. Sreenivasulu, M. Jayasimhadri, A.S. Rao, Judd-Ofelt parametrization and radiative analysis of Dy^{3+} ions doped Sodium Bismuth Strontium Phosphate glasses, J. Lumin. 215 (2019) 116693. https://doi.org/10.1016/j.jlumin.2019.116693.

[40] S. Kumari, A.S. Rao, R.K. Sinha, Investigations on photoluminescence and energy transfer studies of Sm^{3+} and Eu^{3+} ions doped $Sr_9Y_2W_4O_{24}$ red emitting phosphors with high color purity for w-LEDs, J. Mol. Struct. 1295 (2024) 136507. https://doi.org/10.1016/j.molstruc.2023.136507.

[41] Q. Long, C. Wang, Y. Li, J. Ding, X. Wang, Y. Wang, Solid state reaction synthesis and photoluminescence properties of Dy^{3+} doped $Ca_3Sc_2Si_3O_{12}$ phosphor, Mater. Res. Bull. 71 (2015) 21–24. https://doi.org/10.1016/j.materresbull.2015.07.001.

[42] S. Pradhan, H. Kaur, M. Jayasimhadri, Photoluminescence and thermal sensing properties of Er^{3+} doped silicate based phosphors for multifunctional optoelectronic device applications, Ceram. Int. 47 (2021) 27694–27701.

Accepted Manuscript

https://doi.org/10.1016/j.ceramint.2021.06.194.

[43] A. George, A. Jose, T. Krishnapriya, P.R. Biju, Synthesis and spectroscopic characterization of $Ba_2Sb_2O_7:Eu^{3+}$ orange-red emitting phosphor for WLEDs, Mater. Lett. 300 (2021) 130266. https://doi.org/10.1016/j.matlet.2021.130266.

[44] R.S. Yadav, S.J. Dhoble, S.B. Rai, Enhanced photoluminescence in Tm^{3+}, Yb^{3+}, Mg^{2+} tri-doped $ZnWO_4$ phosphor: Three photon upconversion, laser induced optical heating and temperature sensing, Sensors Actuators, B Chem. 273 (2018) 1425–1434. https://doi.org/10.1016/j.snb.2018.07.049.

[45] Monika, R.S. Yadav, A. Rai, S.B. Rai, NIR light guided enhanced photoluminescence and temperature sensing in Ho^{3+} / Yb^{3+}/Bi^{3+} co-doped $ZnGa_2O_4$ phosphor, Sci. Rep. 11 (2021) 1–17. https://doi.org/10.1038/s41598-021-83644-9.

[46] V. Chauhan, P. Deshmukh, S. Satapathy, P.C. Pandey, Results in Physics Greenish-yellow emission from rare-earth free Li_+ doped zinc vanadate phosphor, Results Phys. 39 (2022) 105689. https://doi.org/10.1016/j.rinp.2022.105689.

[47] Monika, R.S. Yadav, A. Bahadur, S.B. Rai, Concentration and pump power-mediated color tunability, optical heating and temperature sensing via TCLs of red emission in an Er^{3+}/Yb^{3+}/Li_+ co-doped $ZnGa_2O_4$ phosphor, RSC Adv. 9 (2019) 40092–40108. https://doi.org/10.1039/c9ra09120c.

[48] L. Wu, Y. Bai, L. Wu, H. Yi, Y. Kong, Y. Zhang, J. Xu, Sm^{3+} and Eu^{3+} codoped $SrBi_2B_2O_7$: A red-emitting phosphor with improved thermal stability, RSC Adv. 7 (2017) 1146–1153. https://doi.org/10.1039/c6ra26752a.

[49] G. Zhu, Z. Li, F. Zhou, C. Wang, S. Xin, A novel temperature sensitive Sm^{3+} doped niobate orange-red phosphor : The synthesis and characteristic luminescent property investigation, J. Lumin. 196 (2018) 32–35. https://doi.org/10.1016/j.jlumin.2017.12.014.

Accepted Manuscript

Figure Captions:

Fig. 1. Flow chart of the synthesis process of CBNO phosphor.

Fig. 2. XRD patterns for undoped, single Ho^{3+}, and Ho^{3+}/Yb^{3}CBNO phosphor

Fig. 3. (a-c) Scanning electron microscope images for undoped, single Ho^{3+}, and Ho^{3+}/Yb^{3}CBNO phosphor (d-f) their corresponding EDX images

Fig. 4. DRS spectra for of the 3.0 mol% Ho^{3+}, and 3.0 mol% Ho^{3+}/5.0 mol% Yb^{3+} doped CBNO phosphors.

Fig. 5. Tauc's plot for direct band for CBNO phosphors.

Fig. 6. (a) PL excitation spectra for CBNO: xHo^{3+} (1.0, 2.0, 3.0, 4.0 & 5.0 mol%) doped phosphors under 545 nm emission wavelength. (b) PL emission spectra of CBNO: xHo^{3+} (1.0, 2.0, 3.0, 4.0 & 5.0 mol%) doped phosphors under excitation wavelength 456 nm.

Fig. 7. Linear emission spectra under 456 nm excitation for the CBNO: xHo^{3+}/yYb^{3+} (x = 3.0 mol%, & y = 1.0, 2.0, 3.0, 4.0 & 5.0 mol%) phosphors.

Fig. 8. UC spectra of the CBNO: xHo^{3+}/yYb^{3+} (x = 3.0 mol%, & y = 1.0, 2.0, 3.0, 4.0 & 5.0 mol%) phosphors under 980 nm excitation.

Fig. 9. (a)Power dependence PL spectra of CBNO: xHo^{3+}/yYb^{3+} (x = 3.0 mol%, y = 5.0 mol%) phosphor under 980 nm laser excitation and (b) Ln-Ln plot of UC intensity vs pump power.

Fig. 10. Schematic Energy level diagram of Ho^{3+} and Yb^{3+} ions.

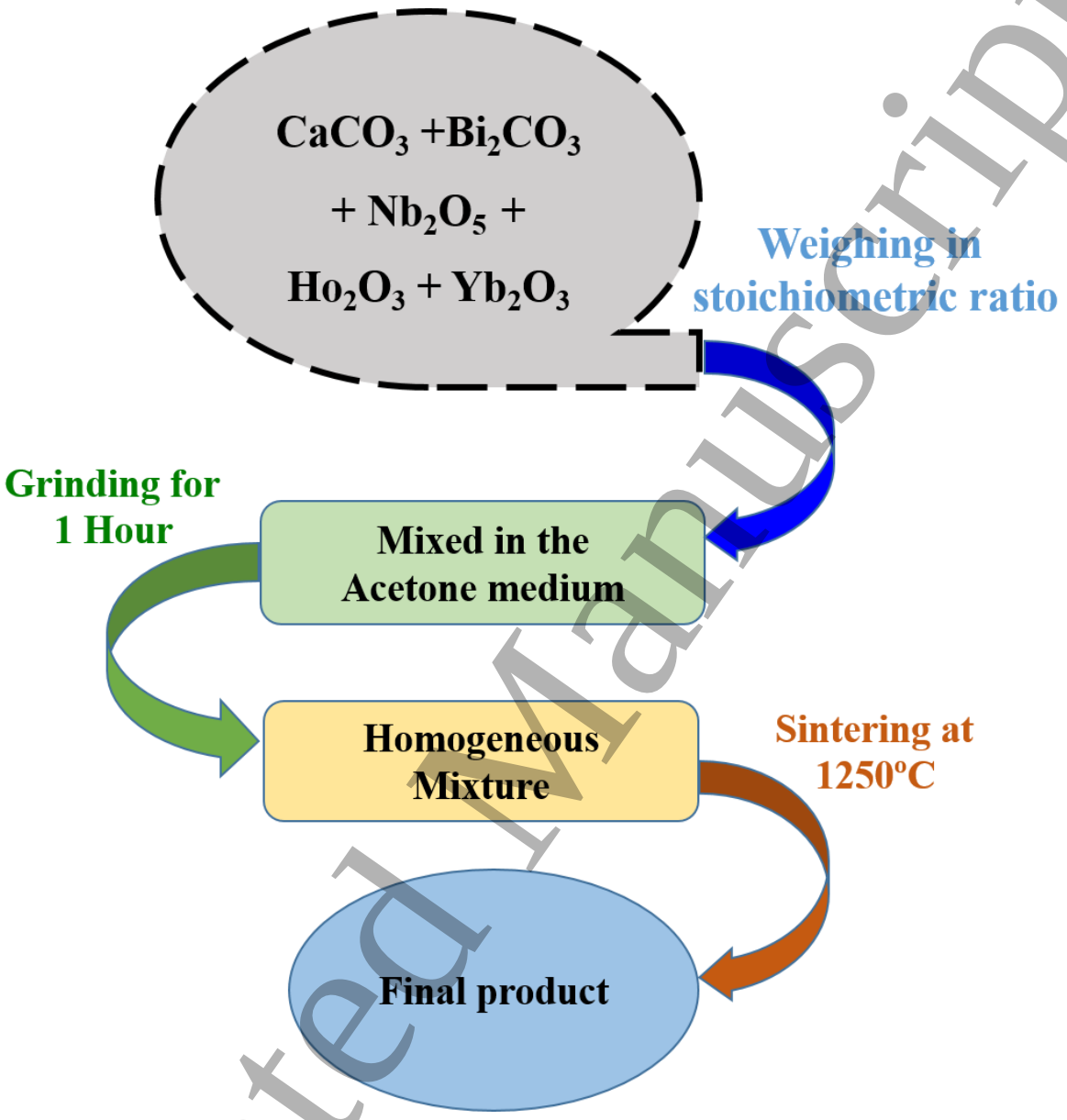
Fig. 11. CIE diagram for CBNO:Ho^{3+}/Yb^{3+} doped phosphors with excitation wavelengths of 456 nm and 980 nm.

Fig. 12. The decay curves of CBNO: xHo^{3+}/yYb^{3+} (x = 3.0 mol%, & y = 1.0, 2.0, 3.0, 4.0 & 5.0 mol%) phosphors under 980 nm excitation.

Fig. 13. (a) TD-PL spectra of CBNO: xHo^{3+}/yYb^{3+} (x = 3.0 mol%, y = 5.0 mol%) phosphor under 980 nm excitation (b) Linear relationship of $\ln [(I_0/I_T) - 1]$ vs $1/k_BT$ plot

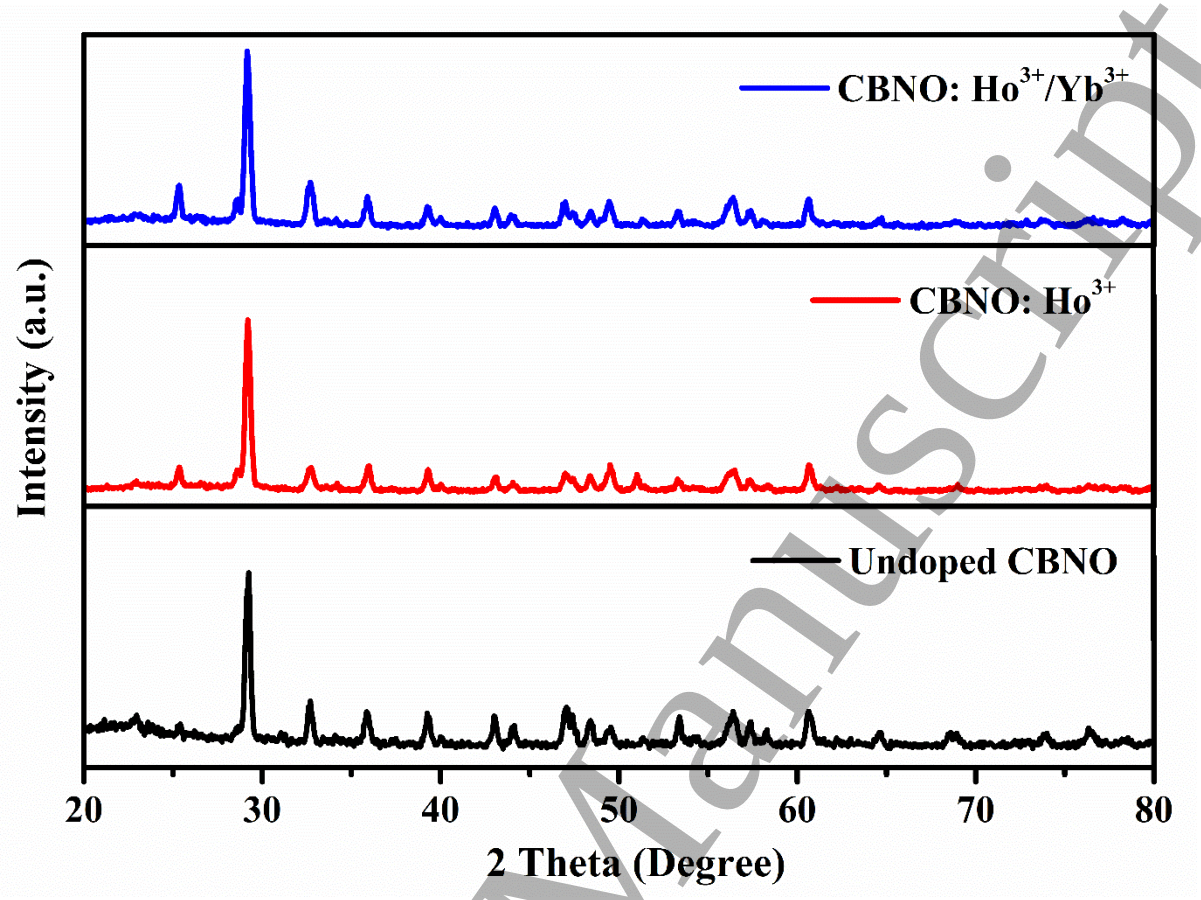
Accepted Manuscript

Fig. 1



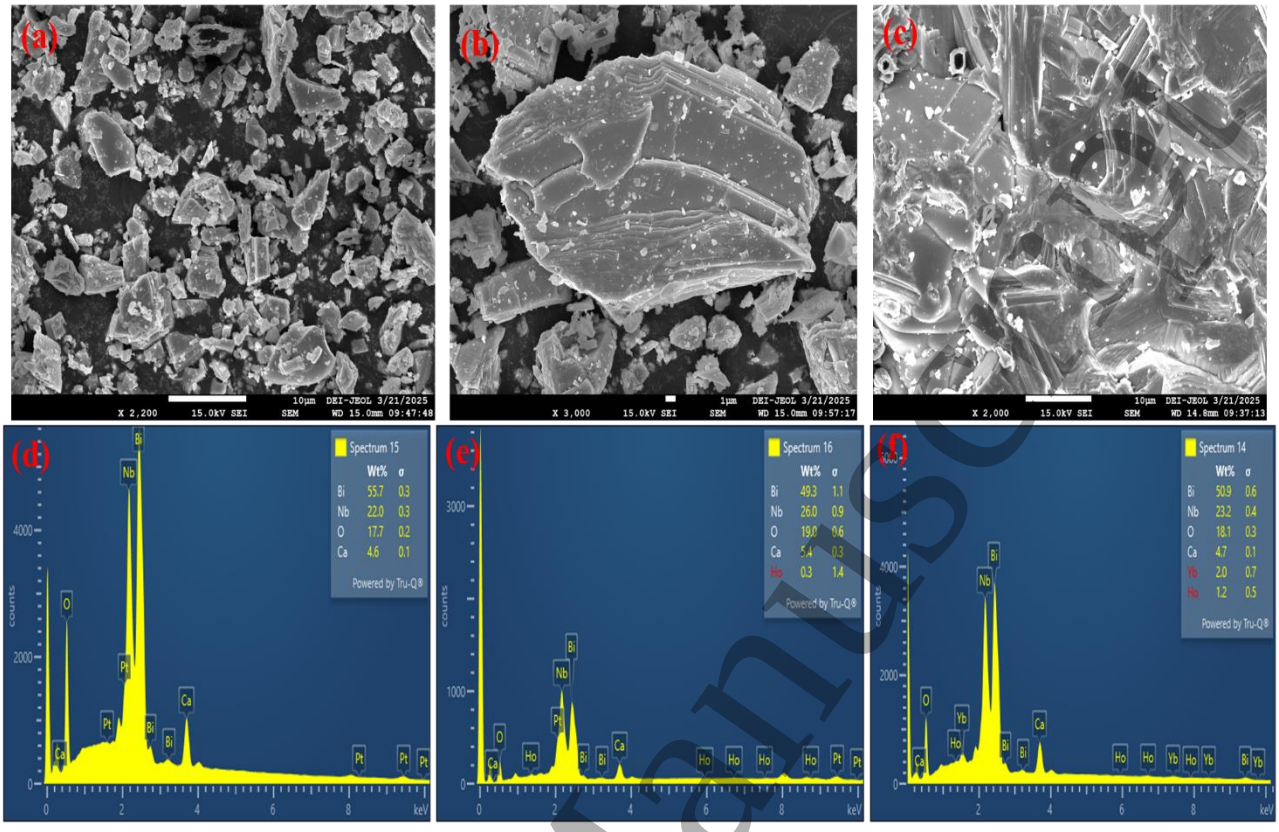
Accepted Manuscript

Fig. 2.



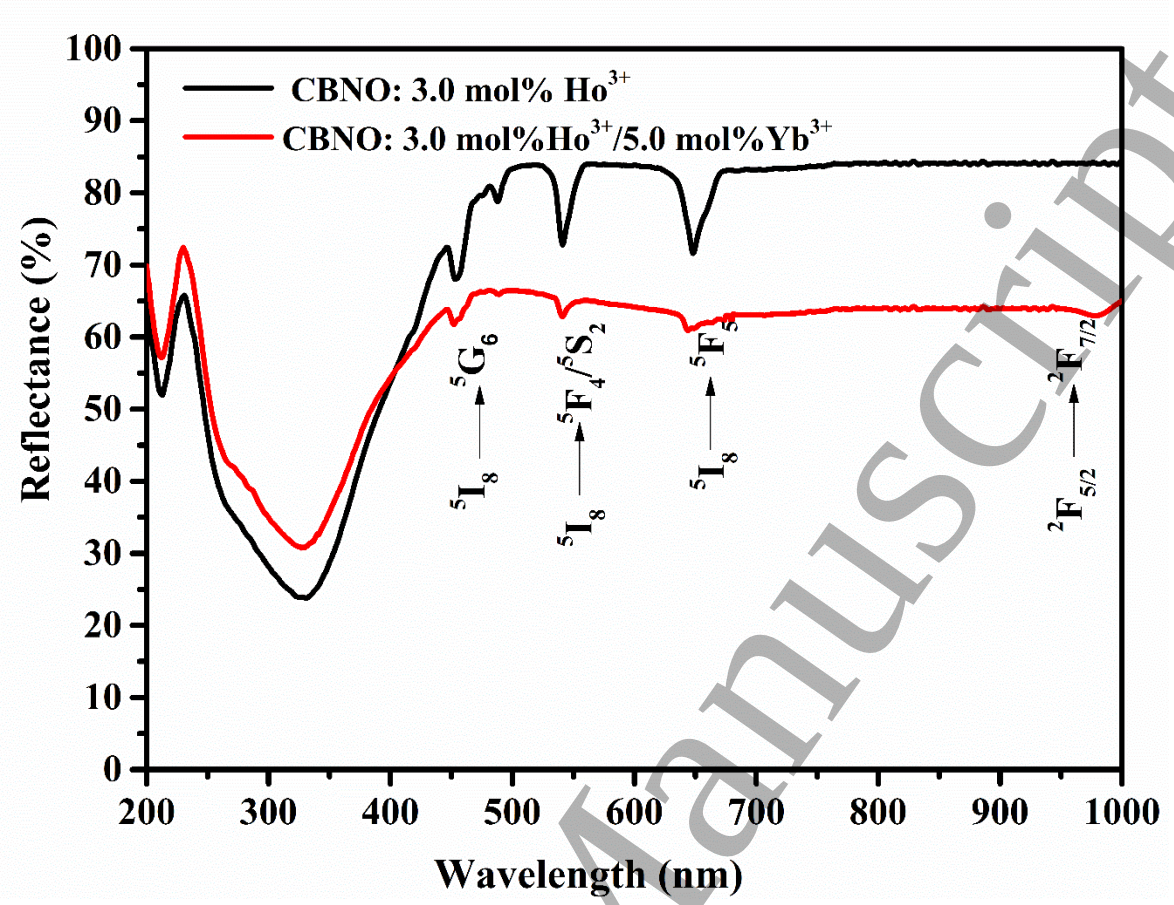
Accepted Manuscript

Fig. 3



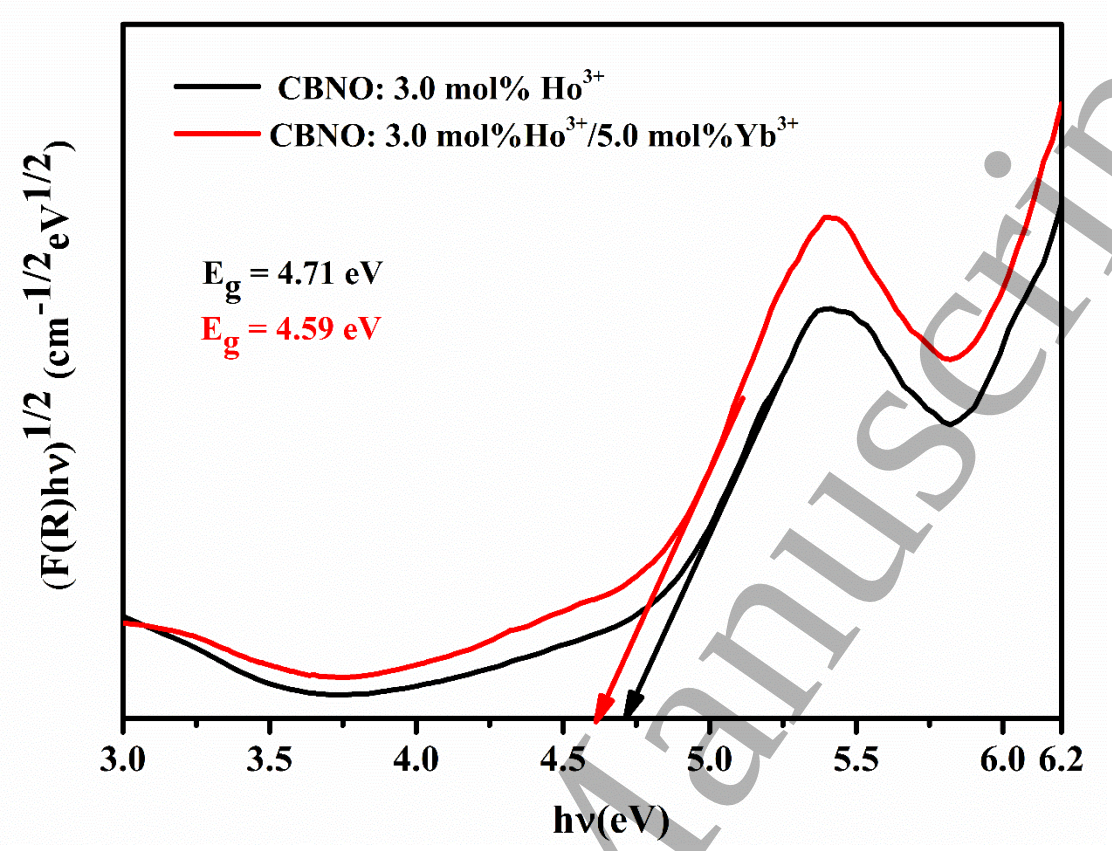
Accepted Manuscript

Fig. 4



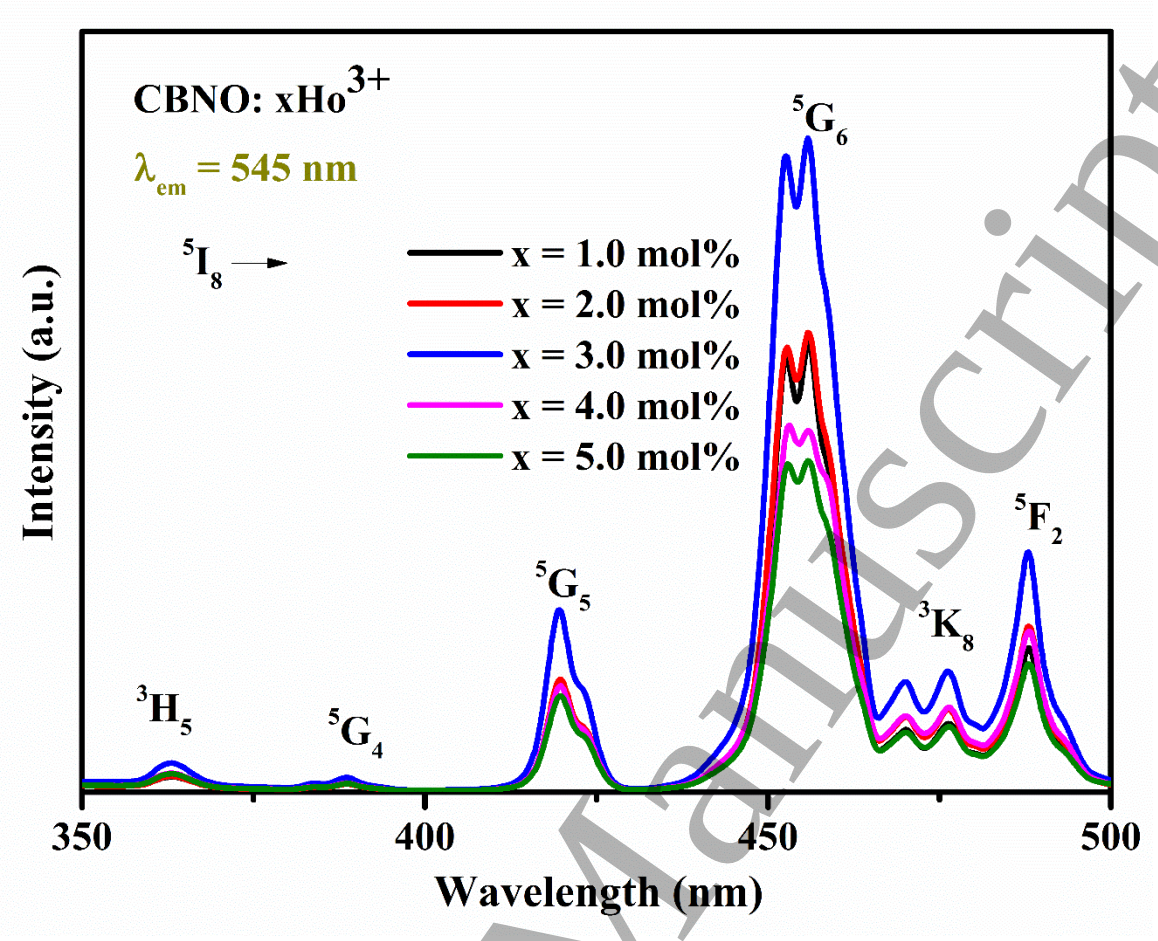
Accepted Manuscript

Fig. 5



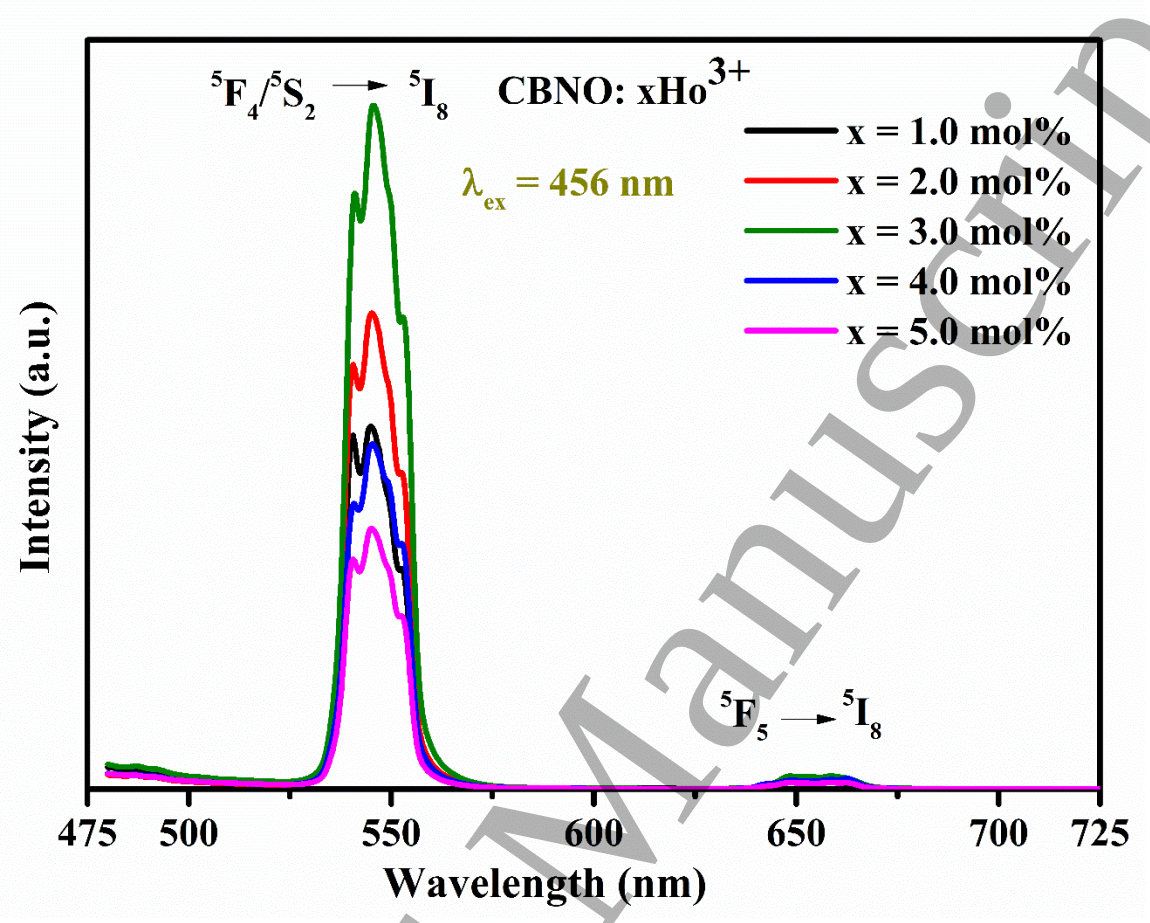
Accepted Manuscript

Fig. 6(a)



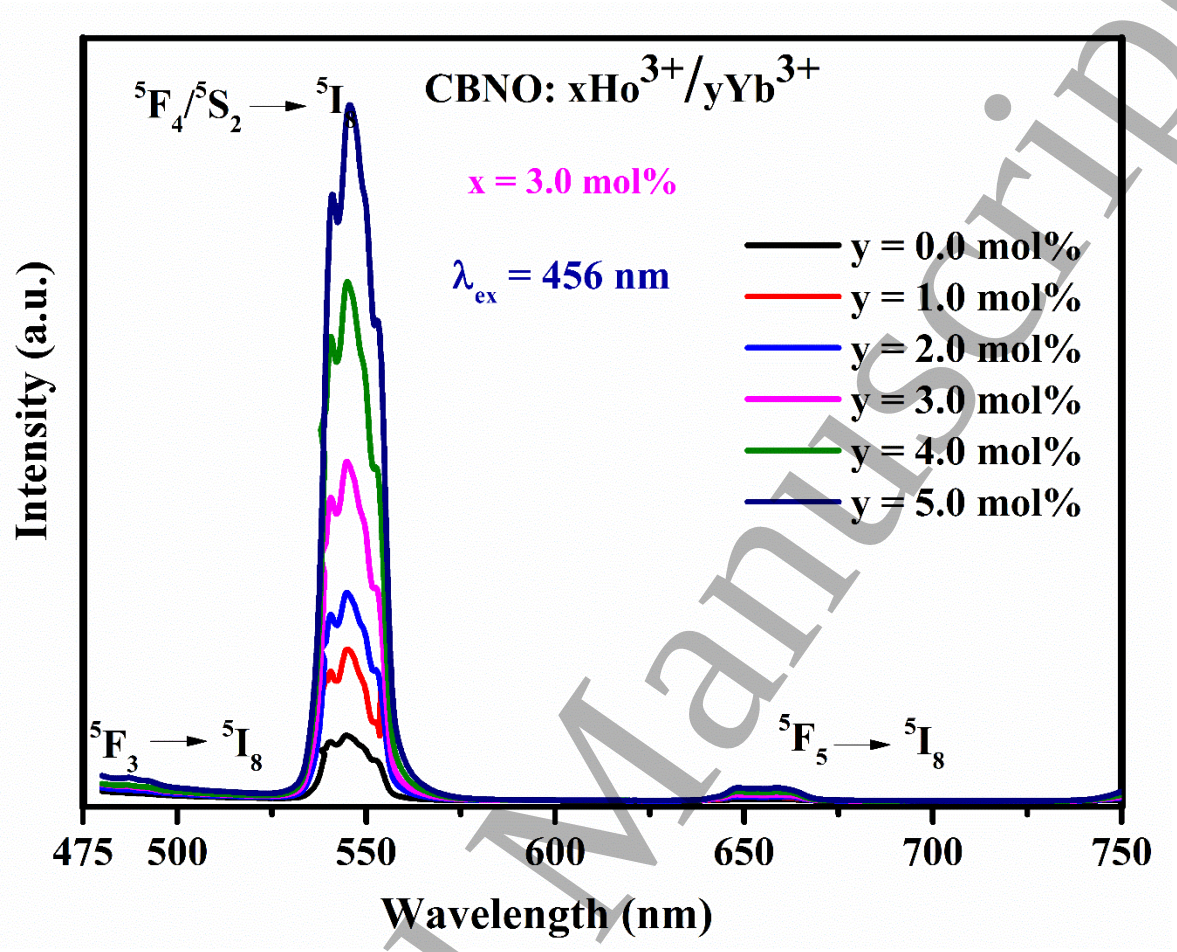
Accepted Manuscript

Fig. 6(b)



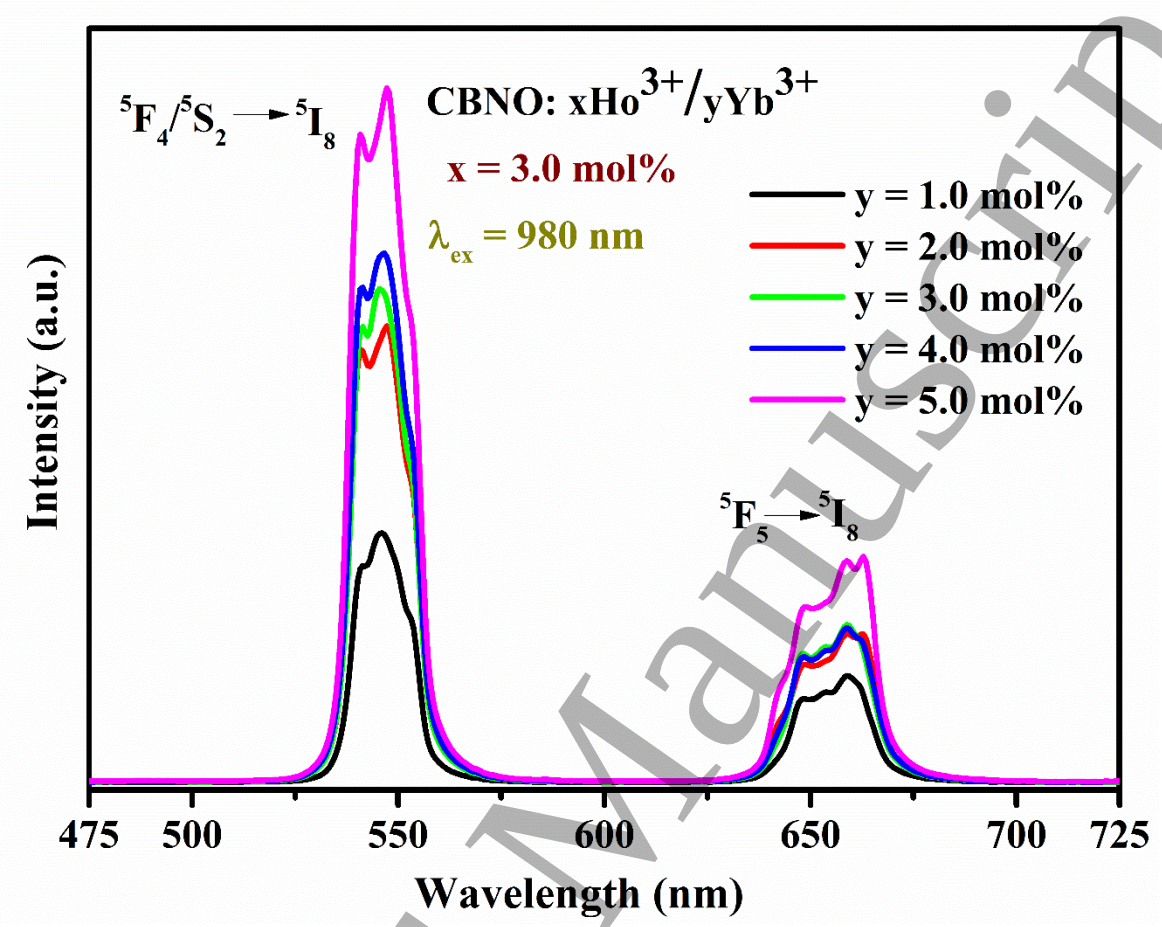
Accepted Manuscript

Fig. 7



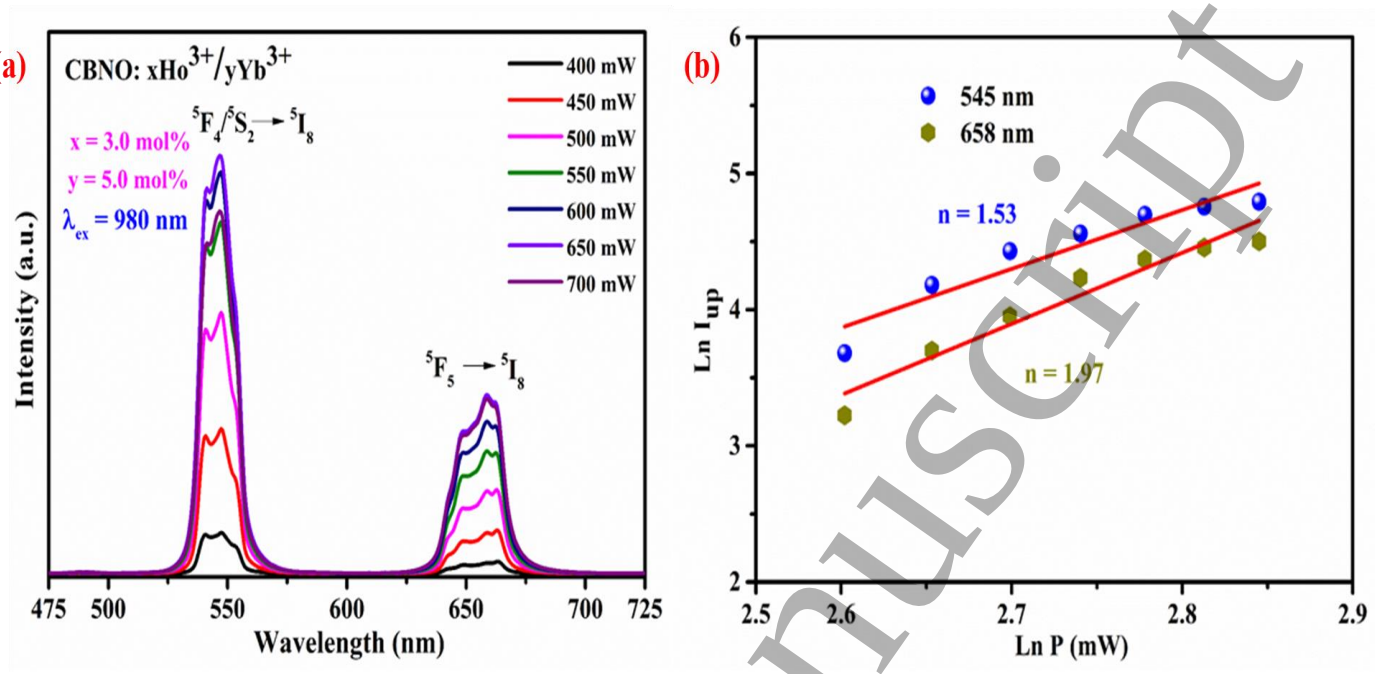
Accepted Manuscript

Fig. 8



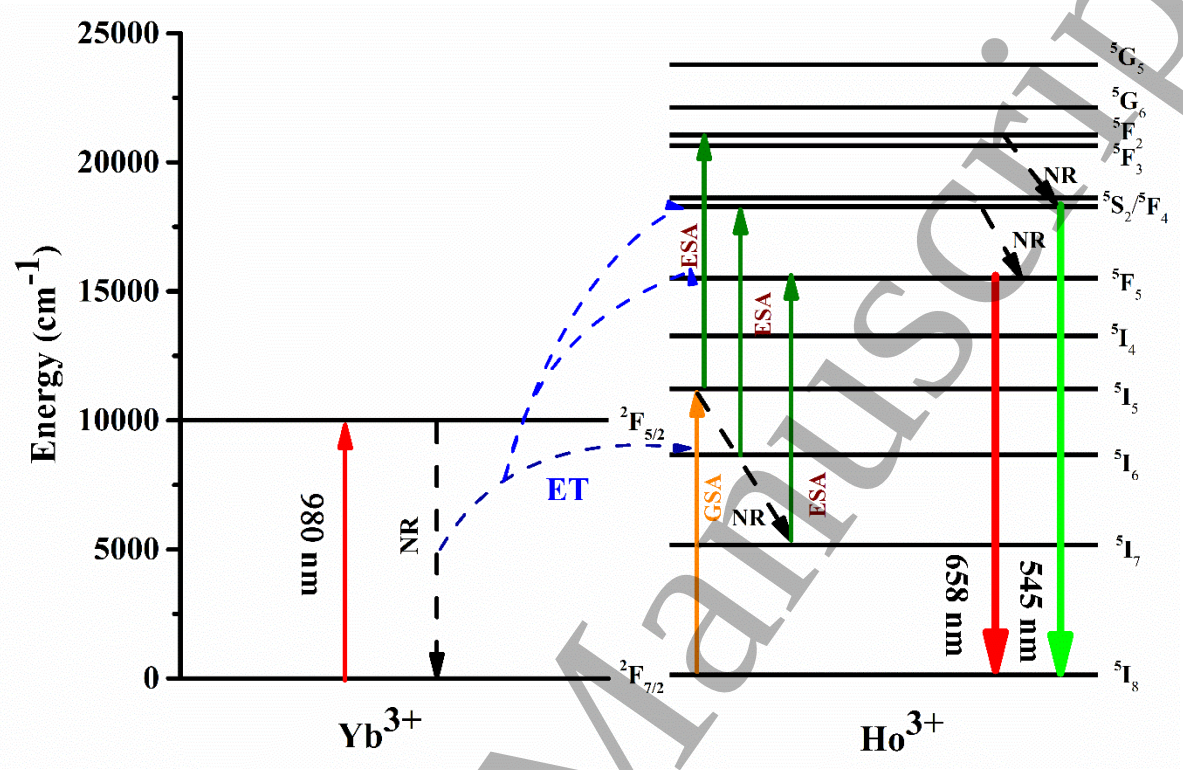
Accepted Manuscript

Fig. 9



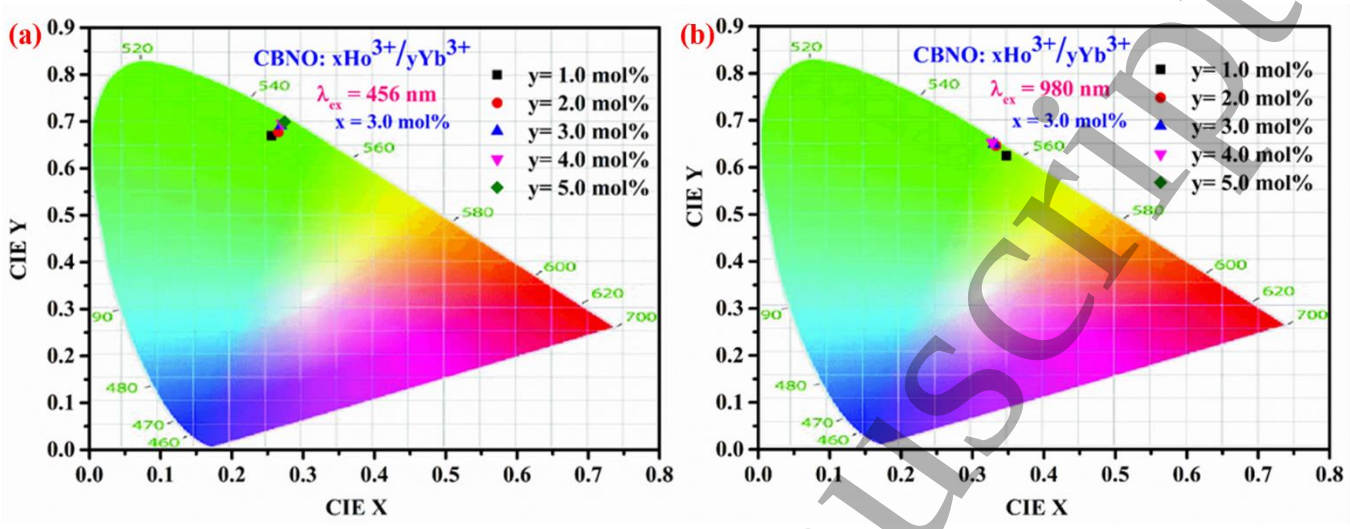
Accepted Manuscript

Fig.10



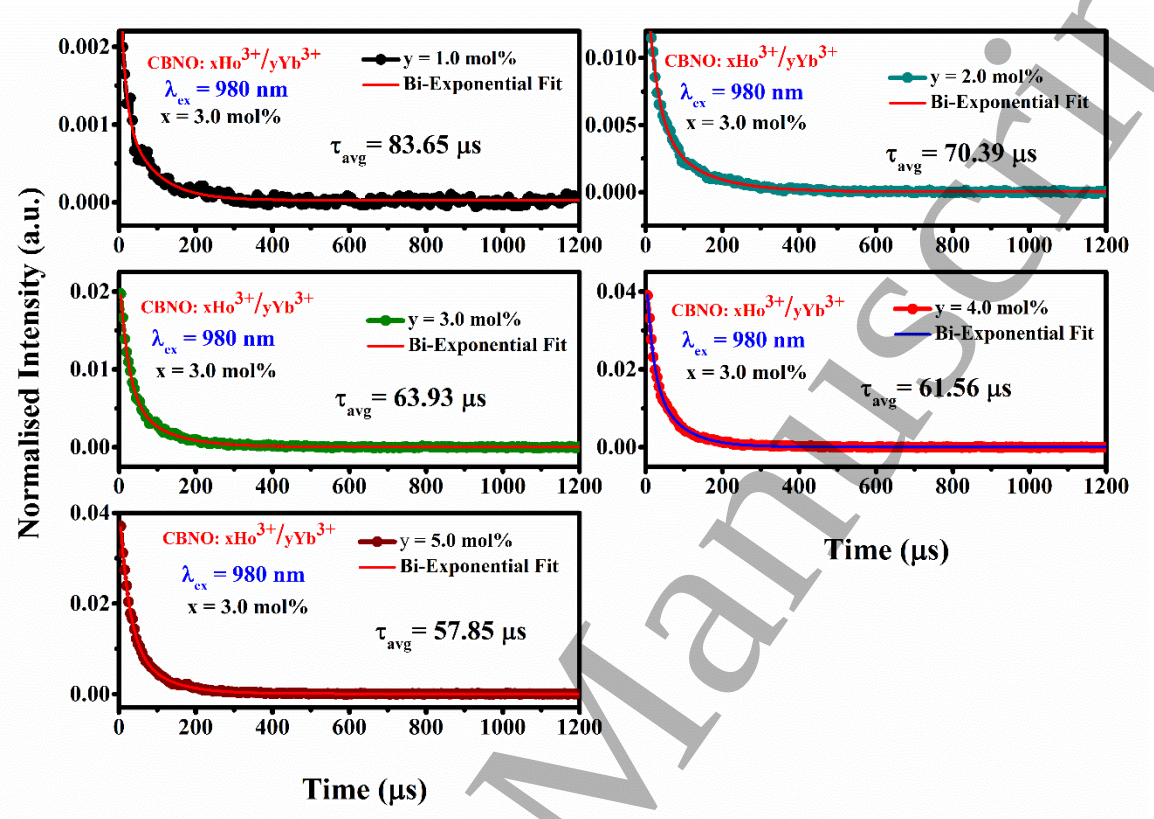
Accepted Manuscript

Fig. 11



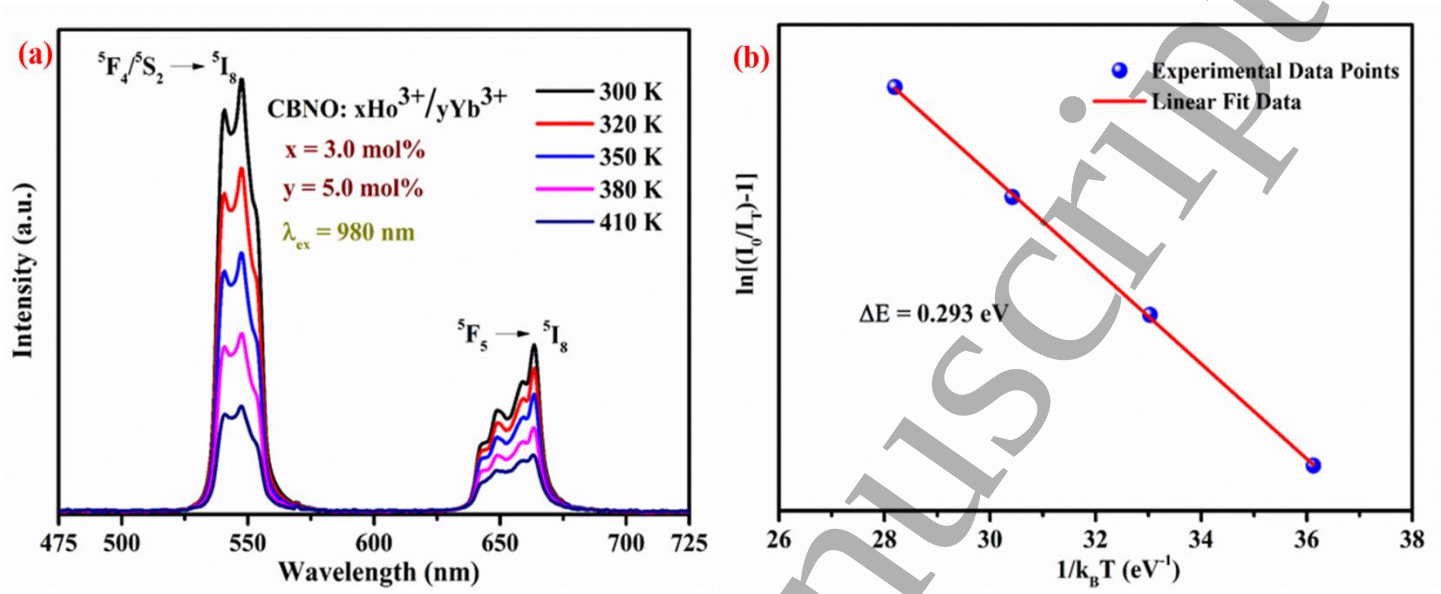
Accepted Manuscript

Fig. 12



Accepted Manuscript

Fig. 13



Accepted Manuscript

Table 1

CIE coordinates, CCT values of CBNO: xHo^{3+}/yYb^{3+} (x = 3.0 mol%, & y = 1.0, 2.0, 3.0, 4.0 & 5.0 mol%) phosphors under 980 nm and 456 nm excitation wavelengths.

Sample ID SYW	**CIE coordinates (x, y) 980 nm Excitation**	**CCT values (K) 980 nm Excitation**	**CIE coordinates (x, y) 454 nm Excitation**	**CCT values (K) 454 nm Excitation**	**Decay Lifetimes (µs)**
y = 1.0mol%	(0.3473, 0.6248)	5289.42	(0.2557,0.6699)	6696.86	83.65
y = 2.0mol%	(0.3329, 0.6462)	5510.43	(0.2653,0.6772)	6522.93	70.39
y = 3.0mol%	(0.3300, 0.6499)	5553.59	(0.2687,0.6910)	6440.94	63.93
y = 4.0mol%	(0.3276, 0.6528)	5589.20	(0.2707,0.6953)	6403.12	61.56
y = 5.0mol%	(0.3288, 0.6538)	5570.37	(0.2740,0.6992)	6345.53	57.85

Accepted Manuscript

Literature review on SRAM cells under Stability, Energy, and Power Constraints

Ram Murti Rawat
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
rammurtirawat@dtu.ac.in

Vinod Kumar
Department of Computer Science & Engineering
Delhi Technological University
Delhi, India
vinod_k@dtu.ac.in

2025 International Conference on Intelligent Control, Computing and Communications (IC3) | 979-8-3315-2749-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/IC363308.2025.10956626

Abstract- The rise of portable devices in our everyday lives has made power optimization a critical focus in contemporary VLSI technology. In recent years, static random-access memory (SRAM) has gained popularity due to its superior performance in VLSI design processes that operate at the submicron or nanoscale. Numerous SRAM cells have been developed to work effectively at lower voltages and with minimized latency. This paper introduces a newly proposed 8T SRAM cell design that offers enhanced read stability when compared to 6T SRAM cells. To tackle the issue of power leakage, two PMOS transistors are incorporated into both the read and write circuits. The proposed 8T SRAM cell demonstrates a marked reduction in leakage current relative to 6T SRAM cells. SRAM serves as the primary memory for small cache-less embedded processors, leading to the development of memories that utilize an SRAM cell optimized for various process parameters, such as energy, stability, and power limitations.

Keywords: low-power SRAM, stability, energy efficiency, power consumption, and static random-access memory (SRAM).

I. INTRODUCTION

The radix-2 system is prevalent in digital logic systems, employing binary representation that utilizes two symbols to denote logic levels. Systems are categorized as multi-valued logic systems when the radix exceeds 2, [1], [2], [3], [4], [5]. For instance, ternary logic employs three symbols, or logic states: zero, one, and two. In the context of the ternary system, $log_3^{(2n)}$ bits are necessary to represent a binary integer comprising n bits. In VLSI circuits, addressing pin-out challenges becomes particularly difficult when managing multi-valued logic. The static random-access memory (SRAM) cell operates as a latch, using back-to-back connected inverters for data storage. It includes a bit line (BL) and a select line, where the BL of the SRAM cell retains either read or write data when the select line is activated [6]. Introducing ternary inverters in place of the inverters within a binary SRAM cell allows for the design of a ternary static random-access memory (TSRAM) cell [7], [8], [9], [10], [11]. The SRAM cell designed for ternary logic aims to facilitate the creation of larger memory arrays by utilizing ternary logic circuits [12]. A traditional ternary inverter is used in a conventional ternary SRAM cell. The under-consideration design is poor in the storage of value and gives a considerable voltage drop that negatively influences the read/write processes within SRAM cells. To overcome this issue, it is crucial to create a ternary inverter cell capable of consistently providing the correct output voltage for each of the three states found in ternary logic [13], [14]. In modern scenarios, the emergence of Carbon Nanotube Field-Effect Transistors (CNTFETs) has introduced complications concerning complexity and stability in SRAM design. To overcome these problems, a new technology named Enhanced Gate Diffusion Input-based Ballistic Wrap Gate CNTFET (EGDI-BWGCNTFET) has been proposed. This method is utilized alongside ternary static random-access memory (T-SRAM) to improve stability, lower standby power usage, decrease voltage drop and effectively store read/write values within SRAM cells [15].

II. LITERATURE SURVEY

The literature survey has one of the main focuses on the literature review. It also provides details about current area techniques. The survey is divided into the following sections.

A. Definition of the Research Questions

In this step, we formulate the research questions (RQs) that will guide our work. We have formulated four Research Questions.

RQ1: What are various designs for stability analysis of low-power SRAM cells?
RQ2: What are the different low-power SRAM architectures?
RQ3: What are various designs for energy- and power-efficient SRAM cells?
RQ4: What are the various parameters for SRAM cell analysis?

B. Conduct Search

The objective of searching is to find the journal or conference paper that investigated the low-power SRAM cells. The search is normally based on some keywords. By choosing the suitable keywords, we can select a keyword. The keywords used to search for research articles are listed below: SRAM, stability analysis, static noise margin (SNM), low power low swing, leakage power, Energy efficiency, and Power consumption.

C. Screening of Papers

To ensure that all the papers considered for review were directly relevant to the research questions, Specific inclusion

979-8-3315-2749-5/25/$31.00 ©2025 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:43:07 UTC from IEEE Xplore. Restrictions apply.

and exclusion criteria were set. Table 1 describes the inclusion and exclusion criteria used in this study.

TABLE. I. INCLUSION AND EXCLUSION CRITERIA

Criteria for Inclusion	Criteria for Exclusion
Studies that focus on improved stability of SRAM cells	Studies not in English
Studies focusing on various approaches for energy-efficient, leakage power, and stability analysis	Studies not related to any of the research questions.
Studies that conduct high-performance SRAM cell analysis of their research	Studies whose findings are unclear and ambiguous

D. Summary of Research Review

Based on the literature review, a summary of all research articles is given below in Table 2. It includes the name of the technique, technology, summary, and observation.

TABLE. II. SUMMARY OF RESEARCH REVIEW

Title and Authors	Publisher and Years	Technology	Summary and observation
8T-SRAM Cell with Improved Read and Write Margins in 65 nm CMOS Technology F. Moradi et al. [16].	IFIP International Federation for Information Processing, 2015	65 nm CMOS TSMC technology	The study suggests that an 8T-SRAM cell improves gate leakage power consumption and write margin compared to a 6T-SRAM cell, but increases subthreshold leakage, highlighting the need for further research.
A Review of Circuit-Level Leakage Minimization Techniques in CMOS VLSI Circuits R. Lorenzo & S. Chaudhury [17].	IETE Technical Review, 2016	NA	This study focuses on leakage power dissipation, a significant issue in high-performance chips and VLSI circuits. It categorizes various techniques and evaluates their reliability under process, voltage, and temperature changes. The analysis and experimental findings guide selecting the most effective method for reducing leakage power. Limitations include potential increased circuit delay, data retention issues, and the need for additional transistors.
Cell Stability Analysis of Conventional 6T Dynamic 8T SRAM Cells in 45NM Technology K. Dhanumjaya et al. [18].	International Journal of VLSI Design & Communication Systems (VLSICS), April 2012	45 nm Technology	This study compares a dynamic column-based power supply 8T SRAM cell to a regular 6T SRAM. N-curve measurements were utilized to determine read stability and write capabilities. According to simulation results, the suggested 8T SRAM cell provides greater read stability, higher read current, and lower leakage current in 45 nm technology than the typical 6T SRAM.
Pre-Charged Local Bit-Line Sharing SRAM Architecture for Near-Threshold Operation T. W. Oh, H. Jeong, J. Park, and S. Jung [19].	IEEE Transactions on Circuits and Systems, October 2017, 2017	22-nm FinFET technology	This paper describes a static random-access memory (SRAM) with pre-charged local bit-line sharing (PCLBS) that functions near the threshold. Prior SRAMs, such as average-8T and full-swing local bit-line (FSLB), used a low-capacitance local bit-line pair to improve read stability. However, the average-8T SRAM has considerable delays because of incomplete local bit-line development.
A low-power SRAM with charge cycling-based read and write assist scheme H. Zhang et al. [20].	2020, IEEE	14nm FinFET	This work describes a bit-line charge cycling-based read-and-write assist circuit for static random-access memory that leverages energy that would otherwise be wasted on bit lines to precharge them in the following cycle. The NBL Circuit is used as a writing help to reduce latency and increase writing skills.
Low power self-controllable voltage level and low swing logic-based 11T SRAM cell for high-speed CMOS circuits K. Gavaskar, U. S. Ragupathy [21].	Analog Integrated Circuits and Signal Processing (2019)	30 nm CMOS technology.	The article discusses the development of a low-power 11T SRAM cell using two voltage sources connected to the bit line and bit bar line. This method effectively removes swing voltage at output nodes, reducing dynamic power loss in SRAM cells. The self-controllable voltage level technique uses a PMOS transistor as a switch and NMOS transistors as series resistors, reducing leakage current and static power dissipation. However, the use of additional transistors may increase the overall area.
A Novel Low Power and Swing Restoration SRAM Logic Circuit Technique, R. M. Rawat [22].	International Journal of Engineering Research & Technology (IJERT), 2018	180 nm Technology	The study aims to enhance read and write operations for 8T SRAM cells, which outperform 6T cells using Swing Restoration Dual Node Voltage, reducing read and write time and improving stability, while utilizing PMOS transistors with narrower widths.

E. Research Gaps

1. During the Read and Write operation in the 6T SRAM cell, poor read stability and write ability are provided due to voltage drops occurring at the node voltages [19], [20], [21], [22].

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:43:07 UTC from IEEE Xplore. Restrictions apply.

2. The standard 6T SRAM cell has poor write performance owing to disturbances in the node voltage at the downscaled cell supply voltage [17].
3. The severe (read disturb) problems that SRAM memory has been facing are high leakage power consumption and instability due to continuous downscaling of the process technology [17].
4. The memory array is responsible for the consumption of a major amount of power of the total power consumed by the chip [20].

F. Research Objectives

The following are the expected objectives:

1. To design SRAM cells for low power and high speed.
2. To develop an improved stability of low-power SRAM cells.
3. To develop a low-leakage SRAM cell.

III. PROPOSED METHODOLOGY

Designing an 8T (8-transistor) static random-access memory (SRAM) cell involves creating a stable memory element using transistors. The 8T SRAM cell is a common design choice for on-chip memory due to its low leakage and good read and write characteristics.

A. Problem statement

In the existing work, the power consumption of the 6T SRAM and 5T CNTFET SRAM cells for all write and read operations is very high [23], [24], [25], [26], [27], 28].

B. Rationale for Undertaking This Research Work

To overcome these challenges, it is essential to present solutions that can successfully overcome the issues at hand.

The current SRAM architecture based on the 5T design exhibits significant drawbacks, including elevated leakage power, inefficient power utilization, and a lack of precise stability [23], [24], [25], [26], [27], 28]. These shortcomings serve as the driving force behind our research endeavors.

C. Our Research Contribution

The primary contributions of this research endeavor are the following: In this research work, an 8T SRAM design utilizing electrostatic doped-gate all-around CNTFETs is proposed for enhancing read/write assist performances with stability. The 8T SRAM Cell design uses an Electrostatic Doped-Gate All Around CNTFET (EDGA-CNTFET) to properly retain data in read/write signals.

In the pull-down network, two electrostatic doped-gate all-around CNTFETs (N5 and N6) are interconnected in a Darlington configuration. The memory cell that has been proposed is made up of EDGA CNTFETs (N1, N2, P1, P2, N5, and N6). Additionally, access transistors N3 and N4, which are also EDGA CNTFETs, are controlled by word lines. This arrangement establishes the storage nodes Q and QB. Routes for writing and reading are provided by the bit lines BL and BLB. The incorporation of Darlington electrostatically doped gate-all-around CNTFET in the SRAM cell's pull-down network enhances the memory cell's power consumption performance during both write and read operations [29]. Subsequently, the proposed 8T SRAM cell design, utilizing electrostatically doped gate-all-around CNTFET (EDGA-CNTFET), is simulated with the HSPICE simulation tool. The performance of the EDGA-CNTFET based on the 8T SRAM cell design is assessed in terms of read SNM, write SNM, read delay, write delay, average write power consumption, and average read. Evaluations are conducted on the power consumption, as well as write and read access times. Then the proposed electrostatic doped-gate all-around CNTFETs (EDGA-CNTFET)-based 8T SRAM cell design is compared with existing method-like designs of differential TG-based 8T SRAM cells for ultralow-power applications [23], reliable 8T SRAM for high-speed searching, and logic-in-memory operations [24], respectively.

D. 6T SRAM Cell

During a read operation of a 6T SRAM cell, the bit line access transistor directly accesses the data storage node, as illustrated in Figure 1. As illustrated in Figure 1, during the reading process, the voltage at the storage node changes due to the interaction of the cross-coupled inverter pair and the bit line. BL and BLB represent bit lines, while WL represents word lines. WL controls the access transistor and facilitates both read and write operations. Bitline acts as a node for both input and output. Data from the SRAM cell is transferred to the sense amplifier during the read operation through the bit line. In this specific technology, the minimum transistor length is 130 nm [30].

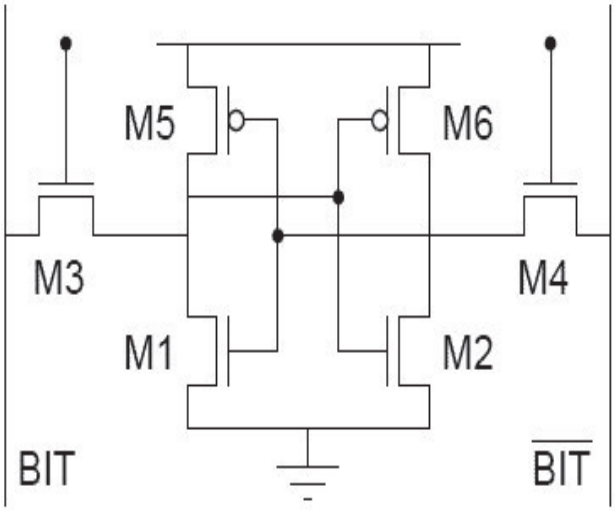


Fig. 1. 6T SRAM Cell [30].

E. Proposed 8T SRAM Cell

The proposed SRAM cell employs dual voltage and swing restoration Inverter (SRI) in order to enhance the node voltage performance during retention, read, and write operations. Delay and stability are also shown in Figure 2. Low-power 6T SRAM and proposed 8T SRAM cell design comparison is made, and it shows that 8T SRAM cell has a better read, write, and retention performance than that of the 6T SRAM cell. These reasons result in an increase in Static Noise Margins (SNM), which improves the read and write capability of the bit cell. The result of this is an increased SNM, which ensures a strong read and write ability of the bit cell.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:43:07 UTC from IEEE Xplore. Restrictions apply.

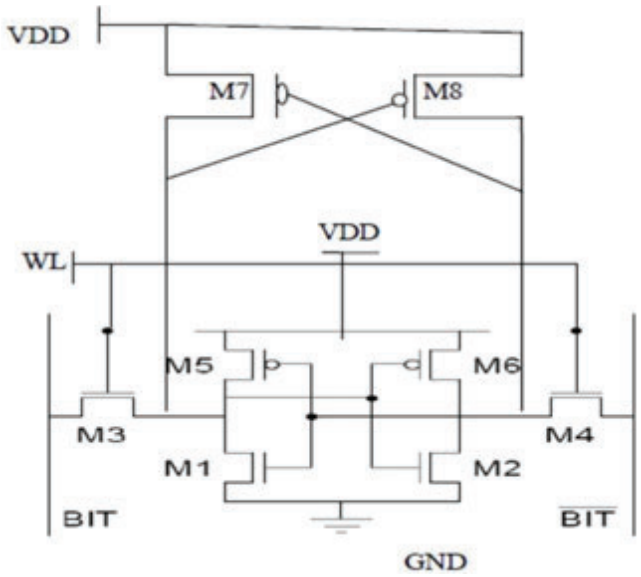


Fig. 2. Proposed 8T SRAM Cell.

IV. DISCUSSIONS

This section discusses the proposed 8T SRAM cell's write ability, read stability, and hold static noise margin.

A. Hold Stability

The most commonly used metric to determine retention stability is the static signal-to-noise ratio of an SRAM cell. Retention stability is usually evaluated when the cell is in the retention state. In this state, the word line is disabled such that the cell is disconnected from the bit line. SNM is defined here as the maximum noise passed to the storage node such that the cell contents don't collapse.

B. Read Stability

The RSNM measures the read stability in the SRAM cell. Comparing the conventional 6T SRAM cell, the isolation of storage nodes in the proposed 8T SRAM cell improves the RSNM.

C. Write Stability

Write stability is examined through the WSNM. The isolation of the nodes that store in the proposed 8T SRAM cell leads to a larger WSNM than does the existing 6T SRAM cell. The write trip point assesses the write capabilities of the SRAM cell, indicating how difficult it is to alter storage nodes. It is determined by gradually increasing the bit-line voltage from 0 to Vdd and recording the instant that the cell flips, causing the Q and Q bar to alter their recorded values. The value of the bit-line voltage at the intersection of internal storage nodes Q and QB bar represents the write trip point.

V. CONCLUSIONS

This article focuses on the existing design approaches for creating SRAM cells. The integrated circuits, find their application in high-speed and low-power SRAM cells. This paper presents a novel design for a proposed 8T SRAM cell, incorporating both differential bit lines and word lines. This innovative configuration effectively reduces leakage current in comparison to other designs, making it especially well-suited for applications where minimizing power consumption due to leakage is a critical consideration. Moreover, the proposed 8T SRAM cell outperforms the 6T SRAM cell. Furthermore, the proposed 8T cell has better read and write stability than the 6T SRAM cell. Overall, the 8T SRAM cell, highlights its suitability for high-performance, power-efficient, and reliable circuit applications. The various approaches utilized in different SRAM cells were investigated, with a focus on low-power SRAM cell design.

REFERENCES

[1] A. Amirany, M. H. Moaiyeri, and K. Jafari, “Process-in-Memory using a Magnetic-Tunnel-Junction synapse and a neuron based on a carbon nanotube Field-Effect transistor,” IEEE Magnetics Letters, vol. 10, pp. 1–5, Jan. 2019, doi: 10.1109/lmag.2019.2958813.

[2] S. Bala and M. Khosla, “Design and performance analysis of low-power SRAM based on electrostatically doped tunnel CNTFETs,” Journal of Computational Electronics, vol. 18, no. 3, pp. 856–863, May 2019, doi: 10.1007/s10825-019-01345-z.

[3] A. Darabi, M. R. Salehi, and E. Abiri, “Newly energy-efficient SRAM bit-cell using GAA CNT-GDI method with asymmetrical write and built-in read-assist schemes for QR code-based multimedia applications,” Microelectronics Journal, vol. 114, p. 105117, Jun. 2021, doi: 10.1016/j.mejo.2021.105117.

[4] T. Khurshid, S. Fatima, F. A. Khanday, F. Bashir, F. Zahoor, and F. A. Hussin, “Carbon nanotube field effect transistor (CNTFET) operational transconductance amplifier (OTA) based design of high-frequency memristor emulator,” International Journal of Numerical Modelling Electronic Networks Devices and Fields, vol. 34, no. 2, Oct. 2020, doi: 10.1002/jnm.2827.

[5] H. Kumar, S. Srivastava, and B. Singh, “Low power, high-performance reversible logic enabled CNTFET SRAM cell with improved stability,” Materials Today Proceedings, vol. 42, pp. 1617–1623, Jul. 2020, doi: 10.1016/j.matpr.2020.06.475.

[6] M. Kumar and J. S. Ubhi, “Design and analysis of CNTFET based 10T SRAM for high performance at the nanoscale,” International Journal of Circuit Theory and Applications, vol. 47, no. 11, pp. 1775–1785, Sep. 2019, doi: 10.1002/cta.2696.

[7] T. S. Kumar and S. L. Tripathi, “Leakage Reduction in 18 nm FinFET based 7T SRAM Cell using Self Controllable Voltage Level Technique,” Wireless Personal Communications, vol. 116, no. 3, pp. 1837–1847, Sep. 2020, doi: 10.1007/s11277-020-07765-6.

[8] D. Nayak, P. K. Rout, S. Sahu, D. P. Acharya, U. Nanda, and D. Tripthy, “A novel indirect read technique-based SRAM with ability to charge recycle and differential read for low power consumption, high stability, and performance,” Microelectronics Journal, vol. 97, p. 104723, Feb. 2020, doi: 10.1016/j.mejo.2020.104723.

[9] P. K. Patel, M. Malik, and T. K. Gupta, “Design of an ultralow power CNTFET based 9T SRAM with shared BL and half select free techniques,” International Journal of Numerical Modelling Electronic Networks Devices and Fields, vol. 32, no. 2, Sep. 2018, doi: 10.1002/jnm.2511.

[10] P. K. Patel, M. M. Malik, and T. K. Gupta, “Performance evaluation of single-ended disturb-free CNTFET-based multi-Vt SRAM,” Microelectronics Journal, vol. 90, pp. 19–28, May 2019, doi: 10.1016/j.mejo.2019.05.015.

[11] R. M. Ponnayan and S. V. Thazhathu, “One instruction set computer with optimised polarity-tunable model of double gate CNTFETs,” IET Circuits Devices & Systems, vol. 14, no. 6, pp. 770–779, Mar. 2020, doi: 10.1049/iet-cds.2019.0143.

[12] A. Rahimi et al., “Hyperdimensional computing nanosystem: in-memory computing using monolithic 3D integration of RRAM and CNFET,” in Elsevier eBooks, 2020, pp. 195–219. doi: 10.1016/b978-0-08-102782-0.00008-3.

[13] P. Rajesh and F. H. Shajin, “Optimal allocation of EV charging spots and capacitors in distribution network improving voltage and power loss by Quantum-Behaved and Gaussian Mutational Dragonfly Algorithm (QGDA),” Electric Power Systems Research, vol. 194, p. 107049, Jan. 2021, doi: 10.1016/j.epsr.2021.107049.

[14] P. Rajesh, F. H. Shajin, and B. N. Kommula, “An efficient integration and control approach to increase the conversion efficiency of high-

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:43:07 UTC from IEEE Xplore. Restrictions apply.

current low-voltage DC/DC converter," Energy Systems, vol. 13, no. 4, pp. 939–958, May 2021, doi: 10.1007/s12667-021-00452-w.

[15] A. Sachdeva and V. K. Tomar, "Design of 10T SRAM cell with improved read performance and expanded write margin," IET Circuits Devices & Systems, vol. 15, no. 1, pp. 42–64, Dec. 2020, doi: 10.1049/cds2.12006.

[16] F. Moradi and J. K. Madsen, "Improved read and write margins using a novel 8T-SRAM cell," 2014 22nd International Conference on Very Large-Scale Integration (VLSI-SoC), 2014, pp. 1-5, doi: 10.1109/VLSI-SoC.2014.7004186.

[17] R. Lorenzo & S. Chaudhury (2016): Review of Circuit Level Leakage Minimization Techniques in CMOS VLSI Circuits, IETE Technical Review, DOI: 10.1080/02564602.2016.1162116

[18] K. Dhanumjaya & M. Sudha, M & M. Prasad & K. Padmaraju (2012). Cell Stability Analysis of Conventional 6T Dynamic 8T SRAM Cell in 45NM Technology. International Journal of VLSI Design & Communication Systems.

[19] T. W. Oh, H. Jeong, J. Park, and S. Jung, "Pre-Charged Local Bit-Line Sharing SRAM Architecture for Near-Threshold Operation," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 64, no. 10, pp. 2737-2747, Oct. 2017, doi: 10.1109/TCSI.2017.2702587.

[20] H. Zhang, S. Jia, J. Yang, and Y. Wang, "A low-power SRAM with charge cycling-based read and write assist scheme," 2020 IEEE 15th International Conference on Solid-State & Integrated Circuit Technology (ICSICT), 2020, pp. 1-3, doi: 10.1109/ICSICT49897.2020.9278282.

[21] K. Gavaskar & U. Ragupathy (2019). Low power self-controllable voltage level and low swing logic-based 11T SRAM cell for high-speed CMOS circuits. Analog Integrated Circuits and Signal Processing. 100. 10.1007/s10470-018-1277-3.

[22] R. M. Rawat (2018). A Novel Low Power and Swing Restoration SRAM Logic Circuit Technique, March 2018, International Journal of Engineering and Technical Research V7(02).

[23] Roy, C., and Islam, A., 2020. Design of differential TG-based 8T SRAM cells for ultralow-power applications. Microsystem Technologies, 26, pp. 3299-3310.

[24] Chen, J., Zhao, W., Wang, Y., Shu, Y., Jiang, W., and Ha, Y., 2022. A reliable 8T SRAM for high-speed searching and logic-in-memory operations. IEEE Transactions on Very Large-Scale Integration (VLSI) Systems, 30(6), pp. 769–780.

[25] Aura, S.R., Huq, S.M., and Biswas, S.N., 2022. Design of High-Speed Dual Port 8T SRAM Cell with Simultaneous and Parallel READ-WRITE Feature. International journal of electrical and computer engineering systems, 13(9), pp. 823-829.

[26] Kim, Y., Li, S., Yadav, N., and Choi, K.K., 2021. A novel ultra-low-power 8t SRAM-based compute-in-memory design for binary neural networks. Electronics, 10(17), p. 2181.

[27] Yokoyama, Y., Nii, K., Ishii, Y., Tanaka, S., and Kobayashi, K., 2022. Disturbance-aware Dynamic Power Reduction in Synchronous 2RW Dual-Port 8T SRAM by Self-Adjusting Wordline Pulse Timing. IEEE Journal of Solid-State Circuits.

[28] Tripathy, M.R., and Jit, S., 2021. Lateral and vertical gate oxide stacking impact on noise margins and delays for the 8T SRAM designed with source pocket engineered GaSb/Si heterojunction vertical TFET: a reliability study. IEEE Transactions on Device and Materials Reliability, 21(3), pp. 372-378.

[29] Seevinck E, List FJ, Lohstroh J. Static-noise margin analysis of MOS SRAM cells. IEEE J Solid-State Circuits 1987;22(5):748–54.

[30] Chang M-F, Chang S-W, Chou P-W, Wu W-C. A 130 mV SRAM with expanded write and read margins for subthreshold applications, IEEE J Solid-State Circuits 2011;46(2):520–9.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:43:07 UTC from IEEE Xplore. Restrictions apply.

Social Network Analysis and Mining (2025) 15:31
https://doi.org/10.1007/s13278-025-01447-6

CASE REPORT

Longitudinal analysis of Covid-19 vaccine-related tweets in india: linking sentiment fluctuations with topic modeling

Seba Susan[1] · Anmol Bansal[1] · Arjun Choudhry[1] · Anubhav Sharma[1]

Received: 22 July 2024 / Revised: 20 February 2025 / Accepted: 21 February 2025
© The Author(s) 2025

Abstract
Twitter (now known as "X") is a popular medium for Covid-19 related discussions. This paper presents a novel case study on sentiment analysis and topic modeling of Covid-19 vaccine-related tweets of users geo-located in India in the duration of 12 December 2020 to 11 November 2021 in the course of which more than half of the country's 1.3 billion population got vaccinated. The sentiment analysis was performed, on day-wise basis, using unsupervised lexicon-driven sentiment analysis tools AFINN and Valence Aware Dictionary and sEntiment Reasoner, as well as BERTweet and Covid-Twitter-BERT transformer models pre-trained on Covid-19 tweets. The models were comparatively evaluated on a smaller annotated dataset of Covid-19 vaccine-related tweets prior to the longitudinal analysis. AFINN was ultimately chosen due to its better performance and ease of use for large unannotated data. AFINN analysis revealed that 51.38% of the 44,130 tweets were neutral, while 38.84% were positive, and 9.78% negative. Latent Dirichlet Allocation was used for topic modeling at the peak points corresponding to large sentiment fluctuations in the positive and negative longitudinal graphs derived using AFINN. The positive and negative vocabularies at peak points were scrutinized to derive insights on national/international events that triggered a change in public opinion. These findings could guide policy makers in gathering intelligence on misinformation and associated sentiments, and planning counter-measures to combat anti-vaccine campaigns. This study informs future strategies to counter vaccine hesitancy through targeted communication aimed at vulnerable groups that experience high anxiety and psychosocial burden during the pandemic.

Keywords Twitter · Covid-19 vaccine · Sentiment analysis · Topic modeling · Longitudinal studies

1 Introduction

Twitter (now known as "X") is an easily accessible social media platform where people constantly express their opinions about national and international events. The Covid-19 pandemic and the ensuing lockdown imposed in several countries caused a significant rise in the number of people turning to social media platforms like Twitter to express their feelings, doubts and apprehensions. The vaccination program underway in several countries is considered universally as one of the most effective ways to counter the spread of the virus (Chakraborty et al. 2023). Several vaccines got worldwide recognition, such as AstraZeneca, Moderna, Pfizer, and Sputnik, while domestic vaccines developed by low- and middle-income countries got approval after a long procedural delay. Naturally, this sparked a widespread discussion on the shortage and efficacy of imported and domestic vaccine brands among the Twitter users in developing countries like India. Twitter, now popularly known as "X", became the most popular social media platform in India for obtaining real-time updates during the Covid-19 pandemic due to its outreach among students, middle-class working group, high-class income group, corporate world, political leaders and celebrities, though constrained by ethnic and demographic biases, and use of English as the primary medium of communication (Jain et al. 2021). The online discussions ranged from hope and willingness to participate in the vaccination program to despair as the deadly Delta variant caused much fatalities globally. Vaccine sentiment analysis and topic modeling from geo-tagged social media posts is now considered significant research (Hussain et al. 2021; Dubey 2021) since it helps to identify the (1)

✉ Seba Susan
seba_406@yahoo.in

[1] Department of Information Technology, Delhi Technological University, Shahbad Daulatpur, Bawana Road, New Delhi, Delhi 110042, India

Published online: 01 April 2025

dynamically evolving public opinion regarding vaccines (2) the specific counties or countries having a more skeptical attitude towards certain vaccines (3) local and international news or events that triggered a change in the public sentiment towards vaccination and the vaccine brands. It is crucial for the government and policy makers to understand the public mood from social media platforms so that timely interventional steps can be introduced for vaccine procurement, vaccine approval, promoting vaccination and reducing vaccine hesitancy through educational campaigns.

Sentiment analysis techniques are broadly classified into unsupervised and supervised approaches. Some popular unsupervised tools are Valence Aware Dictionary and sEntiment Reasoner (VADER) (Hutto and Gilbert 2014), TextBlob (Loria et al. 2014) and AFINN (Nielsen 2011) that assign a sentiment score to each tweet (the polarity of the score indicates the sentiment−positive or negative). Supervised methods perform text classification using bag-of-words feature representations (Susan and Keshari 2019) and sequential word embeddings (Nguyen et al. 2020; Müller et al. 2023), that are learned using machine learning classifiers and transformers, respectively. A common issue with the Twitter data related to the Covid-19 pandemic is the lack of sentiment annotations by human analysts, and the smaller size of datasets with skewed or imbalanced class profiles, that makes supervised learning a challenging task. The use of domain-specific pre-trained transformer models such as BERTweet (Nguyen et al. 2020) and Covid-Twitter-BERT (CT-BERT) (Müller et al. 2023) that are pre-trained on Covid-19 tweets, mitigates the issue to some extent.

1.1 Motivation for the study

Vaccinations started in India, the most populous country in the world, on 16 January 2021. India with its population of 1.3 billion was the second country most affected by the pandemic (Mirza et al. 2024). It also became, on 21st October 2021, the second country to cross the 1 billion mark of the number of doses of vaccines administered. 50% of the country's population had got vaccinated by November 2021. The second Covid-19 wave, that was attributed to the Delta (B.1.617.2) variant, wreaked havoc in India in the months of May–June 2021 with an increase in the numbers of hospitalizations and fatalities while the vaccination program was still underway in the country. The second wave of Covid-19 had peaked on 7 May 2021 with 4.14 lakh daily cases, as observed from the graph in Fig. 1.

It would be therefore a constructive case study to monitor the public receptiveness towards the vaccination program by analyzing the trending topics in social media on the days when sentiments (positive or negative) of the Twitter users peaked. The stakeholders of such a sentiment-cum-topic analysis study are government officials, policy makers and vaccine manufacturers who could devise timely measures to counter vaccine hesitancy based on shifting opinions on social media.

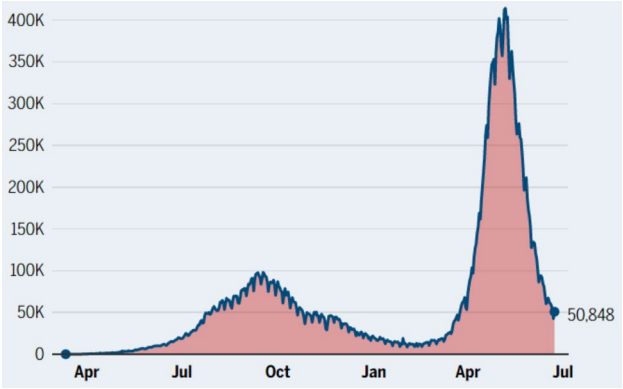


Fig. 1 The peak (right-side) in the number of daily confirmed cases during the second Covid-19 wave in India in the months of April to July 2021 [data provided by Ministry of health and family welfare, Government of India]

2 Research questions

Specifically, this study attempts to answer the following research questions:

RQ1: How to effectively use unsupervised sentiment analysis tools for predicting day-wise sentiments of Covid-19 vaccine-related tweets of users geo-located in India?

RQ2: How to effectively extract the topics of discussion on the days when positive and negative sentiments peaked on Twitter?

RQ3: How to correlate the extracted topics and sentiments with events of national and international importance that may have been the cause and effect of the outburst of opinions on social media platforms?

To answer the research questions, this study presents a novel approach for longitudinal sentiment analysis and dynamic topic modeling of geo-tagged tweets of users from India in the duration from 12 December 2020 to 11 November 2021 while the vaccination program was rolled out and the deadly second Covid-19 wave progressed through the country. While AFINN was used for sentiment analysis, Latent Dirichlet Allocation (LDA) was used for dynamic topic modeling at positive and negative sentiment peaks in the longitudinal graphs. To the best of our knowledge, there is no other work that has dynamically analyzed the sentiments of Twitter users in India before and after vaccinations started in India on 16 January 2021, for an almost one-year duration encompassing the rise and abatement of

the devastating second wave, along with topic analysis of the peak points in the longitudinal sentiment graph.

2.1 Main contributions

The main contributions of the paper are summarized as follows:

1. We performed longitudinal sentiment analysis (over time) using AFINN to identify positive and negative sentiment fluctuations in the duration of 12 December 2020 to 11 November 2021 for Twitter users in India. The time window was selected to coincide with the rollout of the vaccination program in India, and the complete duration of the second wave of the Coronavirus in India, until the point when the country had recovered from its aftereffectsby November 2021.

2. The number of positive, negative and neutral tweets were summed up separately for each day, creating three longitudinal graphs, one for each sentiment. The major peaks in the positive and negative sentiments in the longitudinal graphs indicate a significant fluctuation in public sentiment.

3. We used LDA for discovering the trending topics under discussion on the specific dates corresponding to peaks in the sentiment graphs. The top-20 salient terms were identified based on their frequencies of occurrences in the top-20 topics detected through LDA at the peak points in the positive and negative sentiment graphs. The positive and negative vocabularies at the sentiment peaks were objectively analyzed to understand the hopes, fears and doubts of the public regarding vaccine efficacy and availability.

4. We manually collated news articles at sentiment peaks in the longitudinal graphs in order to identify the national/international events that may have caused the sentiment fluctuations and outburst of opinions. The impact of the public discussions on social media on society and governing authorities were qualitatively analyzed.

The rest of the paper has the following organization structure. Related research is presented in Sect. 3, the proposed methodology is detailed in Sect. 4, the experimental setup and the results are discussed and analyzed in Sect. 5, and the final conclusions are summarized in Sect. 6.

3 Related work

Sentiment analysis is a reliable means for understanding the dynamically evolving sentiments of the public who take to social media for exulting over sports events (Vashishtha and Susan 2019), discussing positive or negative reviews of some place, object or event (Ghosh et al. 2021), or expressing apprehensions about natural calamities (Mendon et al. 2021). Mining public sentiments from social media is an essential part of disaster management as argued by the authors in Ragini et al. (2018). The authors extracted textual features such as bag-of-words and part-of-speech tagging, and compared different classifier models for determining the public opinion to philanthropic aids received in response to a disaster. A disaster-specific word embedding trained on the Nepal and Italy earthquakes was adopted by the authors in Behl et al. (2021) to determine public sentiments from Covid-19 related tweets. Some recent works have focused on the Covid-19 vaccination drive in different countries and its impact on social media. Most of these case studies focus on finding the overall sentiment (positive, negative, neutral) in a particular time frame. Dubey (Dubey 2021) analyzed the public sentiments and emotions towards Covishield and Covaxin, the two vaccines popularly administered in India, and proved that there was more positivity associated with the Covishield vaccine. The author used the National Research Council (NRC) Emoticon lexicon to detect eight emotions and sentiments (positive or negative) in tweets posted over five days, specifically, in between 14th January 2021 to 18th January 2021, coinciding with the launch of the vaccination program in India. In a related work (Batra et al. 2021), Batra et al. linked emotions such as joy, surprise, etc. expressed in tweets related to Covid-19 with ongoing events in six countries. Long-short term memory (LSTM) and its variants were used for the prediction. In a related work (Hu et al. 2021), US-based tweets were evaluated using VADER for detecting spatio-temporal patterns in sentiments, and the emotions of tweets were analyzed using National Research Council Canada Lexicon (NRCLex). The authors in Jayasurya et al. (2021) analyzed temporal graphs of sentiments predicted using 14 machine learning models. They used a dataset which had been annotated using VADER and TextBlob, two of the most popular unsupervised sentiment analysis systems. The outcome was a set of dates on which highest numbers of positive, negative and neutral tweets were recorded. Umair et al. (Umair et al. 2023) performed sentiment analysis using bidirectional encoder representations of text along with Naïve Bayes and Support Vector Machine (SVM) classifiers to detect sentiments of tweets posted by users on Twitter. The authors highlighted the utility of Geographic Information System (GIS) to gain the sentiments of people in a specific region.

Previous works in literature (Bansal et al. 2023, 2022) have explored the use of pre-trained transformer models for classifying tweets related to Covid-19 vaccines. However, such applications have limited scope due to insufficient amount of annotated data available for training. Supervised learning models, on the whole, require lots of training data, with annotated sentiment labels, which is found lacking in the case of Covid-19 vaccine-related posts on social media. Text oversampling (Choudhry et al. 2022) does not improve the results of supervised learning for imbalanced datasets, as proved in Bansal et al. (2023). Unsupervised lexicon-based

sentiment analysis tools such as VADER, AFINN and TextBlob are popularly used in the absence of sufficient ground truth annotations. Their popularity is due to (i) the lack of annotated Twitter datasets pertaining to Covid-19 related discussions (ii) the dynamic and evolving nature of the sentiments derived from sequential tweets (Vashishtha and Susan 2019) that renders supervised machine learning a challenging task (Behl et al. 2021). Public optimism in the US and UK towards the vaccination program was evaluated from 300,000 Facebook and Twitter posts in Nguyen et al. (2020) using a weighted ensemble of VADER, TextBlob and BERT. It was noted that while VADER and TextBlob predicted well for the positive sentiments, BERT performed better for the negative sentiments. 10% of the tweets were labeled manually by the team for facilitating supervised learning and computation of precision scores. Niu et al. (Niu et al. 2022) classified the sentiments of tweets pertaining to Japanese users before and after the vaccination started in Japan. The sentiments were classified using Amazon Web Services (AWS) and were found to be more negative in nature due to apprehensions and confusion regarding the vaccination process.

Topic modeling of Covid-19 related tweets would help to understand the hotspots of discussion in a particular time frame or on a specific date. LDA is the most popular unsupervised tool used for topic modeling of Covid-19 tweets, primarily due to its ease of use and effective performance (Lin 2023; Lyu et al. 2021; Ahammad 2024). Phase-wise application of LDA aids in understanding the dynamic changing content of topics under discussion. Dynamic topic modeling was performed on worldwide tweets using LDA in Sharma et al. (2022). This project was implemented over five temporal phases in the duration of 20 December 2020 to 16 October 2021. The top-10 significant terms or unigrams were plotted for each temporal phase in the form of bar charts. The topmost discussion that surfaced from the topic analysis was the appraisal and criticism of different vaccine brands.

Sentiment analysis of Covid-19 tweets may be complemented by topic analysis for providing a more interpretable quantitative and qualitative analysis of social media posts. Examples of such an agglomeration include XLNet transformer model and Word Clouds (Bansal et al. 2022), pre-trained BERTweet transformer model and Word Clouds (Bansal et al. 2025), and VADER and Word Clouds (Verma et al. 2022). Word clouds are a simple and effective means for extracting significant terms in a document. The font size in the word cloud represents the count of the word in the document. Kwok et al. collected 31,000 tweets of Australian users in the timespan of January to October 2020, and performed sentiment analysis and topic modeling using LDA (Kwok et al. 2021). A pre-defined library in R was used to classify the tweets into either positive or negative sentiment, and one of the eight emotions: fear, anger, anticipation, trust, joy, surprise, disgust, and sadness. As per their findings, two-thirds of the tweets were positive and only one-third were negative or skeptical of the vaccines. Melton et al. performed sentiment analysis and topic modeling on Reddit posts on a month-by-month basis (Melton et al. 2021). Their findings reiterated previous research on the sentiments being primarily positive towards the vaccination process. The topic modeling results revealed that users were mostly concerned with the side-effect of vaccines than any other topic. An interesting related research area is stance analysis (Martínez et al. 2023; Küçük and Arıcı 2022; Yanto et al. 2021) wherein the respondents are classified based on their stance on certain issues pertaining to the Covid-19 vaccination program. This would require annotation of tweets with pre-defined stance labels (Melton et al. 2021) (Martínez et al. 2023) or a stance analysis from general public by asking them to reply to pre-defined questions with labeled stances (Küçük and Arıcı 2022). This problem is different from the sentiment and topic analysis problem tackled in this paper wherein the emotional polarity of social media posts is the focus of investigation. Table 1 summarizes a few case studies following distinctive approaches for sentiment analysis and/or topic modeling of Covid-19 vaccine-related discussions on social media.

Previous case studies on Covid-19 vaccine-related social media posts have several limitations. Firstly, most works have conducted isolated studies on sentiment analysis and topic modeling, and ignore the strong correlation between sentiment labels and the topic of discussion. Secondly, dynamic topic modeling is not the usual norm, and clusters of topics are usually detected for the entire database corresponding to a predefined time window. Thirdly, very few studies have attempted to correlate the detected topics with national/international events that may have triggered a discussion on social media.

In the current case study, we seek to advance over existing methods by introducing dynamic topic modeling of vaccine-related discussions at sentiment fluctuations in the longitudinal sentiment graphs, in order to identify the events of national/international importance that led to a surge in public opinion. Moreover, we chose a timeline encompassing the vaccination rollout in India and coinciding with the spread of the second Covid-19 wave.

4 Proposed methodology

This section presents the proposed strategy for longitudinal sentiment analysis and topic modeling at sentiment fluctuations in the longitudinal graphs. Figure 2 visualizes the

Table 1 Summary of some distinctive studies on sentiment analysis and topic modeling of Covid-19 vaccine-related discussions on social media

Author & Year	Dataset used	Sentiment analysis method	Topic modeling method	Key findings
Lyu et al. (2021)	1,499,421 tweets collected between 11 March 2020 to 31 January 2021	Weekly mean sentiment scores computed using National Research Council of Canada Emotion Lexicon	LDA	16 topics were detected that were grouped into 5 themes. In a separate study, it was found that the sentiments were majorly positive as the vaccinations progressed
Marcec and Likic (2022)	701,891 tweets on AstraZeneca, Pfizer and Moderna vaccines from 1 December 2020 to 31 March 2021	Longitudinal sentiment analysis (day-wise) using AFINN	N/A	It was observed that while the sentiments for Pfizer and Moderna appeared to be positive and stable, the AstraZeneca vaccine garnered more negative sentiments over time
Qorib et al. (2023)	42,796 tweets from 26 September 2021 to 7 November 2021	Five machine learning classifiers and three vectorization techniques for daily sentiment analysis of tweets annotated using Azure machine learning, VADER and TextBlob	N/A	The authors observed a daily increase in the positive sentiments computed by VADER indicating a growing warmer response of the public towards the vaccination program
Bansal et al. (2023)	6,000 annotated tweets available from online dataset	Domain-specific transformer models:—CT-BERT and BERTweet that are pre-trained on Covid-19 tweets	N/A	Domain-specific transformers outperform generic transformers for the sentiment analysis task. Also, text oversampling is not beneficial for learning from skewed class distributions
Lin (2023)	226,373 tweets from 1 December 2020, to 15 November 2021	N/A	LDA with agglomerative clustering for detecting top-20 topics in a phase-wise manner	The dendogram of topic clusters revealed that discussions centered on vaccine brands, booster doses and vaccine effectiveness
Lindelöf et al. (2023)	5,163 negative tweets from 1 March 2020 to 31 July 2021	Negative tweets were segregated using support vector machine	BERTopic model was used for topic modeling of negative tweets	The topics detected were primarily about conspiracy theories, vaccine safety and vaccine side-effects
Boucher et al. (2023)	596,987 tweets on HPV vaccine during the Covid-19 pandemic from January 2019 to May 2021	BERT transformer model fine-tuned on random samples annotated by analysts	Thematic analysis of tweets was performed using hierarchical agglomerative clustering of topic themes and TF-IDF computation	The method heavily relied on the analysis of vaccine-hesitant group (negative tone detected) and vaccine-confident group (neutral tone detected) in social networks
Catelli et al. (2023)	1,602,940 tweets from January 2021 to February 2022	Count of features based on co-occurrences of noun phrases and sentiment words for eg. (AstraZeneca, hospitalized)	MultiWordNet lexical database for the qualitative analysis of tweets	It was concluded that while positive tweets were more about scientific breakthroughs, negative tweets were more about delayed vaccine development and vaccine safety
Ahammad (2024)	10,728 news headlines and posts	VADER	LDA	Negative sentiment analysis revealed 20 topics on vaccine effectiveness, while positive sentiment analysis revealed 18 topics on complacency in vaccination procedure

Table 1 (continued)

Author & Year	Dataset used	Sentiment analysis method	Topic modeling method	Key findings
Bansal et al. (2025)	65,164 and 11,190 tweets associated with Covaxin and Covishield vaccines, respectively, from January 2021 to May 2021 (phase 1) and June 2021 to November 2021 (phase 2)	VADER	Word-cloud analysis of positive and negative tweets	Analysis revealed that the number of negative tweets were more than positive tweets due to discussion on "low availability of vaccines" as indicated from the pre-dominant n-grams in word clouds

workflow of the proposed approach. Each step is detailed in the following sub-sections. We start with a description of data collection and text pre-processing steps, before detailing the sentiment analysis and topic modeling procedures.

4.1 Data collection

We gathered geographically-filtered tweets originating from India from the "Covid-19 all vaccines tweet" dataset provided by Gabriel Preda on Kaggle (https://www.kaggle.com, gpreda, all-covid19-vaccines-tweets[Accessed last on 31st Dec 2021]. 2021). This dataset contains world-wide geotagged tweets related to seven Covid-19 vaccines (Pfizer, Sinopharm, Sinovac, Moderna, Oxford/AstraZeneca, Covaxin, Sputnik-V) collected using Twitter API from public tweet streaming data. The dataset has 16 feature columns (id, user_name, user_location, user_description, user_created, user_followers, user_friends, user_favourites, user_verified, date, text, hashtags, source, retweets, favorites, is_retweet). A visualization of the dataset, illustrating the 16 features and the values they assume, is shown in Fig. 3.

4.2 Text pre-processing

The user_location attribute of the dataset is used to determine the country from which the tweet originated, and the tweets originating from India were filtered out by regular expression-based pattern matching of the keyword "india". We dropped the tweets having ambiguous user locations which amounted to 29% of the 141,645 worldwide tweets available in the dataset. We collected a total of 44,130 vaccine-related tweets originating from India in the duration of 12 December 2020 to 11 November 2021. A natural language toolkit was used to pre-process the data. The tweets were first pre-processed and cleaned using regular expressions prior to the sentiment analysis stage. Since the supervised models and lexicons used in the study are trained on the English language, therefore, we removed all non-English words from the tweets. All characters were converted to lower case and punctuations were removed. All URLs, hashtags, emails and non-printable characters were removed. Further, in addition to removing the English stopwords, the following commonly-occurring keywords were identified and removed from the tweets: '#covid', '#coronavirus', '#coronavirusoutbreak', '#coronavirusPandemic', '#covid19', '#covid_19', '#epitwitter', '#ihavecorona''. The deletion of these irrelevant details was performed to further enhance the results of sentiment analysis.

4.3 Sentiment analysis of geo-tagged tweets

AFINN is a rule- and lexicon-based unsupervised sentiment analysis tool created by Finn Årup Nielsen (Nielsen 2011)

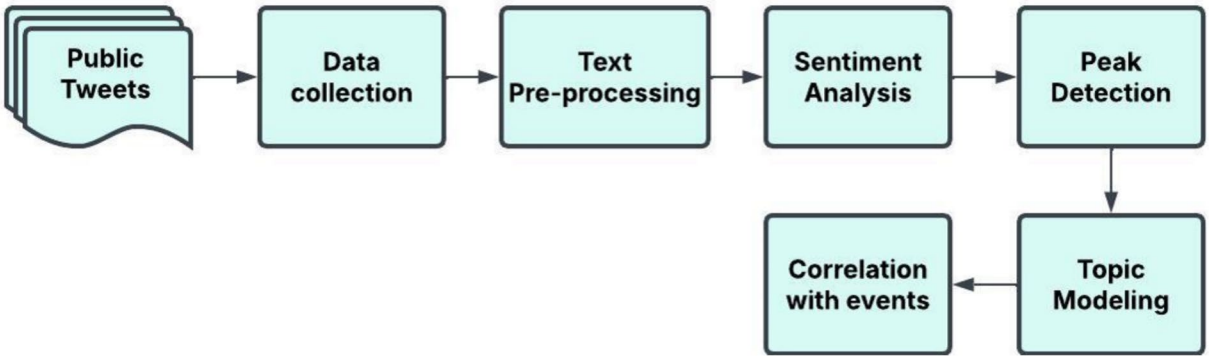


Fig. 2 Workflow of the proposed approach

that is popularly used for the sentiment analysis of social media posts. It computes a compound score for each sentence ranging from -5 (most negative) to + 5 (most positive). If the score is positive, it indicates a positive sentiment. If the score is negative, it indicates a negative sentiment. We consider the AFINN lexicon for the sentiment analysis of geo-tagged tweets since apart from being efficient, it is unsupervised and does not require training; it is difficult to acquire annotated tweet datasets associated with vaccine-related discussions. In the majority of cases, tweets that contained words like "good", "effective", "protects", "approve", "proud" etc. were correctly labeled as positive. For instance, the tweet "*Pfizer's Coronavirus Vaccine Protects Against New Variants Found in UK, South Africa: Study*" was correctly classified as positive with a score of + 1. Likewise, negative tweets were characterized by words like "dead", "adverse", "suffered", "infection". For instance, the tweet "*Two dead in #Norway, received #Pfizer's coronavirus vaccine recently*" was correctly classified as negative with a score of -3. Another instance is the tweet "*No mention of efficacy …which is 50%… lowest amongst all Vaccine #COVID19Vaccines #Moderna #PfizerBioNTech #CovidVaccine #Covishield*", that is correctly predicted as negative (score of -2). On the other hand, non-committal posts like "*Would you vaccinate or not?*", or those tweets stating facts, such as "*Pfizer-BioNTech to produce 2 billion doses of COVID-19 vaccine in 2021* | via *@IndiaTVNews*", were correctly classified as neutral (score of 0).

4.4 Topic modeling at sentiment peaks and correlation with events

The following steps were undertaken for analyzing the sentiment outputs. Longitudinal graphs (Valdez et al. 2020) of positive, negative and neutral sentiments versus time were plotted on a day-wise basis between 12 December 2020 to 11 November 2021. The number of positive, negative and neutral tweets were summed up separately for each day, creating three longitudinal graphs, one for each sentiment. The peaks in the sum of positive and negative sentiments in each of the longitudinal graphs indicate a fluctuation in public opinion. The five highest peaks corresponding to the positive sentiment and the five highest peaks corresponding to the negative sentiment are derived from the positive and negative longitudinal graphs, respectively. The rank is determined from the decreasing order of the day-wise tweet count in the respective longitudinal graph (positive or negative). For determining the sentiment peaks corresponding to significant sentiment fluctuations in the graphs, we follow previous research (Nguyen et al. 2013) that extracts the highest peak points for positive and negative sentiments, respectively. We then performed dynamic topic modeling over sentiment peaks (both positive and negative) in the longitudinal graphs in order to understand the events that influence public receptiveness to government policies related to the vaccination program. Positive and negative vocabularies of the trending topics were extracted at each sentiment peak for identifying the topics that may have led to a fluctuation in public sentiments. Salient terms having high term frequencies were extracted from the top-20 topics detected using LDA. The top-20 salient terms at the peak points were also visualized using bar charts indicating the overall term frequency which is a popular indicator of the significant keywords in a document (Susan and Keshari 2019; Sicilia-Garcia et al. 2002; Garner et al. 2019; Sharma et al. 2023). Dynamic topic modeling over moving time windows helps to identify various events of national/international importance that may have influenced discussions and sentiment fluctuations on social media. To set the LDA hyperparameters, we followed the guidelines in literature (Panichella 2021). The number of topics was set to 20, alpha which influences topic distribution in each document was set to 0.05 (= 1/[number of topics]), beta which influences term

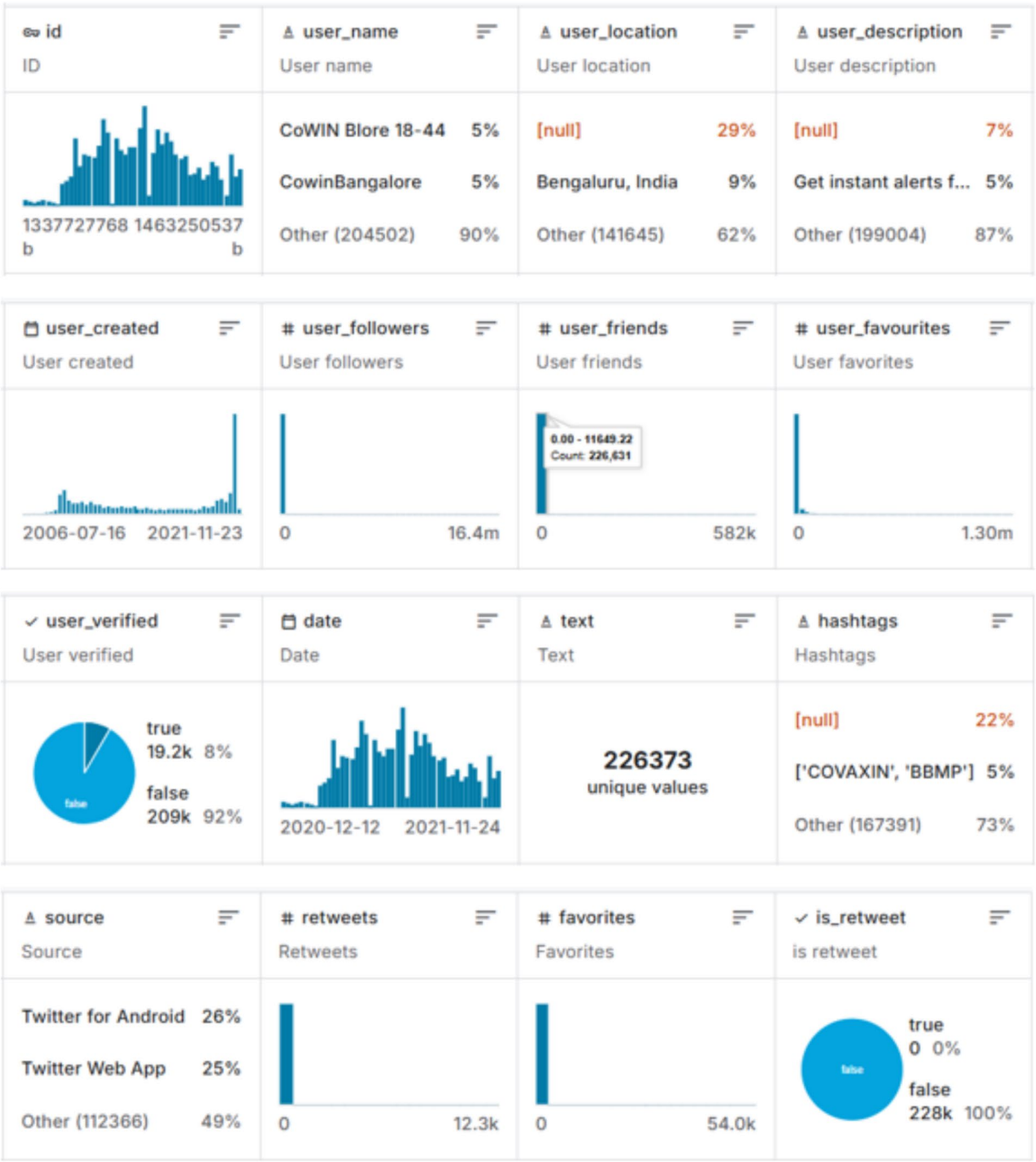


Fig. 3 Visualization of the dataset illustrating the 16 features and the values they assume

distribution in each topic was set to 0.05 (= 1/[number of topics]), and number of iterations was set to 200. We found that the list of salient terms did not change significantly when the number of topics was varied between 10 to 30. For correlating the detected topics with national/international events, a manual curation of leading news articles published in online portals of leading newspapers, magazines, World Health Organization (WHO) website (for constant updates on vaccine brands), and Reuters international news agency, is performed, pertaining to the specific dates on which sentiment peaks are detected

in the longitudinal graphs. A manual approach is taken due to the complexity of real-world events and to ensure correctness in linking contexts.

5 Results and discussions

This section presents the results and analysis of the experimentation. All experiments were conducted in Python (version 3.8) on a 2.8 GHz Intel core PC. The code used in our experiments and the associated data are made available online[1] to facilitate future research.

5.1 Performance comparison of popular sentiment analysis systems on a smaller annotated tweets dataset

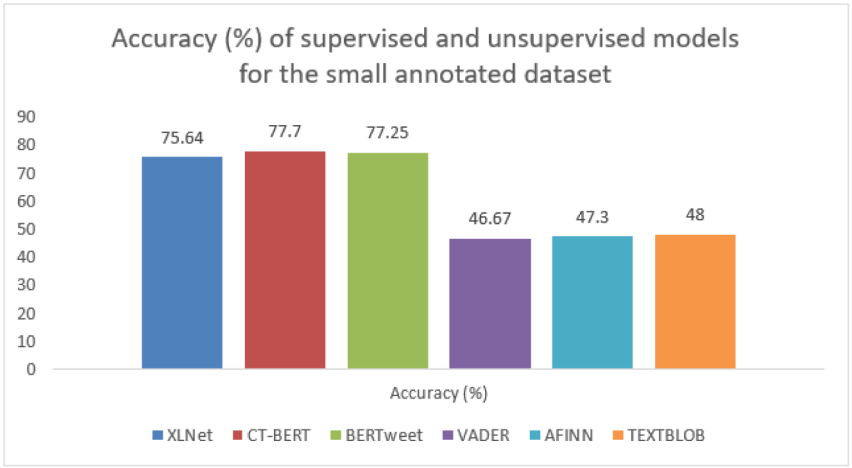
We first present a comparative analysis of some popular supervised (XLNet, CT-BERT and BERTweet transformer models) and unsupervised (VADER, AFINN, TextBlob) sentiment analysis systems trained on a smaller annotated subset of the dataset which was also released by Gabriel Preda (https://www.kaggle.com, gpreda, all-covid19-vaccines-tweets[Accessed last on 31st Dec 2021]. 2021). This comparison serves as a precursor to the ultimate sentiment analysis performed on the complete set of 44,130 unannotated tweets using AFINN. There are a total of 6,000 tweets in the smaller annotated dataset out of which the majority (3,680 tweets) are neutral tweets, 1,900 are positive tweets while the negative tweets are in a minority (420 tweets). A five-fold cross-validation was performed for the supervised models using different 80:20 train-test splits of the dataset. Training using supervised models is expected to be biased towards the neutral sentiment due to class-imbalance, while the negative sentiments are likely to be mis-classified due to inadequate representation. It was proved in previous research (Bansal et al. 2023) that over-sampling using Language Model based Oversampling Technique (LMOTE) and Synthetic Minority Oversampling Technique (SMOTE) does not improve the performance of supervised models when the minority samples are few, as in this case. The reason is that data augmentation based on limited training data (few minority samples), generally produces noise, which may degrade the overall performance. In case of the unsupervised models—VADER, AFINN and TextBlob, no hyperparameters are involved since they are lexicon-based and do not require training. These models generate a compound sentiment score whose value determines the sentiment polarity of the input text. For instance, for a given sentence, VADER generates a compound score for each sentence in the range + 1 (most positive) to − 1 (most negative). The AFINN tool computes a compound score in the range + 5 (most positive) to − 5 (most negative), while TextBlob computes a compound score in the range + 1 (most positive) to − 1 (most negative). CT-BERT and BERTweet are pre-trained transformer models trained on English tweets on the Covid-19 pandemic. For the supervised models, Adam optimizer was used with a learning rate of 2e−5. The loss function is the cross-entropy loss function. While CT-BERT is trained on 160 million Covid-19 related English tweets, BERTweet is trained on 850 million English tweets with 5 million tweets belong to the Covid-19 pandemic. We additionally fine-tuned the BERTweet and CT-BERT models on the smaller annotated dataset, for enhanced performance while predicting the sentiments of the larger unannotated dataset comprising 44,130 tweets. From a comparison of the accuracy, F1-score and Mathew's correlation coefficient (MCC) scores of all six models, shown in Fig. 4, we observe that BERTweet and CT-BERT among the supervised models, and VADER and AFINN among the unsupervised models, showcase the best performance on the basis of F1-score and MCC which are the best indicators of model performance under class imbalance. Therefore, we plot the longitudinal sentiment graphs for the 44,130 unannotated tweets using all four models AFINN, VADER, BERTweet, and CT-BERT, in the next phase of our experiments.

5.2 Sentiment analysis of geo-tagged tweets

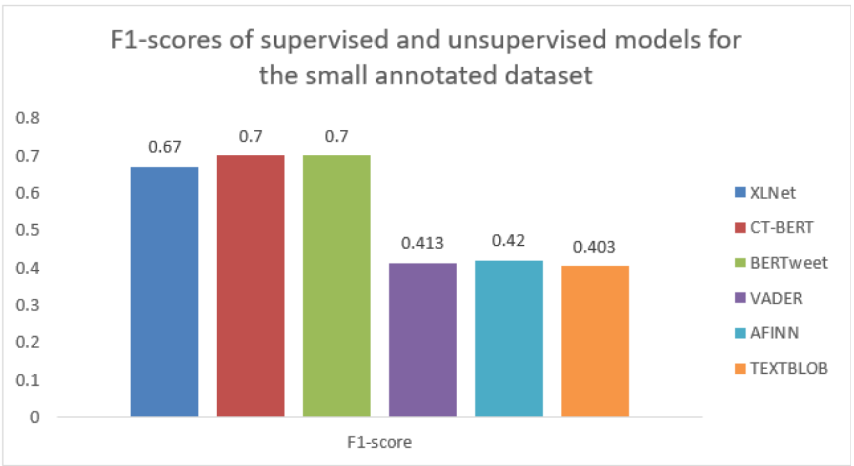
A temporal analysis of the positive and negative tweets over the duration of one year reveals the dynamically evolving emotional mindset of the people towards the ongoing vaccination program as the country progressed through the devastating second wave caused by the Delta variant of SARS-Cov-2 coronavirus in the months of April-July 2021. The neutral tweets are not considered for the analysis since they do not reflect the emotional mindset of the users. Figures 5 and 6 illustrate the longitudinal sentiment graphs corresponding to day-wise aggregated counts of positive, negative and neutral tweets, computed using VADER and AFINN, respectively, in the duration of 12 December 2020 to 11 November 2021. Likewise, Figs. 7 and 8 illustrate the longitudinal sentiment graphs corresponding to the pre-trained models BERTweet and CT-BERT.

In the initial months from December 2020 to February 2021, the number of vaccine-related tweets were relatively low in all the four graphs since vaccinations were a new concept. The vaccination program was launched in India on 16 January 2021. All three sentiments are shown in the longitudinal graphs to comprehend the relative scales of the three sentiments. The neutral tweets that mostly state facts about Covid-19 vaccines form a significant chunk of the total

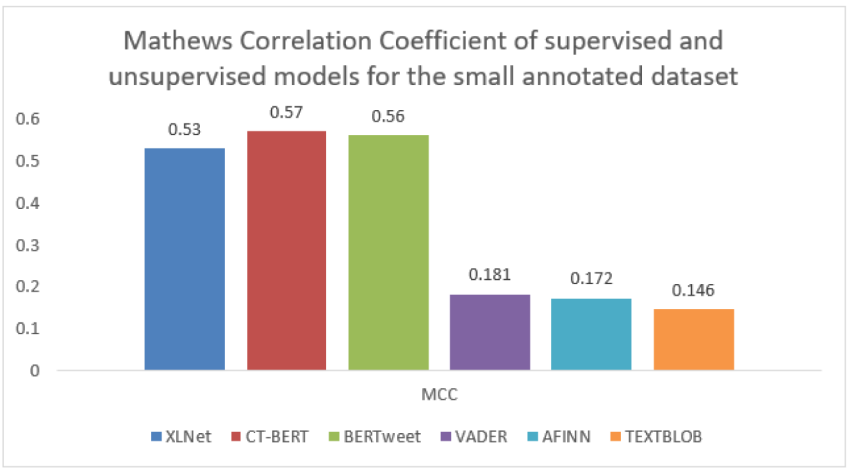
[1] https://github.com/Ace117MC/Topic-Modeling-of-Social-Media-Posts-at-Sentiment-Fluctuations-in-Longitudinal-Sentiment-Graphs



(a)



(b)



(c)

Fig. 4 Accuracy, F1-score and MCC scores of supervised (XLNet, CT-BERT, BERTweet) and unsupervised (VADER, AFINN, TextBlob) sentiment analysis systems for the smaller annotated dataset

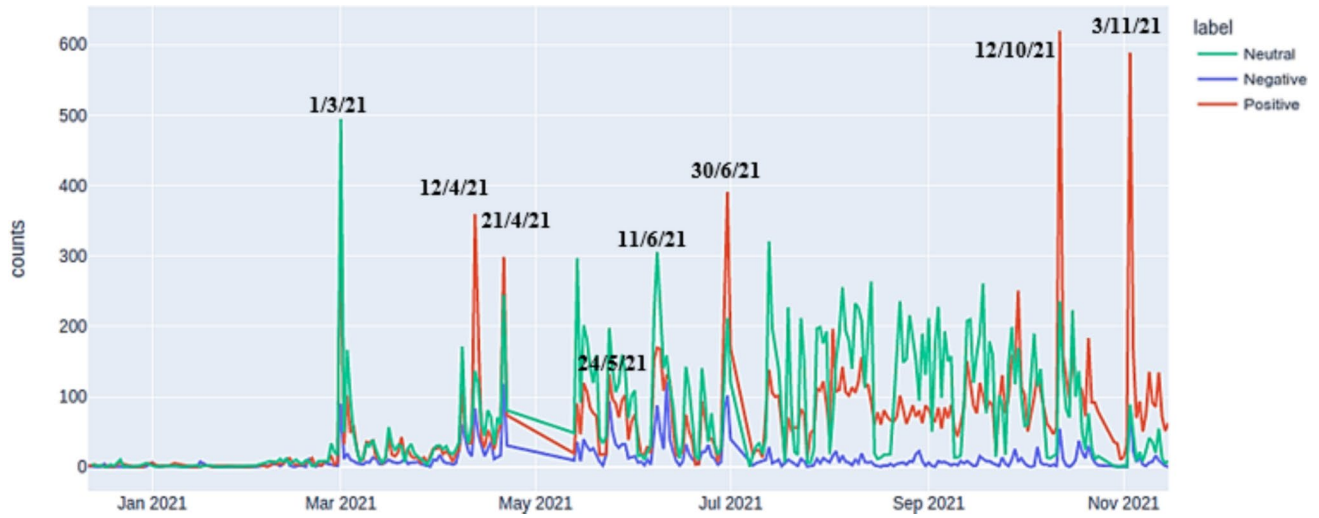


Fig. 5 Sentiment analysis graph using VADER showing the day-wise aggregated count of positive, negative and neutral tweets in the duration of 12 December 2020 to 11 November 2021. The highlighted dates indicate the peaks in the longitudinal sentiment graphs

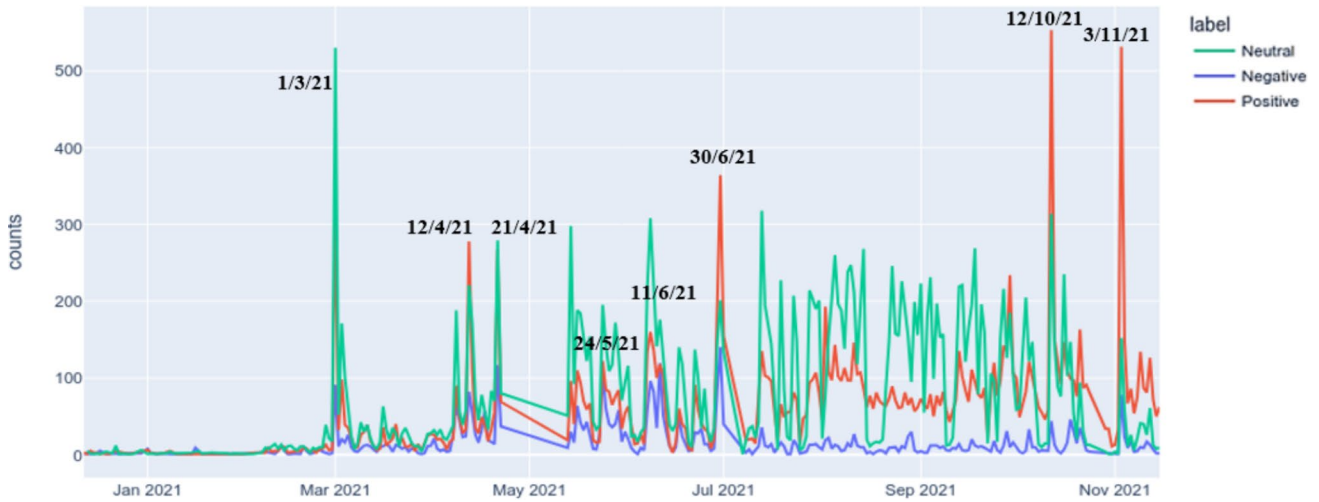


Fig. 6 Sentiment analysis graph using AFINN showing the day-wise count of positive, negative and neutral tweets in the duration of 12 December 2020 to 11 November 2021. The highlighted dates indicate the peaks in the longitudinal sentiment graphs

number of tweets. The number of positive tweets is, at most times, higher as compared to the number of negative tweets, which indicates that the majority of users hold a positive attitude towards vaccination. The negative tweets have a higher concentration in the duration of April 2021 to July 2021 during the progression of the second wave in India.

VADER and AFINN detect the negative tweets efficiently, especially during the second wave in April-July 2021, as observed from Figs. 5 and 6, respectively. The results of AFINN were comparatively superior to VADER in the experiments conducted on the smaller annotated dataset (Fig. 4). A comparison of the graphs in Figs. 5 and 6 proves that the negative sentiments are more well-defined in the AFINN graph with a reduced gap between the positive and negative graphs during the second wave (April-July 2021), as compared to the VADER graph where a significant gap exists between the positive and negative sentiments even at the peak of the second wave (April–July 2021). Therefore, we consider the AFINN results for topic modeling in the next phase of experiments.

An examination of the BERTweet and CT-BERT longitudinal graphs in Figs. 7 and 8 reveal that most of the tweets were predicted as neutral, and a small minority as positive, while a miniscule population of tweets were labeled as negative. The supervised models, therefore, are prone to biases and inaccuracies, and do not classify positive and negative

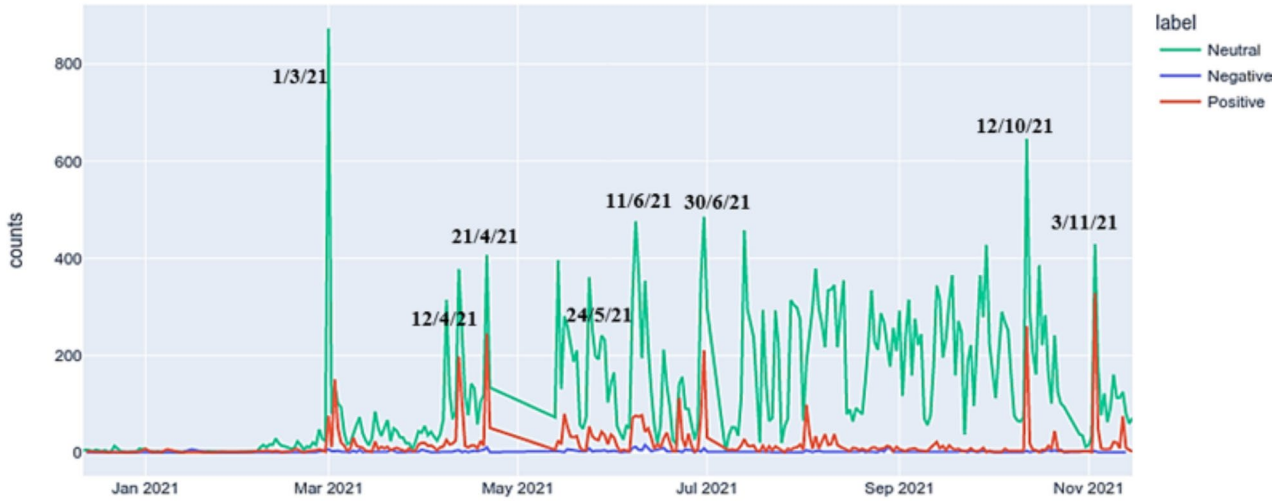


Fig. 7 Sentiment analysis graph using BERTweet showing the day-wise count of positive, negative and neutral tweets in the duration of 12 December 2020 to 11 November 2021. The highlighted dates indicate the peaks in the longitudinal sentiment graphs

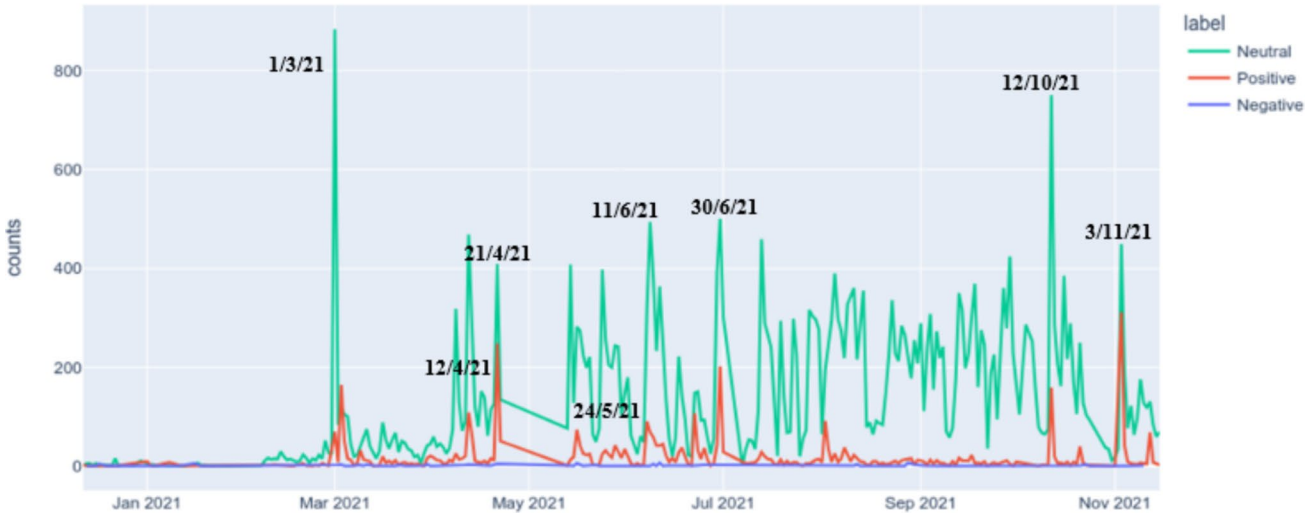


Fig. 8 Sentiment analysis graph using CT-BERT showing the day-wise count of positive, negative and neutral tweets in the duration of 12 December 2020 to 11 November 2021. The highlighted dates indicate the peaks in the longitudinal sentiment graphs

sentiments as well as the unsupervised models in this particular test scenario. This disparity becomes more evident in the months of October–November 2021, towards the end of the time window, when the majority of the tweets were predicted as neutral by the supervised models and positive by the unsupervised models. This result was on expected lines since the performance of the supervised models was bound to be affected by the imbalanced class profile of the small annotated dataset on which it was fine-tuned on (Sect. 5.1). The positive sentiment peaks in the graphs in Figs. 5, 6, 7, 8 were found to mostly coincide, though the actual counts significantly differed.

A statistical significance test (Cohen's Kappa) revealed that AFINN and VADER results showed near perfect agreement (Kappa = 0.8204), whereas there was a notable disagreement of predictions between AFINN and CT-BERT (Kappa = 0.09728), and between AFINN and BERTweet (Kappa = 0.1251). There was moderate agreement of predictions between the two supervised models CT-BERT and BERTweet (Kappa = 0.669) illustrating the unstable and outcome of supervised learning for a limited dataset with skewed class distribution.

AFINN is ultimately chosen as the sentiment analysis system for determining the sentiment peaks for topic modeling due to its better performance and ease of use for large

unannotated data. Out of the 44,130 tweets, 22,672 tweets (51.38%) were detected as neutral, 17,141 tweets (38.84%) were detected as positive, and 4317 tweets (9.78%) were detected as negative, by AFINN. The number of positive sentiments significantly increased after the second wave subsided in July 2021, as observed from the higher peaks in October–November 2021 (Fig. 5), indicating a higher optimism and willingness to participate in the vaccination program.

5.3 Topic modeling at sentiment peaks

We analyzed the positive and negative sentiment graphs over time, and identified the topics of discussion at sentiment fluctuations in the graphs using LDA which is an unsupervised machine learning model for topic modeling. The five highest peaks corresponding to the positive sentiment (Pos) and the five highest peaks corresponding to the negative sentiment (Neg) are shown in Table 2. The rank is determined from the decreasing order of the day-wise tweet count in the respective longitudinal graph (positive or negative). For determining the sentiment peaks corresponding to significant sentiment fluctuations in the graphs, we followed previous research (Nguyen et al. 2013) that analyzes the top-ranked peak points for positive and negative sentiments, respectively. As per Table 2, the sentiment peaks (positive and negative), determined by the above procedure, were observed on the specific dates−1 March 2021, 12 April 2021, 21 April 2021, 24 May 2021, 11 June 2021, 30 June 2021, 12 October 2021 and 3 November 2021. These dates are also shown highlighted in the longitudinal graphs in Figs. 5, 6, 7, 8. We perform topic modeling on the tweets associated with the days corresponding to the sentiment peaks listed in Table 2. Table 3 shows the results of topic analysis on the specific dates that correspond to peak points in the positive and negative sentiment graphs. The top-20 frequently occurring terms, also known as salient terms, in the top-20 topics detected through LDA are used for the analysis. The frequencies of the salient terms in the topics detected at sentiment fluctuations are visualized in the form of bar charts in Figs. 9, 10, 11, 12, 13, 14, 15, 16 for the dates entered in Table 3 and highlighted in Figs. 5, 6, 7, 8. The green bars in Figs. 9, 10, 11, 14, 15, 16 indicate positive sentiments while the red bars in Figs. 12, 13 represent negative sentiments. These bar charts help to infer the most discussed topics on Twitter, on specific dates, which may have caused the spikes in the sentiment graph.

For correlating the detected topics with national/international events, a manual curation of leading news articles published in online portals of leading Indian newspapers (The Hindu, New Indian Express), Indian magazines (India Today), WHO website (www.who.int), and Reuters news

Table 2 Sentiment peaks in day-wise longitudinal graphs of positive and negative sentiments

	1st positive peak	2nd positive peak	3rd positive peak	4th positive peak	5th positive peak	1st negative peak	2nd negative peak	3rd negative peak	4th negative peak t	5th negative peak
Date	12/10/21	3/11/21	30/6/21	1/3/21	12/4/21	30/6/21	11/6/21	21/4/21	24/5/21	1/3/21
Tweet count	553	531	364	335	278	140	119	117	107	91

The five highest peaks corresponding to the positive sentiment and the five highest peaks corresponding to the negative sentiment are shown date-wise, the rank being determined from the respective positive and negative tweet counts

Table 3 Topic modeling using LDA at sentiment peaks (specific dates mentioned in Table 2) in the longitudinal graphs in the duration of 12 December 2020 to 11 November 2021 (positive and negative sentiment labeling using AFINN)

Date	Positive tweets T_p	Negative tweets T_n	Positivity = T_p-T_n	Trending topic	Positive vocabulary	Negative vocabulary	Correlation with major news/events	Overall impact on public and government
1/3/2021	335	91	244 (high)	News regarding the Prime Minister	india pm took vaccine	vaccine pm jab fears hesitancy	Prime Minister of India takes first dose of Covaxin (https://www.thehindu.com/news/national/pm-modi-takes-first-jab-of-covaxin-against-covid-19/article33958677.ece)	Boost in public confidence for the vaccination program in the face of fears about vaccine safety
12/4/2021	278	82	196 (high)	Russian vaccine emergency use approval	approval emergency sputnik vaccine india	emergency authorization shortage sputnik vaccine india	Approval for Sputnik-V vaccine for emergency use in India (https://www.thehindu.com/sci-tech/health/expert-panel-recommends-granting-approval-to-covid-19-vaccine-sputnik-v-for-emergency-use-in-india/article34301838.ece)	Boost in public confidence for imported vaccines in the face of concerns about vaccine shortage

Table 3 (continued)

Date	Positive tweets T_p	Negative tweets T_n	Positivity = T_p-T_n	Trending topic	Positive vocabulary	Negative vocabulary	Correlation with major news/events	Overall impact on public and government
21/4/2021	267	117	150 (high)	Domestic vaccine trial completion for multiple Covid-19 variants	covaxin effective neutralises multiple variants icmr	covaxin efficacy severe covid	Indian vaccine Covaxin completes phase 3 trial with 78% efficacy on multiple variants as per Indian Council of Medical Research (ICMR) (https://www.thehindu.com/sci-tech/health/updated-data-from-covaxin-phase-3-trial-shows-78-efficacy/article34376018.ece)	Boost in public confidence for domestic vaccine with a growing segment expressing concerns of Covaxin efficacy in the face of the severe Covid-19 s wave
24/5/2021	122	107	15 (low)	Imported vaccine sale refusal	good news russia vaccine coming	vaccines sell refused directly delhi moderna vaccine	Pfizer, Moderna refuse to sell Covid-19 vaccines to Delhi government directly (https://www.newindianexpress.com/cities/delhi/2021/may/24/nowpfizer-moderna-refuse-to-sell-covid-19-vaccines-to-delhi-government-directly-2306861.html)	Drop in public morale and negative opinions forces the government to seek alternatives to importing Pfizer and Moderna vaccines

Table 3 (continued)

Date	Positive tweets T_p	Negative tweets T_n	Positivity = T_p-T_n	Trending topic	Positive vocabulary	Negative vocabulary	Correlation with major news/events	Overall impact on public and government
11/6/2021	118	119	− 1 (very low)	Domestic vaccine emergency use disapproval by US-based FDA	first dose vaccine got vaccinated	us approval fda rejects emergency covaxin	Setback for Covaxin, US Food and Drug Administration (FDA) says no to emergency use, suggests 'full approval' route (https://www.indiatoday.in/coronavirus-outbreak/story/setback-for-covaxin-us-fda-rejects-emergency-use-nod-1813521-2021-06-11)	Setback for government as FDA rejects emergency use of the domestic vaccine Covaxin. Negativity on social media spurs the government to seek global approval of Covaxin
30/6/2021	364	140	224 (high)	Domestic vaccine efficacy proved for the variants under circulation	covaxin neutralises delta slots dose free availability	covaxin irregularities brazil suspend deal	Covaxin effectively neutralizes the Delta variant (https://www.thehindu.com/news/international/indias-covaxin-effectively-neutralises-delta-variant-of-covid-says-nih/article35047706.ece) Brazil suspends Covaxin deal (https://www.reuters.com/world/india/brazil-suspend-indian-vaccine-deal-graft-allegations-probed-2021-06-29/)	Significant boost in public confidence for domestic vaccine efficacy against the Delta variant with increasing queries on the next available slots for free dose of the vaccine, in the face of slight criticism on Brazil suspending the deal for Covaxin procurement

Table 3 (continued)

Date	Positive tweets T_p	Negative tweets T_n	Positivity = T_p-T_n	Trending topic	Positive vocabulary	Negative vocabulary	Correlation with major news/events	Overall impact on public and government
12/10/2021	553	44	509 (very high)	Domestic vaccine approval for kids	slots dose covaxin age children approved years	authorisation emergency children age confusion evaluation underway	Covaxin gets emergency use approval for kids aged 2–18 years (https://www.indiatoday.in/coronavirus-outbreak/story/covaxin-gets-emergency-approval-kids-2-18-years-1863866-2021-10-12)	Increased numbers of children getting vaccinated in the face of concerns and doubts regarding the process
3/11/2021	531	78	453 (very high)	WHO approval obtained for domestic vaccine emergency use	listing congratulations gets approval finally approved	emergency listing india vaccine granted finally	WHO issues emergency use listing for Covaxin (https://www.who.int/news/item/03-11-2021-who-issues-emergency-use-listing-for-eighth-covid-19-vaccine)	Boost in public confidence on getting global approval for domestic vaccine emergency use

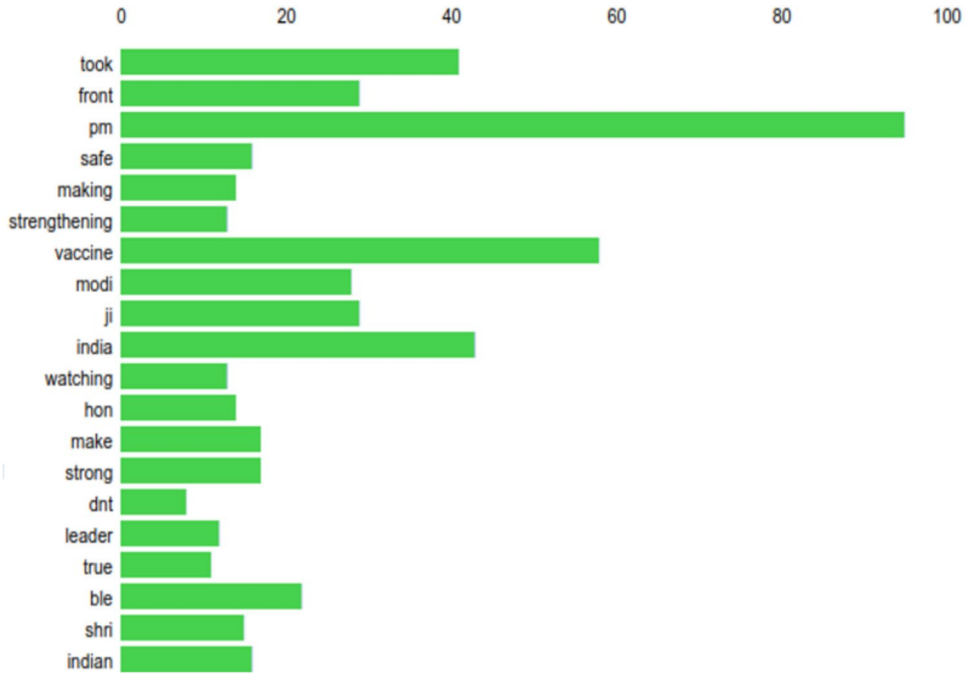


Fig. 9 Salient term frequencies in the top-20 topics detected from the positive sentiment tweets using LDA on 1 March 2021

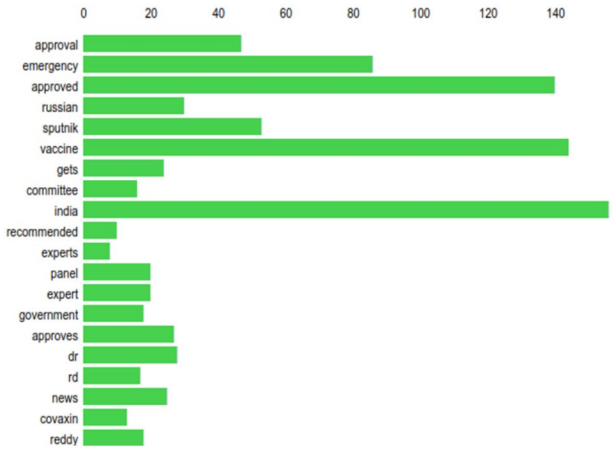


Fig. 10 Salient term frequencies in the top-20 topics detected from the positive sentiment tweets using LDA on 12 April 2021

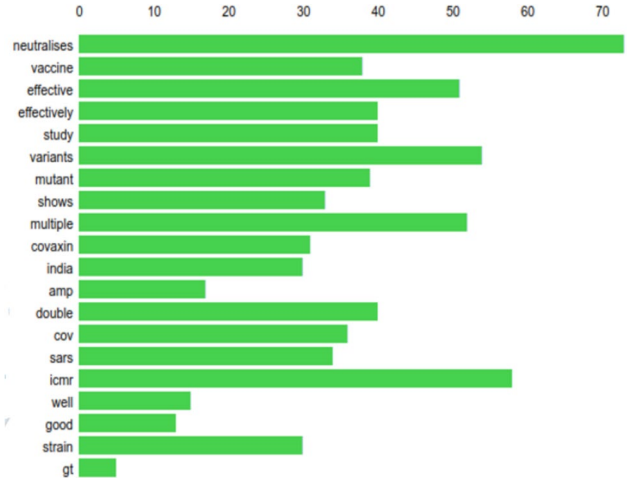


Fig. 11 Salient term frequencies in the top-20 topics detected from the positive sentiment tweets using LDA on 21 April 2021

agency (www.reuters.com), is performed, pertaining to the specific dates on which sentiment peaks are detected in the longitudinal graphs. Table 3 enlists the significant national/international events (with citations to news articles) that occurred on these dates that may have influenced a shift in public opinion. The impact of each event and its discussion on boosting public confidence and framing government policies is summarized in the last column of Table 3.

As noted from Fig. 9, the trending topic on 1 March 2021 was− Indian Prime Minister/PM took the first dose of vaccine (identified by the highest term frequency topics: 'india', 'pm', 'took', 'vaccine'); the most frequent salient term being 'pm'. An overall positive outlook towards vaccination was experienced due to the much-publicized event.

The most-occurring salient terms on 12 April 2021 was observed to be 'approval', 'emergency', 'sputnik', 'vaccine', 'India'. The trending topic can be surmised to be the approval of the Russian vaccine Sputnik-V for emergency use in India, at a time when the country was in the initial stage of the second wave and the number of Covid-19 confirmed cases was rapidly increasing. The news came as a relief to many citizens and resulted in a positive sentiment spike on 12 April 2021.

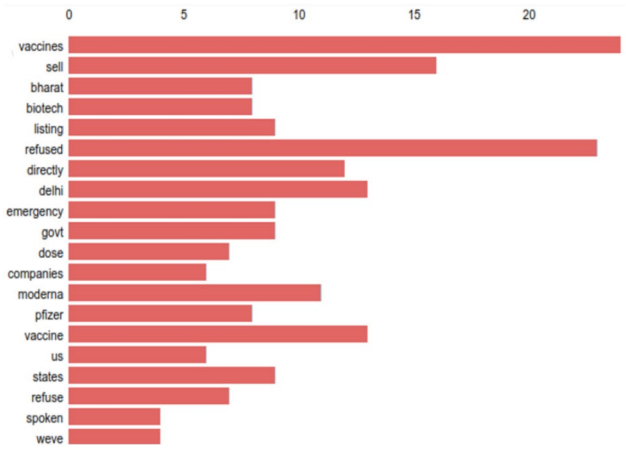


Fig. 12 Salient term frequencies in the top-20 topics detected from the negative sentiment tweets using LDA on 24 May 2021

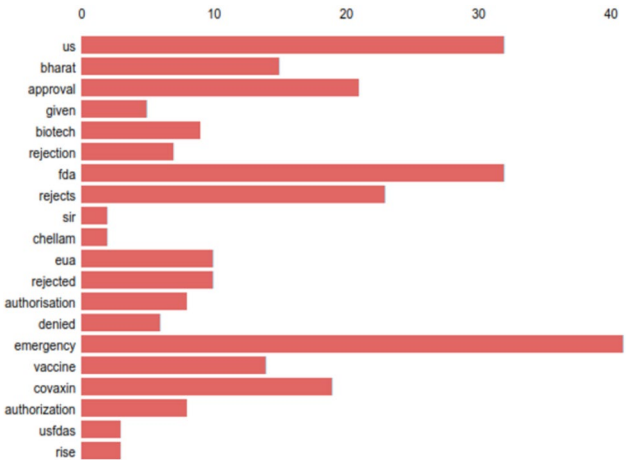


Fig. 13 Salient term frequencies in the top-20 topics detected from the negative sentiment tweets using LDA on 11 June 2021

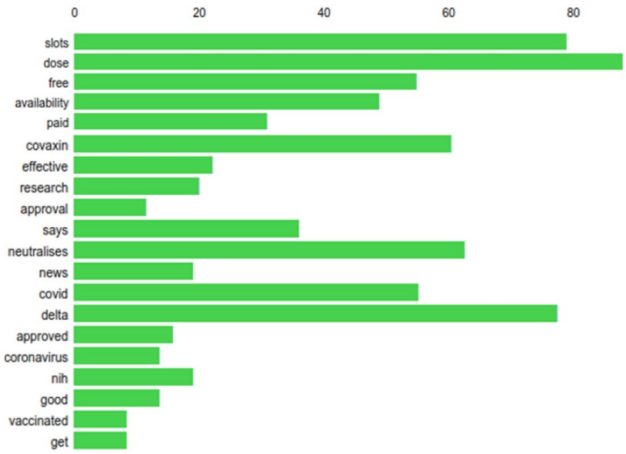


Fig. 14 Salient term frequencies in the top-20 topics detected from the positive sentiment tweets using LDA on 30 June 2021

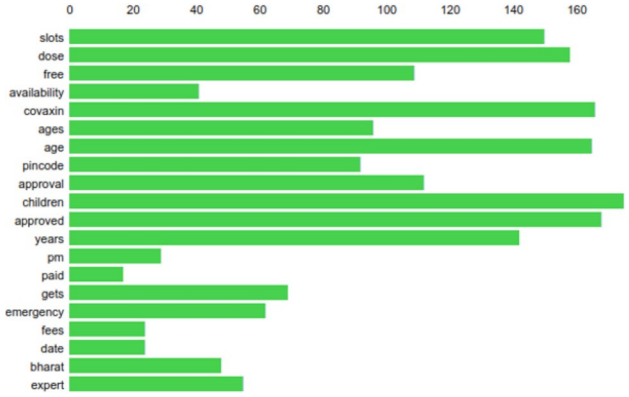


Fig. 15 Salient term frequencies in the top-20 topics detected from the positive sentiment tweets using LDA on 12 October 2021

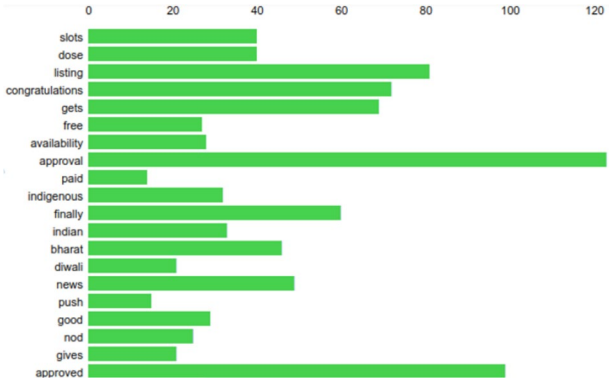


Fig. 16 Salient term frequencies in the top-20 topics detected from the positive sentiment tweets using LDA on 3 November 2021

As per Fig. 11, the top-salient terms on 21 April 2021 were 'covaxin', 'effective', 'neutralises', 'multiple', 'variants', 'icmr'. The trending theme was the positive news of the domestic vaccine Covaxin neutralizing multiple Covid-19 variants as per news released by the Indian Council of Medical Research (ICMR). The public received this news with optimism, as the news served to mitigate fears and apprehensions about the efficacy of the domestic brand.

Fig. 12 reveals the trending topics of 'vaccines', 'sell', 'refused', 'directly', 'delhi', 'moderna', 'vaccine', which corroborates with the news of Pfizer refusing to sell vaccines to the Delhi government on 24 May 2021. At this point the second wave had advanced throughout the country, and the overall outlook was negative and the positivity was low as indicated by the vocabulary at negative sentiment peaks, shown in Fig. 12.

The number of negative tweets spiked on 11 June 2021 (lowest positivity of -1) when an US-based organization (FDA) did not allow the approval of the domestic vaccine Covaxin for emergency use (topics: 'us', 'approval', 'fda',

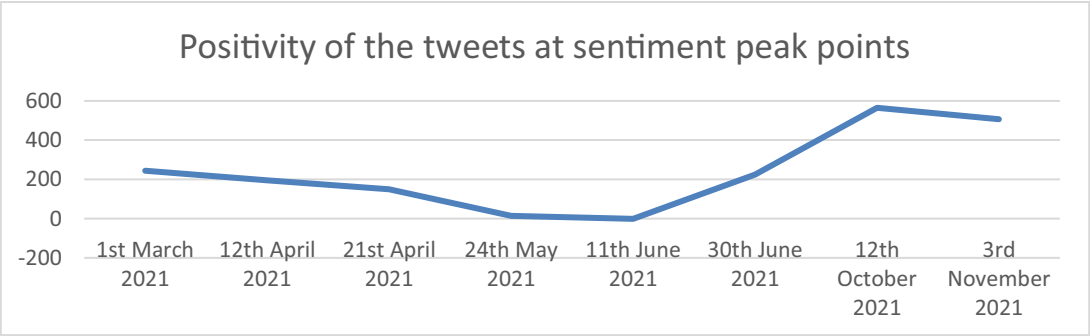


Fig. 17 Positivity graph showing the number of positive tweets in surplus of the number of negative tweets on the specific dates corresponding to sentiment peaks (shown in Table 3)

'rejects', 'emergency', 'covaxin') at a time when the second Covid-19 wave was at its peak (Fig. 13). The public morale and confidence were at their lowest at this point.

The salient terms detected on 30 June 2021 were−'covaxin', 'neutralises', 'delta', 'slots', 'dose', 'free', 'availability'. At a time when the number of Covid-19 cases was slowly declining, news of the efficacy of the domestic vaccine Covaxin neutralizing the Delta variant that caused the second wave, and its free availability for the public, boosted the public confidence once again, as evident from the positive vocabulary in Fig. 14.

The highest number of positive tweets (highest positivity score of 565) were recorded on 12 October 2021 (Fig. 15) when the domestic vaccine Covaxin got approval for administration to children aged between 2 to 18 years (topics: 'slots', 'dose', 'covaxin', 'age', 'children', 'approved', 'years'), indicating that is an example of policy-making in the right direction.

Once the second wave had significantly abated, the positivity towards vaccines increased and rose to an all-time high in the months of October-November 2021 as the public became more enthusiastic towards the vaccination program and expressed willingness to get vaccinated. The most frequently occurring salient terms on the date 3 November 2021 were−'listing', 'congratulations', 'gets', 'approval', 'finally', 'approved'. This points to the event of Covaxin getting global approval by World Health Organization (WHO). A positive spike in sentiments is observed on this occasion at the backdrop of an undercurrent of optimism and positivity at having overcome the worst of the deadly second wave.

The column of trending topics in Table 3 summarizes the primary themes (detected using LDA) that sparked an emotional discussion on Twitter. A scrutiny of the themes or trending topics in Table 3, over time, reveal a low confidence in the vaccination program and vaccine efficacy at the beginning. The positivity measure in Table 3 is computed as the difference between the daily count of positive and negative tweets, and is shown plotted as a function of time in Fig. 17.

As noted from Table 3 and the positivity curve in Fig. 17, a positive spike in sentiments was observed with news of the Prime Minister getting vaccinated dominating the discussions on 1 March 2021 (topics: 'india', 'pm', 'took', 'vaccine'). However, the positivity reduced significantly during the second Covid-19 wave which affected India in the months of April–July 2021. This was followed by days of apprehensions of imported vaccine availability as observed from the low positivity surrounding the news of Pfizer and Moderna deal cancellation (24/5/2021). Concerns about vaccine approval and vaccine efficacy of the domestic vaccine Covaxin predominated the discussions in the interim period. Public confidence got to an all-time high in October–November 2021 with news of vaccine availability for kids (12/10/2021) and WHO approval of the domestic brand Covaxin (3/11/2021). Additionally, the trending topics at sentiment peaks would help the government to frame policies to combat vaccine hesitancy and boost public confidence in the vaccination program. This includes alternative and faster strategies for vaccine availability, imported vaccine procurement, and obtaining global approval for domestic vaccine brands by speeding up the trials for vaccine approval, as observed from the major national/international events listed in Table 3 corresponding to high positivity values (12/10/2021, 3/11/2021).

5.4 Discussions

To summarize, in order to address the research questions raised in Sect. 2, we propose a novel approach for longitudinal sentiment analysis with dynamic topic modeling at positive and negative sentiment peaks. The answers to specific research questions are provided as follows. *RQ1: How to effectively use unsupervised sentiment analysis tools for predicting day-wise sentiments of Covid-19 vaccine-related tweets of users geo-located in India?* We performed

sentiment analysis of tweets originating from India, using AFINN, on a day-wise basis, in the duration of 12 December 2020 to 11 November 2021 when both the vaccination program and the second Covid-19 wave were simultaneously experienced in the country. *RQ2: How to effectively extract the topics of discussion on the days when positive and negative sentiments peaked on Twitter?* LDA was used for dynamic topic modeling at sentiment peaks corresponding to significant fluctuations in the positive and negative longitudinal sentiment graphs. *RQ3: How to correlate the extracted topics and sentiments with events of national and international importance that may have been the cause and effect of the outburst of opinions on social media platforms?* The topics detected using LDA are correlated with manually collated news articles on the days when positive and negative sentiments peaked to identify the national/international events that may have led to an outburst of opinions on social media. The main topics of discussion detected at the sentiment fluctuations using LDA are summarized as: (1) the Prime Minister takes a dose of the vaccine (2) approval for emergency use of Sputnik-V vaccine (4) concerns about the availability of vaccines and delay in getting approval for Covaxin, the domestic vaccine, from US-based FDA (5) euphoria about WHO approval of Covaxin and its approval for kids.

In general, the public sentiments expressed on Twitter, mostly positive, had an impact on boosting the public receptiveness towards the vaccination program. The government policy makers can derive useful insights from the positive and negative vocabularies of trending topics at the sentiment fluctuations, to take timely steps such as vaccine procurement, vaccine approval, and improve vaccine trials for efficacy evaluation, and above all, communicate all progress achieved to an anxious public who always tend to keep their hopes high even in the face of adversity as evident from the numbers of positive tweets as compared to the numbers of negative tweets.

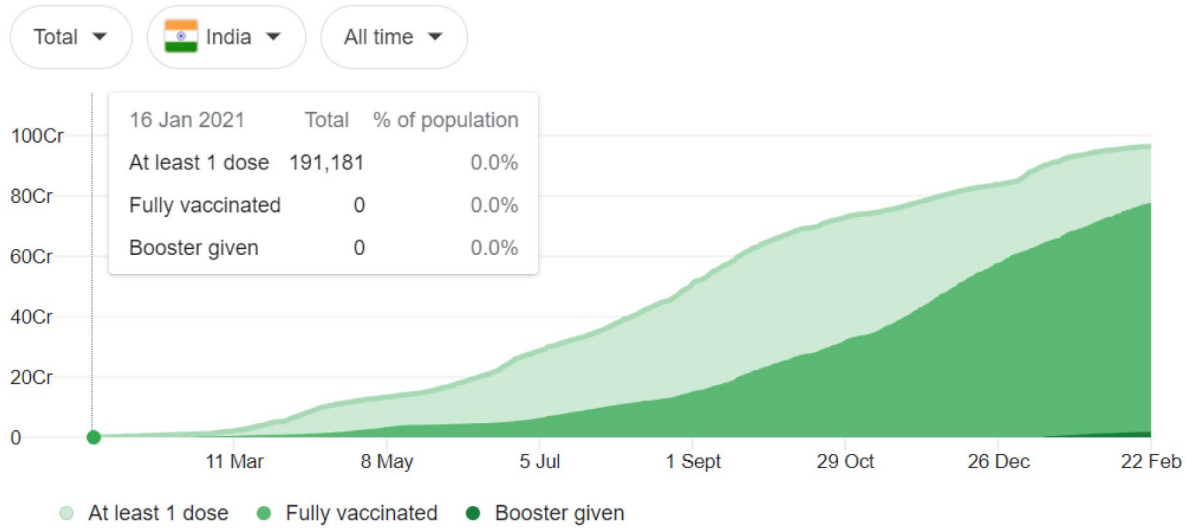
5.5 Correlating sentiments on social media with actual vaccination trends in India

The vaccination trends in India are shown in Fig. 18. This graph agrees with our previous observations. After a slow start in January 2021, the pace of vaccinations started picking up in March, and declined a bit during the second wave, and finally picked up pace in October and November, with the peak of vaccinations occurring around 21 October 2021, when the mark of 1 billion vaccinations (first or second dose) was crossed. The overall public sentiments towards vaccinations can thus be deemed to be positive. Notably, on 11 November 2021, 54.1% of the Indian population had got at least the first dose of vaccine. Recent studies on the public perception towards vaccination has established that there is increasing positivity and trust in regard of Covid-19 vaccinations (Saleh et al. 2023). However, some aspects that inculcate fear and doubts still remain, primarily regarding vaccine safety and the spread of disinformation. It is suggested that social media platforms may be used effectively by governments to disseminate fake information about vaccines. For instance, if negative sentiments spike, authorities can launch targeted social media campaigns or expedite vaccine shipments.

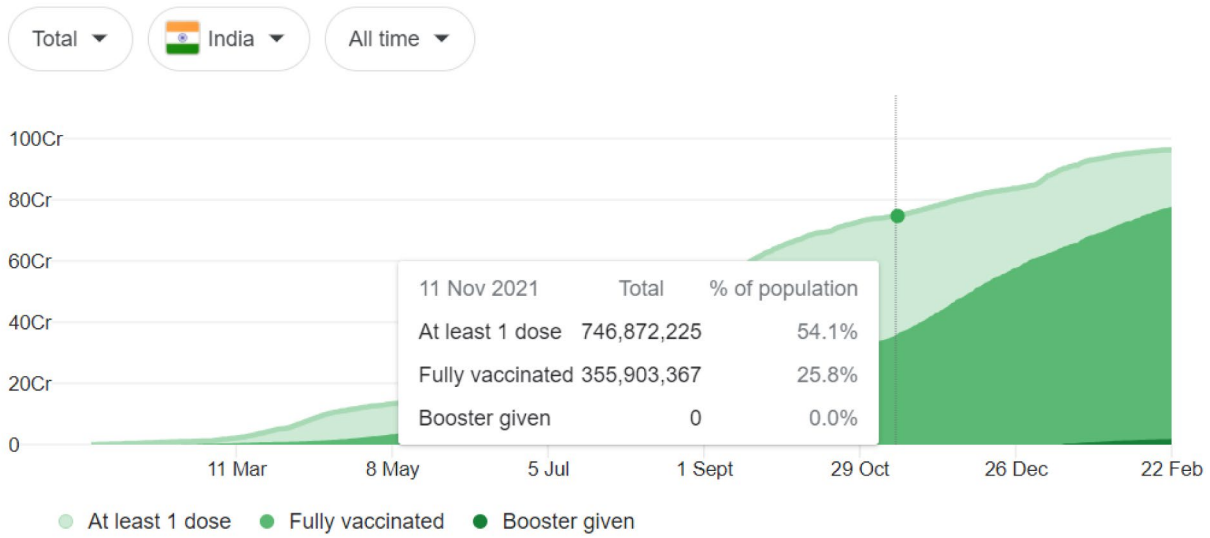
5.6 Challenges and limitations

Our work is not without its limitations. The English tweets that are analyzed for sentiments are not representative of the sentiments of the entire population of a country like India that has many languages and dialects. In the current experiments, the authors segregated and analyzed English tweets due to the ease of application of sentiment analysis tools. Unsupervised sentiment analysis tools like AFINN perform sentiment analysis of English texts only. Even pretrained transformer models like BERTweet and CT-BERT are trained on English tweets. Most languages lack a lexicon connecting words with sentiments or emotions. Though supervised models with neural networks at the core have been proposed for sentiment analysis of Hindi tweets (Gupta et al. 2021), such approaches require a large dataset annotated with sentiment labels for training the models. For handling multi-lingual texts, Google Translate API (https://translate.google.com) is often integrated with sentiment analysis tools attuned to English texts which may not lead to a desirable outcome due to translation-induced errors (Rashid et al. 2024; Mohammad et al. 2016). The machine translation service provided by Google can achieve both forward and backward translation. However, the translated text is prone to lingual errors as proved by researchers in Mondal et al. (2019) who worked on the translation of English to Bengali and vice versa. Future research in the development of multilingual sentiment analysis systems based on multilingual transformer models such as mBERT, XLM, and XLM-R (Bhattacharjee et al. 2024; Doddapaneni et al. 2021), or use of efficient translation APIs such as Yandex translation API (Miller 2023) would aid in effective translation of social media posts to Indic languages like Hindi. Additionally, the use of word embeddings based on FastText encoding of the n-gram components of a word makes it possible for pretrained models like ALBERT to handle the morphological variations in the multi-lingual text in Indian tweets in a more effective manner (Kakwani et al. 2020).

Another limitation is the algorithm-induced biases that contribute to the misinterpretation of sentiments in vaccine-related discussions. Unsupervised models like AFINN and VADER do not require training, are not affected by class imbalances, and are relatively fast and easy to execute



(a)



(b)

Fig. 18 Vaccination trends [data provided by Ministry of health and family welfare, Government of India]. Details provided for **a** vaccination data on 16 Jan 2021, **b** vaccination data on 11 Nov 2021

algorithms. However, they do have some limitations. A manual scrutiny of the sentiments predicted by AFINN revealed that the use of certain words that were semantically ambiguous, sometimes, contributed to a wrong prediction. For instance, it was found that the word "positive" that was used to indicate a corona-positive lab result resulted in the tweet being labeled as positive by AFINN, while it actually indicated to a negative sentiment. As an example, the tweet "*New #COVID19 strain: 20 #UK returnees to #India test positive for the new variant.*" was incorrectly labeled as positive (score of + 2) by AFINN. The misinterpretation of "positive" and "negative" was responsible for most of the mis-classifications. Similarly, some tweets that contained the word "emergency" in the context of the emergency use authorization of a vaccine was mis-interpreted by AFINN as a negative sentiment. One instance is the tweet "*#WHO*

Grants Pfizer-BioNTech #Coronavirusvaccine For Emergency Use #PfizerBioNTech" which was wrongly predicted as negative (score of -1) by AFINN. An interesting conflict of predictions was observed when on the positive occasion of WHO granting approval for emergency use of vaccines, one associated tweet "*WHO Grants 'Emergency Validation' For Pfizer-BioNTech Covid Vaccine, Paving Way For Global Access*" was classified as negative by VADER and positive by AFINN. Another tweet referring to the same event "*The @WHO World Health Organization has approved the #PfizerBioNTech #vaccine for emergency use, paving the way for lower and middle-income countries to begin immunizing their populations against #Covid*" was classified as positive by VADER, though it was classified as neutral by AFINN due to the occurrence of both positive (approved) and negative words (emergency) in the same tweet. This establishes that AFINN can derive the emotional tone of the text in a more effective manner than VADER, agreeing with the numerical quantification of results on the annotated subset of tweets in Fig. 4.

A small rule-based post-processing step may be added to correct such semantic errors. Examples of rules that may be integrated with lexicon-based methods are:

(i) The co-occurrence of words "test" and "positive" in the tweet would result in the label being modified as "negative".
(ii) The co-occurrence of words "test" and "negative" in the tweet would result in the label being modified as "positive".
(iii) The co-occurrence of words"approve", "emergency" and "use" in the tweet would result in the label being modified as "positive".

The performance of BERTweet and CT-BERT depend, to a large extent, on the size of the annotated training set and whether its class distribution is balanced or not. BERTweet and CT-BERT are supervised models pre-trained on COVID-19 text, and hence they are able to distinguish phrases specific to Covid-19 that are present in the limited training set. However, these models are sensitive to the insufficiency of training data and an imbalanced class distribution which are common issues with pandemic-related tweets.

Additionally, there is a lack of annotated tweets for Covid-19 vaccination based discussions on Twitter that makes supervised sentiment analysis a difficult task. A small annotated version of the dataset containing 6,000 tweets is available for supervised classification (https://www.kaggle.com, gpreda, all-covid19-vaccines-tweets [Accessed last on 31st Dec 2021]. 2021). However, since we have concentrated on a larger duration for the analysis, covering a larger number of tweets (44,130), therefore we decided to go for unsupervised sentiment analysis to detect the sentiment peaks. We chose AFINN since it achieved higher performance than VADER or TextBlob. AFINN has been used by several researchers in the past to gain useful insights into the public opinion (Nielsen 2011; Marcec and Likic 2022), and we took the cue from these previous work for our experiments as well.

6 Conclusion

The vaccination drive for Covid-19 pandemic is underway worldwide, with 67% of the world's population having received both primary doses of the vaccine as on 31st December 2023.[2] Initial public skepticism has given way, over time, to a more positive outlook towards vaccination which is considered as the one of the most effective means for achieving immunity. However, vaccinating the complete population in under-developed and developing countries is a dream yet to be realized. In the loom of the uncertainty over the vaccination process, government policies, availability of vaccines, side-effects of vaccines, and the emergence of new mutants of the Coronavirus, people tend to vent their frustration, fears, hopes, and opinion on social media platforms. This paper evaluates the vaccine-related tweets posted on Twitter from users in India in the critical period of 12 December 2020 to 11 November 2021 as India battled and recovered from the deadly second wave of the Coronavirus, while the vast population of 1.3 billion Indian citizens were in the process of getting convinced on getting vaccinated. We trace the public sentiments in this duration through longitudinal sentiment graphs plotted on a day-wise basis using AFINN which is an unsupervised sentiment analysis tool (and cross-verified using VADER, and pre-trained transformer models BERTweet and CT-BERT), and find the positive and negative vocabularies topics of discussion on the dates which led to the sentiment fluctuations using LDA and salient term frequency analysis. Our findings indicate that, overall, the outlook of the users is positive towards the vaccination program, as indicated by the positivity curve plotted over time. Eight time points along the longitudinal sentiment graphs indicate sentiment fluctuations corresponding to peaks in the positive and negative sentiments plotted on a day-wise basis. The main topics of discussion detected at the sentiment fluctuations using LDA are summarized as: (1) the Prime Minister takes a dose of the vaccine (2) approval for emergency use of Sputnik-V vaccine (4) concerns about the availability of vaccines and delay in getting approval for Covaxin, the domestic vaccine, from US-based FDA (5) euphoria about WHO approval of Covaxin and its approval for kids. The public response got warmer and more

[2] https://data.who.int/dashboards/covid19/vaccines

and more positive in October–November 2021 by when a majority of the country's population had received the first dose of the vaccine as the country progressively recovered from the second Covid-19 wave. The case study conducted in this research is unique in the following aspects: (1) nearly a year-long data, (2) major national events linked with sentiment fluctuations, and (3) combination of sentiment modeling and LDA for topic analysis. AFINN analysis revealed that 51.38% of the 44,130 tweets were neutral, while 38.84% were positive, and 9.78% negative.

These findings could guide policy makers in gathering intelligence on misinformation and associated sentiments, and planning counter-measures to combat anti-vaccine campaigns. Real-time monitoring of negative peaks can prompt immediate governmental responses, communication strategies, or logistic adjustments. Our work suggests a strategy which would enable early detection of the loss of confidence in vaccines, so that suitable measures such as imported vaccine procurement, domestic vaccine approval, and imported and domestic vaccine efficacy trials be conducted at a faster pace in order to restore public optimism. This study informs future strategies to counter vaccine hesitancy through targeted communication aimed at vulnerable groups that experience high anxiety and psychosocial burden during the pandemic. The limitations of this study include exclusion of non-English tweets, lack of annotated data which constrains the use of supervised models, and skewed or imbalanced data in which negative tweets are significantly less than positive or neutral tweets. As the future scope of this work, we suggest subjectivity analysis of tweets (Yaqub et al. 2020) as an auxiliary step to sentiment analysis, since it would reflect the beliefs, emotions and speculations of the users, in addition to the polarity detection achieved by sentiment analysis tools. This would aid in accessing the emotional states of users, and the public in general, facilitating improved decision-making based on opinions expressed on the Web. The results could also be enhanced by using a Covid-19 based text summarization model (Karotia and Susan 2023) in conjunction with sentiment analysis for effective opinion summarization and topic analysis from social media posts. Other possible future extensions include using domain-adaptive pretraining on larger vaccine-labeled corpora, cross-lingual sentiment detection, advanced stance detection, or real-time monitoring dashboards.

Author contributions S.S wrote the main manuscript and supervised the research. A.B conducted the primary investigation. All authors were involved in the conceptualization, methodology, data curation, and software validation.

Funding No funding is associated with this research.

Data availability The data is made available online at: https://github.com/Ace117MC/Topic-Modeling-of-Social-Media-Posts-at-Sentiment-Fluctuations-in-Longitudinal-Sentiment-Graphs.

Code availability The code is made available online at: https://github.com/Ace117MC/Topic-Modeling-of-Social-Media-Posts-at-Sentiment-Fluctuations-in-Longitudinal-Sentiment-Graphs [Python, ver. 3.8].

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

References

Ahammad T (2024) Identifying hidden patterns of fake COVID-19 news: An in-depth sentiment analysis and topic modeling approach. Natural Language Process J 6:100053

Bansal A, Choudhry A, Sharma A and Susan S (2023 Adaptation of domain-specific transformer models with text oversampling for sentiment analysis of social media posts on Covid-19 vaccines. Comput Sci, 24(2)

Bansal A, Choudhry A, Sharma A, Susan S (2025) Sentiment and Word Cloud Analysis of Tweets Related to COVID-19 Vaccines before, during, and after the Second Wave in India. In: Emerging Topics in Pattern Recognition and Artificial Intelligence (pp. 229–257)

Bansal A, Susan S, Choudhry A and Sharma A (2022). Covid-19 vaccine sentiment analysis during second wave in india by transfer learning using XLNet. In: pattern recognition and artificial intelligence: third international conference, ICPRAI 2022, Paris, France, June 1–3, 2022, Proceedings, Part II (pp. 443-454). Cham: Springer International Publishing

Batra R, Imran AS, Kastrati Z, Ghafoor A, Daudpota SM, Shaikh S (2021) Evaluating polarity trend amidst the coronavirus crisis in peoples' attitudes toward the vaccination drive. Sustainability 13(10):5344

Behl S, Rao A, Aggarwal S, Chadha S, Pannu HS (2021) Twitter for disaster relief through sentiment analysis for COVID-19 and natural hazard crises. Int J Disaster Risk Reduction 55:102101

Bhattacharjee S, Gain B, Ekbal A (2024) Domain dynamics: evaluating large language models in english-hindi translation. In: proceedings of the ninth conference on machine translation (pp. 341–354)

Boucher JC, Kim SY, Jessiman-Perreault G, Edwards J, Smith H, Frenette N, Badami A, Scott LA (2023) HPV vaccine narratives on Twitter during the COVID-19 pandemic: a social network, thematic, and sentiment analysis. BMC Public Health 23(1):694

Brazil to suspend Indian COVID-19 vaccine deal as graft claims probed | Reuters (Accessed last on 31 Dec 2021)

Catelli R, Pelosi S, Comito C, Pizzuti C, Esposito M (2023) Lexicon-based sentiment analysis to detect opinions and attitude

towards COVID-19 vaccines on Twitter in Italy. Comput Biol Med 158:106876
Chakraborty C, Bhattacharya M, Dhama K (2023) SARS-CoV-2 vaccines, vaccine development technologies, and significant efforts in vaccine development during the pandemic: the lessons learned might help to fight against the next pandemic. Vaccines 11(3):682
Choudhry A, Susan S, Bansal A, Sharma A, (2022) TLMOTE: A topic-based language modelling approach for text oversampling. In: The International FLAIRS Conference Proceedings (Vol. 35)
Coronavirus | Updated data from Covaxin phase 3 trial shows 78% efficacy - The Hindu [Accessed last on 31st Dec 2021]
Covaxin gets emergency use approval for kids aged 2–18 years - Coronavirus Outbreak News (indiatoday.in) [Accessed last on 31st Dec 2021]
Doddapaneni S, Ramesh G, Khapra MM, Kunchukuttan A, Kumar P (2021) A primer on pretrained multilingual language models. arXiv preprint arXiv:2107.00676
Dubey AD (2021) Public sentiment analysis of covid-19 vaccination drive in India. Available at SSRN 3772401
Expert panel approves Russia's Sputnik V COVID-19 vaccine in India - The Hindu [Accessed last on 31st Dec 2021]
Garner J, Crossley S, Kyle K (2019) N-gram measures and L2 writing proficiency. System 80:176–187
Ghosh M, Gupta K and Susan S (2021) Aspect-based unsupervised negative sentiment analysis. In: Intelligent Data Communication Technologies and Internet of Things: Proceedings of ICICI 2020 (pp. 335–344). Springer Singapore
Gupta V, Jain N, Shubham S, Madan A, Chaudhary A, Xin Q (2021) Toward integrated CNN-based sentiment analysis of tweets for scarce-resource language—Hindi. Trans Asian Low-Resourc Language Inf Process 20(5):1–23
https://www.kaggle.com/gpreda/all-covid19-vaccines-tweets [Accessed last on 31st Dec 2021]
https://www.thehindu.com/news/national/pm-modi-takes-first-jab-of-covaxin-against-covid-19/article33958677.ece [Accessed last on 31st Dec 2021]
Hu T, Wang S, Luo W, Zhang M, Huang X, Yan Y, Liu R, Ly K, Kacker V, She B, Li Z (2021) Revealing public opinion towards COVID-19 vaccines with Twitter data in the United States: spatiotemporal perspective. J Med Internet Res 23(9):e30854
Hussain A, Tahir A, Hussain Z, Sheikh Z, Gogate M, Dashtipour K, Ali A, Sheikh A (2021) Artificial intelligence–enabled analysis of public attitudes on facebook and twitter toward covid-19 vaccines in the united kingdom and the united states: observational study. J Med Internet Res 23(4):e26627
Hutto C and Gilbert E (2014) Vader: a parsimonious rule-based model for sentiment analysis of social media text. In: Proceedings of the international AAAI conference on web and social media (Vol. 8, No. 1, pp. 216–225)
India's Covaxin effectively neutralises Delta variant of Covid, says U.S.' health research institute - The Hindu [Accessed last on 31st Dec 2021]
Jain N, Malviya P, Singh P, Mukherjee S (2021) Twitter mediated Sociopolitical communication during the COVID-19 pandemic crisis in India. Front Psychol 12:784907
Jayasurya GG, Kumar S, Singh BK, Kumar V (2021) Analysis of public sentiment on COVID-19 vaccination using twitter. IEEE Trans Comput Soc Syst 9(4):1101–1111
Kakwani D, Kunchukuttan A, Golla S, Gokul NC, Bhattacharyya A, Khapra MM, Kumar P (2020) IndicNLPSuite: Monolingual corpora, evaluation benchmarks and pre-trained multilingual language models for Indian languages. In: Findings of the Association for Computational Linguistics: EMNLP 2020 (pp. 4948–4961)
Karotia A, Susan S (2023) CovSumm: an unsupervised transformer-cum-graph-based hybrid document summarization model for CORD-19. *The Journal of Supercomputing*, pp.1–23
Küçük D and Arıcı N (2022) Sentiment analysis and stance detection in Turkish tweets about COVID-19 vaccination. In: handbook of research on opinion mining and text analytics on literary works and social media (pp. 371–387). IGI Global
Kwok SWH, Vadde SK, Wang G (2021) Tweet topics and sentiments relating to COVID-19 vaccination among Australian Twitter users: machine learning analysis. J Med Internet Res 23(5):e26953
Lin BS (2023) Visualizing change and correlation of topics with LDA and agglomerative clustering on COVID-19 vaccine tweets. IEEE Access 11:51647–51656
Lindelöf G, Aledavood T, Keller B (2023) Dynamics of the negative discourse toward COVID-19 vaccines: topic modeling study and an annotated data set of Twitter posts. J Med Internet Res 25(1):e41319
Loria S, Keen P, Honnibal M, Yankovsky R, Karesh D, Dempsey E (2014) Textblob: simplified text processing. Secondary TextBlob Simplified Text Process 3:2014
Lyu JC, Han EL, Luli GK (2021) COVID-19 vaccine–related discussion on Twitter: topic modeling and sentiment analysis. J Med Internet Res 23(6):e24435
Marcec R, Likic R (2022) Using twitter for sentiment analysis towards AstraZeneca/Oxford, Pfizer/BioNTech and Moderna COVID-19 vaccines. Postgrad Med J 98(1161):544–550
Martínez RY, Blanco G, Lourenço A (2023) Spanish Corpora of tweets about COVID-19 vaccination for automatic stance detection. Inf Process Manage 60(3):103294
Melton CA, Olusanya OA, Ammar N, Shaban-Nejad A (2021) Public sentiment analysis and topic modeling regarding COVID-19 vaccines on the Reddit social media platform: a call to action for strengthening vaccine confidence. J Infect Public Health 14(10):1505–1512
Mendon S, Dutta P, Behl A, Lessmann S (2021) A hybrid approach of machine learning and lexicons to sentiment analysis: Enhanced insights from twitter data of natural disasters. Inf Syst Front 23:1145–1168
Miller MD (2023) Translation Tools and Techniques. In Discovering Hidden Gems in Foreign Languages, pp. 173–225. Cham: Springer International Publishing
Mirza S, Arvinden VR, Rophina M, Bhawalkar J, Khan U, Chothani B, Singh S, Sharma T, Dwivedi A, Pandey E, Garg S (2024) Impact of COVID-19 outbreak on healthcare workers in a Tertiary Healthcare Center in India: a cross sectional study. Sci Rep 14(1):1504
Mohammad SM, Salameh M, Kiritchenko S (2016) How translation alters sentiment. J Artific Intell Res 55:95–130
Mondal H, Mondal S, Mondal S (2019) Feasibility of using "Google Translate" in adaptation of survey questionnaire from English to Bengali: a pilot study. Indian J Soc Psychiatry 35(2):119–124
Müller M, Salathé M, Kummervold PE (2023) Covid-twitter-bert: a natural language processing model to analyse covid-19 content on twitter. Frontiers Artific Intell 6:1023281
Nguyen T, Phung D, Adams B, Venkatesh S (2013) Event extraction using behaviors of sentiment signals and burst structure in social media. Knowl Inf Syst 37:279–304
Nguyen DQ, Vu T and Nguyen AT (2020) BERTweet: A pre-trained language model for English Tweets. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations (pp. 9–14)
Nielsen FÅ (2011. A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. In: 1st Workshop on Making Sense of Microposts: Big things come in small packages (pp. 93–98).

Niu Q, Liu J, Kato M, Shinohara Y, Matsumura N, Aoyama T, Nagai-Tanima M (2022) Public opinion and sentiment before and at the beginning of COVID-19 vaccinations in Japan: twitter analysis. JMIR Infodemiology 2(1):e32335

Now, Pfizer, Moderna refuse to sell COVID-19 vaccines to Delhi government directly- The New Indian Express [Accessed last on 31st Dec 2021]

Panichella A (2021) A systematic comparison of search-based approaches for LDA hyperparameter tuning. Inf Softw Technol 130:106411

Qorib M, Oladunni T, Denis M, Ososanya E, Cotae P (2023) Covid-19 vaccine hesitancy: Text mining, sentiment analysis and machine learning on COVID-19 vaccination Twitter dataset. Expert Syst Appl 212:118715

Ragini JR, Anand PR, Bhaskar V (2018) Big data analytics for disaster response and recovery through sentiment analysis. Int J Inf Manage 42:13–24

Rashid RA, Ab Hamid SH, Fahmi F (2024) A Multilingual corpus for panic and worry in code-mixed tweets by VADER sentiment analysis. J Appl Data Sci 5(3):1038–1051

Saleh SN, McDonald SA, Basit MA, Kumar S, Arasaratnam RJ, Perl TM, Lehmann CU, Medford RJ (2023) Public perception of COVID-19 vaccines through analysis of Twitter content and users. Vaccine 41(33):4844–4853

Setback for Covaxin, US FDA rejects emergency use nod, suggests 'full approval' route - Coronavirus Outbreak News (indiatoday.in) [Accessed last on 31st Dec 2021]

Sharma A, Susan S, Bansal A, Choudhry A (2022) Dynamic topic modeling of covid-19 vaccine-related tweets. In: 2022 the 5th international conference on data storage and data engineering (pp. 79–84)

Sharma M, Choudhary G, Susan S (2023) Resume Classification using Elite Bag-of-Words Approach. In: 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 1409–1413). IEEE

Sicilia-Garcia EI, Ming J and Smith FJ (2002) Extension of Zipf's law to words and phrases. In: COLING 2002: The 19th International Conference on Computational Linguistics

Susan S, Keshari J (2019) Finding significant keywords for document databases by two-phase maximum entropy partitioning. Pattern Recogn Lett 125:195–205

Umair A, Masciari E, Ullah MH (2023) Vaccine sentiment analysis using BERT+ NBSVM and geo-spatial approaches. J Supercomput 79(15):17355–17385

Valdez D, Ten Thij M, Bathina K, Rutter LA, Bollen J (2020) Social media insights into US mental health during the COVID-19 pandemic: Longitudinal analysis of Twitter data. J Med Internet Res 22(12):e21418

Vashishtha S and Susan S (2019) Fuzzy logic based dynamic plotting of mood swings from tweets. In: Innovations in Bio-Inspired Computing and Applications: Proceedings of the 9th International Conference on Innovations in Bio-Inspired Computing and Applications (IBICA 2018) held in Kochi, India during December 17–19, 2018 9 (pp. 129–139). Springer International Publishing

Verma R, Chhabra A, Gupta A (2022) A statistical analysis of tweets on covid-19 vaccine hesitancy utilizing opinion mining: an Indian perspective. Soc Netw Anal Min 13(1):12

WHO issues emergency use listing for eighth COVID-19 vaccine [Accessed last on 31st Dec 2021]

Yanto TA, Octavius GS, Heriyanto RS, Ienawi C, Nisa H, Pasai HE (2021) Psychological factors affecting COVID-19 vaccine acceptance in Indonesia. Egypt J Neurol Psychiatry Neurosurg 57:1–8

Yaqub U, Sharma N, Pabreja R, Chun SA, Atluri V, Vaidya J (2020) Location-based sentiment analyses and visualization of Twitter election data. Digital Govern Res Pract 1(2):1–19

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

2025 3rd IEEE International Conference on Industrial Electronics: Developments & Applications (ICIDeA) | 979-8-3315-3320-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICIDEA64800.2025.10962846

Modified Sigma-Pi neural network for dynamical system identification

Kartik Saini
Department of Electrical Engineering
Delhi Technological University
New Delhi-110042, India
kartiksaini_2k21phdee18@dtu.ac.in

Narendra Kumar
Department of Electrical Engineering
Delhi Technological University
New Delhi-110042, India
narendrakumar@dtu.ac.in

Rajesh Kumar
Department of Electrical Engineering
National Institute of Technology Kurukshetra
Kurukshetra 136119, India
rajeshmahindru23@nitkkr.ac.in

Bharat Bhushan
Department of Electrical Engineering
Delhi Technological University
New Delhi-110042, India
bharat@dce.ac.in

***Abstract*—This study introduces a newly developed Modified Sigma-Pi Neural Network (MSPNN) architecture designed to identify the dynamics of the plant. The proposed structure is an extension of the classical sigma-pi neural network (SPNN) and includes a self-feedback loop at the output node of the output layer. Since sigma-pi artificial neural networks contain additive and multiplicative units in their structure. We apply the back-propagation optimization algorithm to train the proposed MSPNN model. The results of the proposed neural network are compared with preexisting models, such as the Jordan neural network, the feedforward neural network, and the sigma-pi neural network. An analysis of the simulation data indicates that the proposed architecture of MSPNN is more effective than the other models.**

***Index Terms*—Identification, Sigma-Pi neural network, dynamical systems**

I. Introduction

The application of dynamical systems has grown significantly in engineering [1]. Nonlinear systems are challenging to solve, so system identification is required to understand mathematical modeling [2], [3]. The system identification process uses various techniques, including ANN, genetic algorithms, fuzzy systems, etc. ANNs are widely used in many aspects of our lives [4], [5]. They are an essential component of modern technology. Their data processing and analysis application is one of the most exciting areas for neural networks [6]. We refer to a class of ANNs as higher-order neural networks (HONNs). The SPNN model belongs to HONNs and has a rapidly expanding field [7]. The recurrent layer of sigma-pi units has product units. We handle these units using a procedure that involves addition and multiplication. In contrast, the output layer has summing units [8]. The weights of the input to the hidden layer remain fixed at unity. The weights recurrent to the output layer are adjustable [9]. An optimization algorithm obtains these tunable weights during the training process [10]. It is possible to optimize the parameters of the ANN model in different ways. The backpropagation algorithm is the most essential optimization method in machine learning [11]. Using the biological concept of recurrent neural networks, sigma-pi neural networks represent an emerging area of study in machine learning [12]. They are an effective tool for handling numerous data, such as signals, pictures, and time series data [13]. Sigma-Pi neural networks are known for their ability to represent complicated relationships, contextual information, and time series dependencies over long periods [14]. A more significant approximation with a faster convergence property, better storage capacity, and stronger fault tolerance capabilities are just a few of the distinctive qualities of higher-order neural networks (HONNs) compared to ordinary neural networks [15]. HONN is a neural network that offers a stronger classification capacity than linear neurons by providing nonlinear decision limits [16]. HONNs simultaneously combine additive and multiplicative units. As an example of HONN, Shin and Gosh put out the Sigma-pi model in 1991, which has a significantly different design from other ANN models [17], [18]. Since they use additive and multiplicative structures compared to many artificial neuron networks [19]. This article introduces a modified sigma-pi neural network (MSPNN) model, which aims to identify the dynamics of the given dynamical plant. During training, a backpropagation optimization approach is applied to reduce identification errors. Different neural networks, such as the sigma-pi neural network, the Jordan network, and the feedforward network, are used to compare the proposed strategy. I have organized this paper as follows. The MSPNN model is described in section 2 in terms of its architecture and mathematical formulation. In section 3, a learning approach for training parameters is described. The results of the simulation are presented in section 4. The conclusion is presented in section 5, and references are given at the end.

979-8-3315-3320-5/25/$31.00 © 2025 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:25:20 UTC from IEEE Xplore. Restrictions apply.

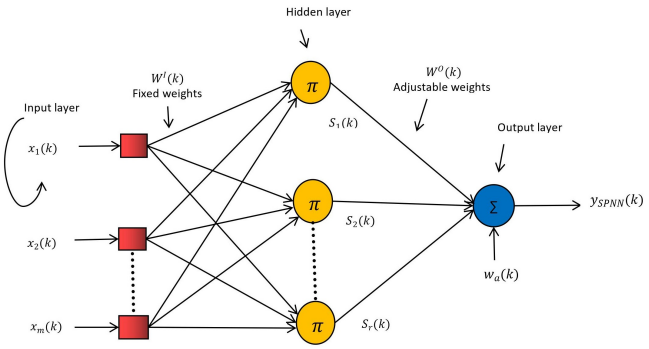


Fig. 1: Sigma-pi neural network model

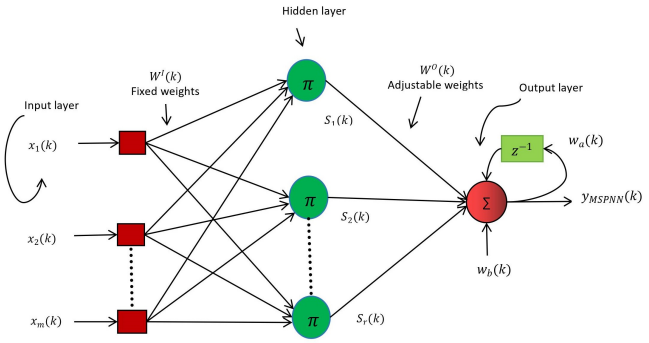


Fig. 2: Modified sigma-pi neural network model

II. Overview of MSPNN and SPNN model

The article introduces the architecture of the SPNN model, comprising three layers: input, hidden, and output, as illustrated in Figure 1. As shown in Figure 2, the proposed MSPNN model is an extension of the standard Sigma-Pi Neural Network model. MSPNN also includes 3 layers: input, hidden, and output. The proposed MSPNN model has a self-feedback loop that sends values later to the output node of the output layer. This is what makes it different from the SPNN model. Here, $X_j(k) = \{x_1(k), x_2(k)...x_m(k)\}$ represents the external input vector m-dimensional. The weights that connect the input layer to the recurrent layer remain fixed to unity. This $w_{qu}^O(k)$ represents the adjustable weights of the uth recurrent layer to the qth output layer. In the architecture of the MSPNN model, an additional self-feedback loop is present with a tunable weight $w_a(k)$ at the output neuron and $w_b(k)$ an adjustable bias weight at the output neuron. When a hidden layer initiates a field, it performs the following calculation:

$$V_{ir}(k) = \prod_{j=1}^{m} x_j(k) \tag{1}$$

The response of hidden neurons is:

$$S_{ir}(k) = V_{ir}(k) \tag{2}$$

The nonlinear activation function (tangent hyperbolic activation function) is applied to each hidden neuron, and then it produces the following output:

$$P\{V_{ir}(k)\} = tanh\{V_{ir}(k)\} \tag{3}$$

output layer performs the result of the output node.

$$V_o(k) = \left\{ \sum_{q=1} P\{V_{ir}(k)\} w_{qu}^O(k) \right\} \tag{4}$$

The $3rd$ layer has a linear activation function, purelin $f(.)$. The proposed model with a linear activation function yields the following output:

$$y_{MSPNN}(k) = f\{V_o(k)\} \tag{5}$$

III. Learning algorithm

A learning approach obtains the parameter update equations. To derive the equations for the weight modifications, we employed the dynamic backpropagation algorithm [20]. It is a method of minimizing an objective function based on gradient descent. We use the chain rule to find the partial derivatives of the error function in relation to the network weights. This method gives us the parameter update equations. During training, the parameters are adjusted to reduce the identification error $e_i(k)$. The objective function of the study is the MSE, or instantaneous mean squared error.

$$E(k) = \frac{1}{2} {e_i}^2(k) \tag{6}$$

$$E(k) = \frac{1}{2} [y(k) - y_{MSPNN}(k)]^2 \tag{7}$$

where $E(k)$ is the mean squared error (MSE), $y(k)$ is the desired output, and $y_{MSPNN}(k)$ is the proposed model's output. For every weight parameter, the updated equations are developed by subtracting the significant component of the objective function from the weight value in the derivative $2nd$ of the objective function. The equation below describes the process of updating the model parameter.

$$W_{new}(k+1) = W(k) - \eta \frac{\partial E(k)}{\partial W(k)} \tag{8}$$

The learning rate, denoted as η, ranges from 0 to 1. The chain rule is used to find the equations for updating the output parameters.

$$\frac{\partial E(k)}{\partial w_{qu}^O(k)} = \frac{\partial E(k)}{\partial y_{MSPNN}(k)} \times \frac{\partial y_{MSPNN}(k)}{\partial V_o(k)} \times \frac{\partial V_o(k)}{\partial w_{qu}^O(k)} \tag{9}$$

$$where \quad \frac{\partial E(k)}{\partial y_{MSPNN}(k)} = -e_i(k), \frac{\partial y_{MSPNN}(k)}{\partial V_o(k)} = 1, \frac{\partial V_o(k)}{\partial w_{qu}^O(k)} = S_{ir}(k) \tag{10}$$

$$\frac{\partial E(k)}{\partial w_{qu}^O(k)} = -e_i(k)\{S_{ir}(k)\} \tag{11}$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:25:20 UTC from IEEE Xplore. Restrictions apply.

TABLE I: Statistical error comparisons after the training

Sr no.	Model	Hidden neurons	Epochs	Samples	Learning rate	MSE	MAE	RMSE
1.	MSPNN	3	600	1000	0.00006	0.00015	0.0024	0.012
2.	SPNN	3	600	1000	0.00006	0.0031	0.0232	0.055
3.	JRNN	3	600	1000	0.00006	0.0085	0.0775	0.092
4.	FFNN	3	600	1000	0.00006	0.0034	0.0630	0.058

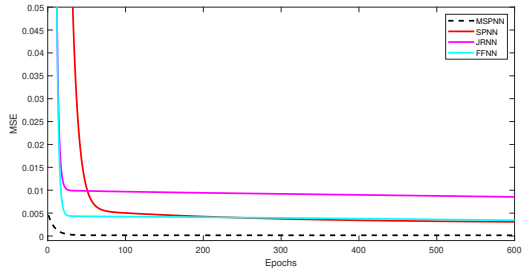


Fig. 3: MSE response obtained with different neural networks

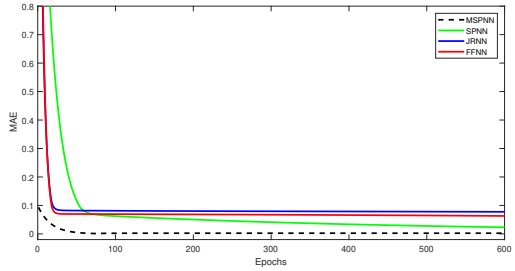


Fig. 4: MAE response obtained with different neural networks

$$w_{qu}^{O}(k+1) = w_{qu}^{O}(k) + \eta e_i(k)\{S_{ir}(k)\} \tag{12}$$

$$\frac{\partial E(k)}{\partial w_a(k)} = \frac{\partial E(k)}{\partial y_{MSPNN}(k)} \times \frac{\partial y_{MSPNN}(k)}{\partial V_o(k)} \times \frac{\partial V_o(k)}{\partial w_a(k)} \tag{13}$$

$$\frac{\partial E(k)}{\partial w_a(k)} = -e_i(k) y_{MSPNN}(k-1) \tag{14}$$

$$w_a(k+1) = w_a(k) + \eta e_i(k)\{y_{MSPNN}(k-1)\} \tag{15}$$

$$w_b(k+1) = w_b(k) + \eta e_i(k)(-1) \tag{16}$$

similarly, for bias weight

$$\frac{\partial E(k)}{\partial w_b(k)} = \frac{\partial E(k)}{\partial y_{MSPNN}(k)} \times \frac{\partial y_{MSPNN}(k)}{\partial V_o(k)} \times \frac{\partial V_o(k)}{\partial w_b(k)} \tag{17}$$

$$\frac{\partial E(k)}{\partial w_b(k)} = -e_i(k)(-1) \tag{18}$$

IV. Simulation Outcomes

One non-linear system is used in the simulation study to compare how well the MSPNN, SPNN, JRNN, and FFNN models work. The initial values of the parameters are selected at random, and the starting learning rate is 0.00006. Based on

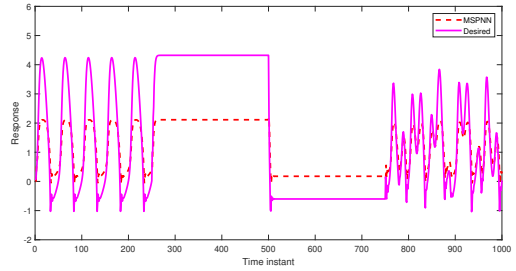


Fig. 5: Validation response obtained from MSPNN

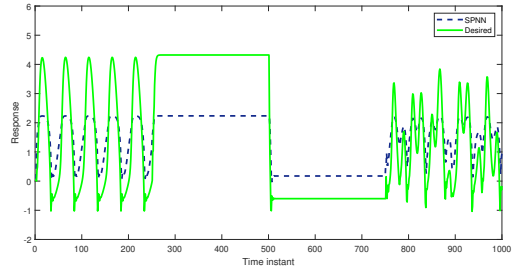


Fig. 6: Validation response obtained from SPNN

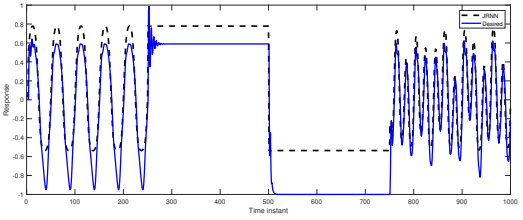


Fig. 7: Validation response obtained from JRNN

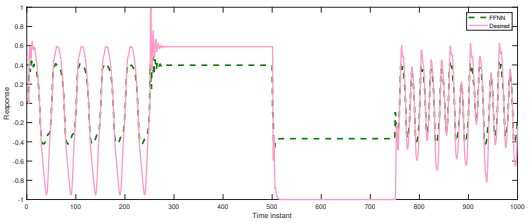


Fig. 8: Validation response obtained from FFNN

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:25:20 UTC from IEEE Xplore. Restrictions apply.

[17], the dynamics of the plant is represented by the non-linear difference equations.

$$y_p(k) = F[y_p(k-1), y_p(k-2)] + r_e(k) \tag{19}$$

where

$$F = \frac{(y_p(k-1)y_p(k-2)(y_p(k-1)+2.5)}{1+y_P^2(k-1)+y_P^2(k-2)} \tag{20}$$

F is considered unknown but differentiable and $r_e(k)$ is the externally applied input. A total of 600 iterations and 1000 samples are utilized to train the system's parameters at an initial learning rate of 0.00006. The external input used is $r_e(k) = 0.3\cos(\pi k/45)$ for $0 < k < 650$ and $r_e(k) = 0.3\sin(\pi k/50) + 0.1\cos(\pi k/45)$ for $650 < k < 1000$. We can see the mean squared error (MSE) and mean average error (MAE) graphs after the training process in Fig. 3 and Fig. 4. Table 1 presents the performance of the statistical error after training. Compared to the other models, the suggested MSPNN model has lower MSE, MAE, and root mean squared error (RMSE) values. Once training is complete, we apply an additional external input to assess the model's capacity for generalization. Here is the updated external input:

$$r_e(k) = \begin{cases} \sin\left(\frac{\pi k}{25}\right) & \text{if } k \leq 250 \\ 1 & \text{if } 250 < k \leq 500 \\ -1 & \text{if } 500 < k \leq 750 \\ 0.3\sin\left(\frac{\pi k}{25}\right) + 0.1\sin\left(\frac{\pi k}{32}\right) & \\ +0.6\sin\left(\frac{\pi k}{10}\right) & \text{if } 750 < k \leq 1000 \end{cases} \tag{21}$$

Figures 5-8 show the validation responses derived from the MSPNN, SPNN, JRNN, and FFNN models. The visual comparison of each model's ability to reveal the dynamics of the intended system throughout the validation stage is shown in these graphics. The MSPNN's validation response demonstrates a closer alignment with the actual system outputs, indicating its superior accuracy and ability to handle complex nonlinear behaviors.

V. Conclusion

This article proposes the MSPNN model, a modified sigma-pi neural network. Dynamical systems with non-linear dynamics can be identified with it. The suggested neural network is a modified version of the current SPNN model. It has a self-feedback loop whose weight can be changed. The plant's previous value and the external input's current value have been the only two inputs used to execute the suggested model. We apply the backpropagation optimization algorithm to train the adjustable weight parameters. This article compares the proposed MSPNN model with several other models (such as SPNN, JRNN, and FFNN) after training. The simulation analysis results are shown in Table 1. They show that the proposed network works more efficiently and has lower error rates (MSE and MAE) than the others. The enhanced functionality of the proposed model is attributed to its architectural design, which includes a self-feedback loop in the output layer, which improves the network's ability to handle temporal dependencies and nonlinear dynamics while reducing approximation errors.

References

[1] F. Ye, J. Yang, A deep neural network model for speaker identification, Applied Sciences 11 (8) (2021) 3603.

[2] S. Mao, E. Sejdić, A review of recurrent neural network-based methods in computational physiology, IEEE Transactions on Neural Networks and Learning Systems (2022).

[3] K. Saini, N. Kumar, B. Bhushan, R. Kumar, Artificial neural network-based adaptive control for nonlinear dynamical systems, International Journal of Adaptive Control and Signal Processing 38 (8) (2024) 2693–2715.

[4] R. Kumar, A lyapunov-stability-based context-layered recurrent pi-sigma neural network for the identification of nonlinear systems, Applied Soft Computing 122 (2022) 108836.

[5] R. Jullapak, A. Thammano, Backpropagation neural network with adaptive learning rate for classification, in: Advances in Natural Computation, Fuzzy Systems and Knowledge Discovery: Proceedings of the ICNC-FSKD 2022, Springer, 2023, pp. 493–499.

[6] D. W. Kim, M. S. Kim, J. Lee, P. Park, Adaptive learning-rate backpropagation neural network algorithm based on the minimization of mean-square deviation for impulsive noises, IEEE Access 8 (2020) 98018–98026.

[7] R. Dash, A. Routray, R. Dash, R. Rautray, Designing an efficient predictor model using psnn and crow search based optimization technique for gold price prediction, Intelligent Decision Technologies 15 (2) (2021) 281–289.

[8] N. A. Husaini, R. Ghazali, N. M. Nawi, L. H. Ismail, The effect of network parameters on pi-sigma neural network for temperature forecasting, in: International Journal of Modern Physics: Conference Series, Vol. 9, World Scientific, 2012, pp. 440–447.

[9] H. V. Dudukcu, M. Taskiran, Z. G. C. Taskiran, T. Yildirim, Temporal convolutional networks with rnn approach for chaotic time series prediction, Applied Soft Computing 133 (2023) 109945.

[10] Q. Fan, Q. Kang, J. M. Zurada, I. Life Fellow, Convergence analysis for sigma-pi-sigma neural network based on some relaxed conditions, Information Sciences 585 (2022) 70–88.

[11] M. F. Ab Aziz, S. A. Mostafa, C. F. M. Foozy, M. A. Mohammed, M. Elhoseny, A. Z. Abualkishik, Integrating elman recurrent neural network with particle swarm optimization algorithms for an improved hybrid training of multidisciplinary datasets, Expert Systems with Applications 183 (2021) 115441.

[12] M. Impraimakis, A convolutional neural network deep learning method for model class selection, Earthquake Engineering & Structural Dynamics 53 (2) (2024) 784–814.

[13] R. Zhilov, Data clustering using sigma-pi neural network, in: 2023 International Russian Automation Conference (RusAutoCon), 2023, pp. 237–242.

[14] S. Behera, A. V. S. Pavan Kumar, S. C. Nayak, Improved firefly based pi-sigma neural network for gold price prediction, in: 2023 OITS International Conference on Information Technology (OCIT), 2023, pp. 202–207.

[15] K. Pandey, S. Bhanacharjee, S. Lau, M. Tushir, A comparative study of fuzzy systems and neural networks for system modeling and identification, in: 2018 2nd IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), IEEE, 2018, pp. 876–880.

[16] L. Samir, G. Said, K. Nora, S. Youcef, Improved pi-sigma neural network for nonlinear system identification, in: 2017 5th International Conference on Electrical Engineering-Boumerdes (ICEE-B), IEEE, 2017, pp. 1–5.

[17] S. N. Kumpati, P. Kannan, et al., Identification and control of dynamical systems using neural networks, IEEE Transactions on neural networks 1 (1) (1990) 4–27.

[18] G. Quaranta, W. Lacarbonara, S. F. Masri, A review on computational intelligence for identification of nonlinear dynamical systems, Nonlinear Dynamics 99 (2) (2020) 1709–1761.

[19] C.-F. Juang, C.-H. Hsu, Structure and parameter optimization of fnns using multi-objective aco for control and prediction, in: 2014 ieee international conference on fuzzy systems (fuzz-ieee), IEEE, 2014, pp. 928–933.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:25:20 UTC from IEEE Xplore. Restrictions apply.

[20] K. Saini, N. Kumar, R. Kumar, B. Bhushan, Recurrent neural network for the identification of nonlinear dynamical systems: A comparative study, in: International Online Conference on Smart Grid Energy Systems and Control, Springer, 2023, pp. 345–353.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:25:20 UTC from IEEE Xplore. Restrictions apply.

Optical and Quantum Electronics (2025) 57:270
https://doi.org/10.1007/s11082-025-08123-1

MXene coated concave shaped microchannel PCF SPR biosensor for the detection of HIV and sickle cell anaemia

Preethi Krishnan[1] · **Akash Khamaru**[1] · **Ajeet Kumar**[1]

Received: 11 December 2024 / Accepted: 19 February 2025
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2025

Abstract
The paper presents the analysis and discussions on a Surface Plasmon resonance (SPR) based Photonic Crystal Fiber (PCF) Refractive index (RI) Biosensor, designed specifically to detect two of the most widespread diseases—HIV-caused AIDS and Sickle cell anaemia. The detection is purely dependent on the RI of the affected cells (1.42 for HIV and 1.38 for Sickle cell anaemia), compared to 1.35 for normal cells. The proposed RI biosensor has concave-shaped microchannels on either side of the PCF acting as a channel, where the plasmonic material interacts with the bio-analytes of HIV-infected and sickle cells. Additionally, a layer of MXene, an emerging 2D plasmonic enhancer is coated on the microchannels, to enhance the sensor performance in terms of sensitivity and confinement loss. Moreover, the optimized structure achieved wavelength sensitivities of 10,571.43 nm/RIU and 3500 nm/RIU for the Gold PCF sensor, and 14,142.86 nm/RIU and 4000 nm/RIU for MXene-Gold PCF sensor for HIV-infected cell and Sickle cell anaemia, respectively. Hence, the proposed SPR biosensor using MXene has the potential for the early detection of HIV and Sickle cell anaemia.

Keywords MXene-coated · Sickle Cell Anaemia · HIV · Concave-Shaped Microchannel · Surface Plasmon Resonance · Confinement Loss

1 Introduction

Surface plasmon resonance (SPR) is an extensively preferred optical technique employed in photonic crystal fibers (PCF) which is an evolved form of optical fiber consisting of periodic arrangement of micro structured air holes for enhanced light confinement (Elabdein et al. 2024). SPR technique entails the collective excitation of surface plasmons which

✉ Ajeet Kumar
ajeetdph@dtu.ac.in

Preethi Krishnan
preethikrishnan_23mscphy68@dtu.ac.in

Akash Khamaru
akashkhamaru_23phdap02@dtu.ac.in

[1] Advanced Photonics Simulation Research Laboratory, Department of Applied Physics, Delhi Technological University, Bawana Road, Delhi 110042, India

Published online: 20 April 2025

are electromagnetic waves that propagates parallel to the metal–dielectric interface. The surface plasmons resonate with the light striking on the metal and leads to a dip in the reflected light intensity thereby changing the reflection conditions, making it very quick to respond to any small refractive index (RI) variations in vicinity. Even small traces of bio-analyte could be detected by exploiting this special feature of the SPR technique (Capelli et al. 2023; Divya et al. 2023).

Of late, this technique has been explored widely by integrating it with PCFs and has opened various new gates in the field of biosensing, as it provides a real time and label-free monitoring. PCFs are fibers having perforated air holes along the length and same are worked as cladding and widely used for applications like supercontinuum generation (Khamaru and Kumar 2024a, b; Garg et al. 2024), terahertz communication and sensing (Singh et al. 2024a, b) etc.

SPR based PCF RI biosensors offer a promising diagnostic tool in the field of medicine (Opoku et al. 2023). It has enabled the detection of various diseases like cancer, malaria, tuberculosis and coronavirus (Singh and Prajapati 2023; Meshginqalam 2022; Divya et al. 2024; Gupta et al. 2023). These diseases are examined based on the RI differences in the bio-analytes between the normal blood cell and the infected blood cell, as these biosensors are highly sensitive to RI variations.

The utilization of plasmonic materials like Gold, Silver or Copper facilitates and provide a platform for the occurrence of SPR (Chaudhary et al. 2022; Zhang et al. 2024; Luo et al. 2022). Gold is the often-favoured plasmonic material due to its excellent chemical stability and biocompatibility. Silver and Copper also exhibit certain qualities like high conductivity which make it good plasmonic materials but in view of real-time fabrication, silver and copper suffers from environmental oxidation. The sensitivity can be further enhanced by using certain 2-D materials like Titanium dioxide (TiO_2) (Zhang et al. 2021), Germanium-antimony-telluride ($Ge_2Sb_2Te_5$) (Zhang et al. 2022), Molybdenum Disulphide (MoS_2) (Kumar et al. 2024), black phosphorous (Gupta et al. 2025) which are known as plasmonic enhancers. These materials, when coupled with the plasmonic material such as gold or silver, yield enhanced sensitivity (Zhang et al. 2021). Nonetheless, these 2D materials put forward certain constraints like reduced conductivity, low adsorption stability, poor biocompatibility and high hydrophobicity. However, there is a class of 2D materials called MXene, which has excellent conductivity, flexibility in functionalization, large surface area and hydrophilic in nature, allowing for uniform dispersion of plasmonic material (Mumtaz et al. 2022).

MXenes are a group of 2D materials discovered by Naguib et al. in 2011 (Jeon et al. 2020) consisting of a transition metal carbide or nitride or carbonitride in a honey-comb like structure. MXenes are synthesized from its precursor MAX phase by selective etching using Hydrofluoric acid etchants and delamination. Further, MXene's properties can be altered based on our objective by changing the surface functional groups (Bhargava Reddy et al. 2022). This paper presents a MXene-Gold coated concave shaped microchannel SPR PCF multipurposed biosensor.

In recent years, many works have been published in the PCF-SPR sensing domain. In 2024, a D-shaped dual core SPR PCF sensor was reported for the precise measurement of a broad analyte RI range of 1.22 to 1.36, with a maximum sensitivity of 22,100 nm/RIU (Fei et al. 2024). Another work on an SPR PCF Gold coated sensor was reported in the same year. It has achieved a sensitivity of 14,100 nm/RIU (Melwin and Senthilnathan 2024). Yimin Mao et al. presented a highly sensitive MXene coated SPR PCF RI sensor for cancer detection which demonstrated the sensitivity enhancing potential of MXene. It reported a maximum wavelength sensitivity of 9286 nm/RIU for breast cancer cell (Mao et al.

2024). Various MXene coated SPR PCF sensors were also proposed by researchers like the one reported in 2023 for bio-analyte detection. Farhan Mumtaz et al. presented a novel D-shaped highly sensitive SPR PCF RI sensor exhibiting a very high wavelength sensitivity of 47,260 nm/RIU at an analyte RI of 1.42 (Mumtaz et al. 2022). Although, SPR PCF MXene coated Biosensors for the detection of various diseases have already been reported in available literatures, none has been reported yet for the detection of both HIV and Sickle cell anaemia using SPR PCF Biosensor. Elizabeth Caroline Britto et al. exhibited a paper on photonic crystal-based nanoscale multipurpose biosensor for detection of HIV, anaemia and brain tumor. It achieved a sensitivity of 1034 nm/RIU for HIV and 1071 nm/RIU for anaemia (Britto et al. 2024). So, the proposed design is focused mainly on the detection bio-analytes of Human Immunodeficiency Virus (HIV) causing AIDS and Sickle Cell Anaemia using SPR PCF RI Biosensor.

HIV is an infection that invades our body's immune system. The extreme stage of infection leads to an extremely dangerous condition called Acquired Immuno-Deficiency Syndrome (AIDS). Globally, around 40 million people are currently affected with this infectious disease among which 1.4 million are children. The death toll ranges up to 50 million in the past 40 years (Parums 2024). The high mortality rate is the consequence of the fact that it has no definite cure. However, there exists many treatment approaches like antiretroviral therapy (ART) that has potentially expanded the life span of the infected individuals by reducing the viral load content in the body (Gandhi et al. 2023). But this is only effective in the early stages of infection, which necessitates the early diagnosis of the HIV infection. There are several diagnostic tools for detecting HIV, such as Enzyme-Linked Immunosorbent Assay (ELISA) and rapid antibody tests (Tavakoli et al. 2017). But these techniques pose certain constraints like limited sensitivity (during acute infection stage), need for labelled antibodies and failure to detect HIV subtypes. The proposed SPR PCF RI biosensor appears to be a superior detection technique in comparison with most conventional techniques, as it provides a label-free detection with a potential of achieving lower detection limits giving higher sensitivity. Another benefit of using SPR PCF RI Biosensor includes its broad specificity, which allows for the detection of multiple HIV subtypes (HIV1 and HIV2). HIV1 and HIV2 are the subtypes of HIV. This paper focusses on HIV1 because it is more common and more easily transmitted compared to HIV2, which is rarer. The RI of normal blood cell bears a value of 1.35, while that of HIV1 is 1.42. This increment in the RI of blood cell after infection occurs as a result of viral load accumulation in the cell, leading to the alteration of the cellular composition (altered protein and lipid concentration) (Kumar et al. 2022; El-assar et al. 2023).

Another suitable application for the proposed sensor is the detection of Sickle Cell Anaemia. It is an inherited disorder resulting in abnormal haemoglobin which in turn alters the appearance of the red blood cells to C or sickle shaped cell. The sickle cell leads to constricted blood flow causing further complications like anaemia and infection. The treatment involves only oral medications to reduce pain and further damage of the cell. It doesn't have a definite cure (Tisdale et al. 2020). However, there are several diagnostic techniques like genetic testing, haemoglobin electrophoresis, Complete blood count (CBC) etc. However, SPR PCF RI Biosensor offers much more promising results as it is a label-free technique providing higher sensitivity. The normal cell RI is, again, 1.35. But the RI of sickle cell anaemia is found to be 1.38 (Kumar et al. 2022). The change RI occurs due to the sickling of cells as it causes Haemoglobin S (HbS) polymerization (Kumar et al. 2022).

The paper proposes the design and structural analysis of a MXene coated concave-shaped microchannel SPR PCF sensor for the detection of HIV infected cell and Sickle cell anemia. The proposed sensor offers a high sensitivity of the discussed biological cells.

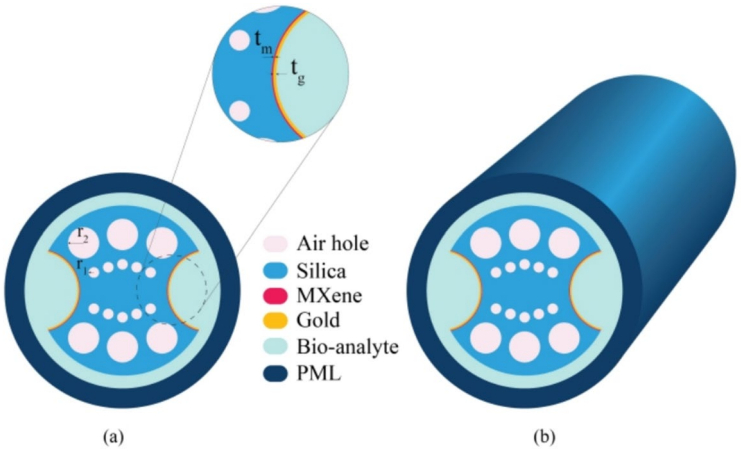


Fig. 1 **a** 2D view of the Microchannel MXene coated sensor, **b** 3D view

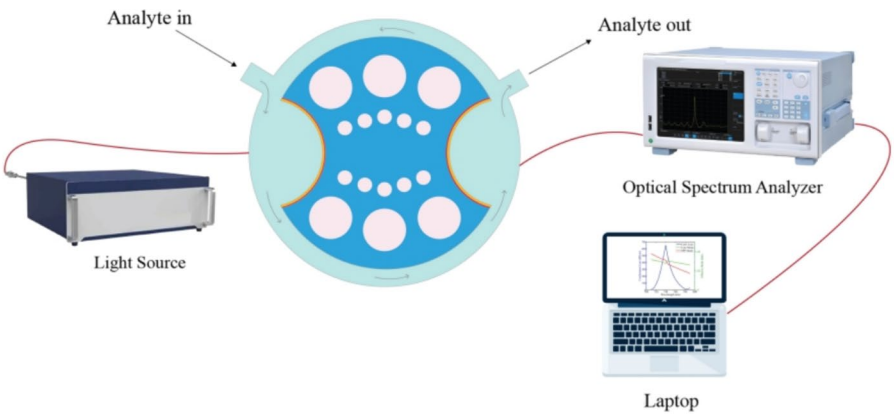


Fig. 2 Proposed experimental arrangement for the presented bio-sensor

Apart from high sensitivity, the designed PCF- SPR sensor is proposed for the for the detection of sickle cell and HIV.

2 Structural design of the proposed concave-shaped microchannel spr pcf sensor

2.1 Structural design

The illustration of the proposed sensor design and the experimental setup is depicted in Fig. 1 and 2. The structural design contains dual sets of airholes arranged in a concentric semi-circular pattern on both the halves of the circular PCF sensor of radius 4200 nm. The radii of the small and big airholes are defined as r_1 and r_2 respectively.

To ensure enhanced interaction of the core mode with plasmonic material, few airholes from both the semicircular concentric sets are eliminated for the incorporation of the concave shaped microchannels on either side of the PCF (Majeed and Ahmad 2024). The arrangement of these airholes is done against a fused silica material background, which forms the core of the PCF. The plasmonic material, gold (t_g) is coated on the concave shaped microchannel on either side of the PCF, which lies in contact with the analyte layer. Additionally, a 14 nm (t_m) layer of MXene (plasmonic enhancer) is sandwiched between silica and gold film of the concave shaped microchannel. This hybrid layer of gold and MXene promotes a more effective coupling and also enhances the field interaction with the analyte. Consequently, this results in an elevated sensitivity with lower detection limits (Mu et al. 2023). The analyte layer has a thickness of 0.5 µm and it contains the bio-analytes of HIV infected cell and Sickle cell anaemia affected blood cells, with a RI of 1.42 and 1.38 respectively. Lastly, a Perfect Matching Layer (PML) of 0.7 µm thickness is added to minimize the electromagnetic wave reflections and for the absorption of radiation energy (Tasnim et al. 2021).

The main core-forming backdrop material is silica, whose RI is given by the wavelength dependent Sellmeier equation,

$$nsi = \sqrt{\left(1 + \frac{\left(B_1 \times \lambda^2\right)}{\lambda^2 - C_1} + \frac{\left(B_2 \times \lambda^2\right)}{\lambda^2 - C_2} + \frac{\left(B_3 \times \lambda^2\right)}{\lambda^2 - C_3}\right)} \tag{1}$$

where $B_1, B_2, B_3, C_1, C_2 and C_3$ are Sellmeier coefficients bearing values of 0.69616300, 0.407942600, 0.897479400, $4.67914826 \times 10^{-3}\mu m^2$, $1.35120631 \times 10^{-2}\mu m^2$, $97.934002\mu m^2$ respectively (Li et al. 2023).

Drude-Lorentz model is used to determine the electrical permittivity of the plasmonic material gold, followed by calculating its square root, resulting in the RI of gold.

$$\varepsilon_{Au} = \varepsilon_{\infty} + \frac{\omega_D^2}{\omega\left(\omega + i\gamma_D\right)} - \frac{\Delta\varepsilon\Omega_L^2}{\left(\omega^2 - \Omega_L^2\right) + i\Gamma_L\omega} \tag{2}$$

where ε_{∞} is given by 5.9673 and ω_D, γ_D and $\Delta\varepsilon$ are the plasma frequency, decaying frequency and weighting factor respectively (Khamaru and Kumar 2024b)

$$n_{Au} = \sqrt{\varepsilon_{Au}} \tag{3}$$

The RI of the 2-D plasmonic enhancer MXene can be determined by fitting the curve using the experimental data of the real and imaginary part of permittivity as a function of wavelength (Dillon et al. 2016). The data specifically represents the permittivity of Ti_3C_2Tx, where T is the surface termination element. Fluorine is taken for this proposed case. We obtain two sets of sixth-order polynomial equations representing the real and imaginary parts of permittivity, ε_1, and ε_2, respectively.

$$\varepsilon_1 = p_1 \times (\lambda^6) + p_2 \times (\lambda^5) + p_3 \times (\lambda^4) + p_4 \times (\lambda^3) + p_5 \times (\lambda^2) + p_6 \times (\lambda) + p_7 \tag{4}$$

$$\varepsilon_2 = q_1 \times (\lambda^6) + q_2 \times (\lambda^5) + q_3 \times (\lambda^4) + q_4 \times (\lambda^3) + q_5 \times (\lambda^2) + q_6 \times (\lambda) + q_7 \tag{5}$$

The real and imaginary part of MXene's RI can be obtained from the following equation (Kumar et al. 2023)

$$n^2 = \frac{\sqrt{\varepsilon_1^2 + \varepsilon_2^2} + \varepsilon_1}{2}, k^2 = \frac{\sqrt{\varepsilon_1^2 - \varepsilon_2^2} - \varepsilon_1}{2} \tag{6}$$

where n and k are the real and imaginary part of the RI. The validity for these equations extends up to around 1770 nm wavelength and is applicable exclusively for 14 nm thickness of MXene.

2.2 Fabrication feasibility

For the practical execution of the suggested sensor various approaches are adopted, first one being the Stack-and-Draw technique for the precise creation of the intended PCF. It is an economical approach involving several steps like preform preparation, consolidation by applying the required temperature and pressure, followed by Drawing process and cooling. This process enables the incorporation of the nanometre ranged airholes into the PCF with ease.

For the homogenous integration of the plasmonic metal layer of gold with the PCF, there exists many methods like Atomic layer deposition (ALD), Sol–Gel method, Physical vapour deposition (PVD), Thermal evaporation and Chemical vapour deposition (CVD). However, CVD is the most commonly used technique for the nano-scale range deposition of gold (Chaudhary et al. 2022).

The plasmonic enhancer MXene is prepared from its precursor MAX phase and it is coupled with gold nanoparticle layer by employing ultrasonication for obtaining uniform integration (Li et al. 2018). This hybrid layer is then attached to the PCF using physical vapour deposition. Since both the plasmonic materials, gold and MXene are biocompatible, the bio-analyte can be passed in through the PCF with ease, facilitating easy analysis. Therefore, the fabrication of the suggested sensor is in accordance with the existing advanced innovations.

3 Simulation results and analysis

3.1 Analysis of the core and SPP mode

The simulation software COMSOL Multiphysics was used for analyzing the sensor performance of the proposed biosensor. Modelling and simulation are done using a numerical approach called Finite Element Method (FEM) using a finer mesh design as shown in Fig. 3.

As demonstrated in the Fig. 4, the phase matching condition is observed when the effective mode indices of both SPP mode (red line) and the core-guided mode (green line) intersects with each other. Here, we observe the dispersion relation corresponding to an analyte RI of 1.38, whose resonant wavelength is observed at 740 nm. Consequently, a sharp confinement loss peak occurs at this resonant point, which corresponds to maximum energy transmission between the core-guided mode and SPP mode in the biosensor. The core mode, SPP mode and SPR mode for gold only probe and MXene-gold probe are displayed in Fig. 5 (a-f). The gold only probe provides maximum loss of 1381.2 dB/cm for analyte of RI 1.38 whereas on incorporating MXene with gold, the maximum loss decreased to

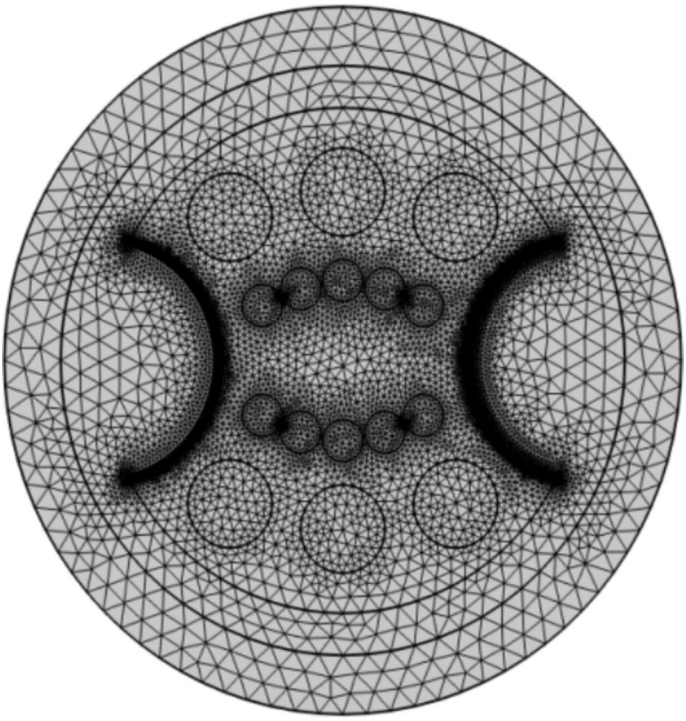


Fig. 3 Meshed view of the proposed design

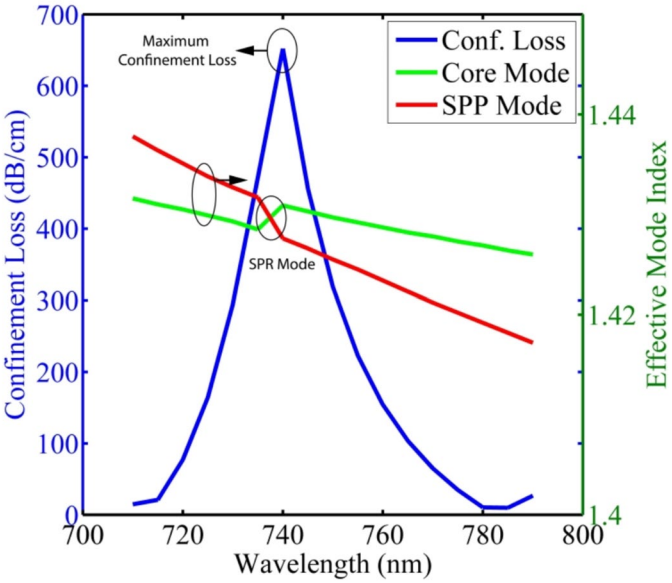


Fig. 4 Dispersion relation of analyte of 1.38 and loss spectrum

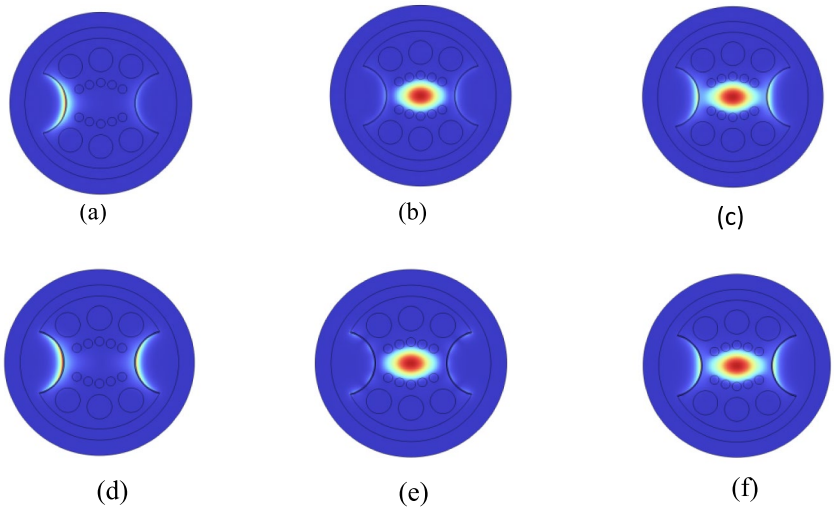


Fig. 5 **a** SPP mode, **b** Eye-shaped Core mode, **c** Coupled mode of Gold only PCF and **d** SPP mode **e** Core mode, **f** Coupled mode of MXene-gold PCF

652.06 dB/cm. Therefore, the MXene-gold probe can be a better substitute for decreasing the loss for lower RI analytes.

3.2 Optimization of the geometrical parameters

The suggested sensor is found to offer promising results in the detection of HIV cell and Sickle cell anaemia. The sensor performance is further amplified by optimization of the structural parameters. Those parameters are optimized which plays an essential role in deciding the sensor's sensitivity and loss spectrum. In this work, three of the parameters are chosen and are subjected to variations. Firstly, the gold layer thickness parameter (t_g) is varied, keeping others constant. An optimized magnitude t_g is obtained and fixed for the further optimization of the radius of small and big air holes. The sensor performance, is quantified by obtaining the value of wavelength sensitivity and also by considering a moderate loss spectrum. Wavelength sensitivity is demonstrated by the equation (Meng et al. 2022),

$$S_\lambda\left(\frac{nm}{RIU}\right) = \frac{\Delta\lambda_{peak}}{\Delta n_a} \tag{7}$$

Confinement loss can be defined as (Hossain et al. 2020)

$$C_L = 8.686 \times \frac{2\pi}{\lambda} \times I_m\left(n_{eff}\right) \times 10^4 \frac{dB}{cm} \tag{8}$$

3.2.1 Influence of the plasmonic enhancer MXene

A 14 nm layer of the 2D plasmonic enhancer MXene is coated along the concave-shaped microchannels on either side of the gold based SPR PCF Biosensor. As we can see in

Fig. 6, initial analysis conducted on the gold (40 nm) only biosensor resulted in a sensitivity of 9714.29 nm/RIU and 3000 nm/RIU whose coupling wavelength occurs at 1295 nm and 705 nm for the RI of 1.42 and 1.38, corresponding to HIV infected cell and Sickle cell anaemia. Further analysis with the MXene-gold hybrid layered PCF SPR sensor revealed a significant surge in the wavelength sensitivity to 13,071.43 nm/RIU and 3500 nm/RIU where coupling occurs at 1550 nm and 740 nm for HIV cell and Sickle cell anaemia. The enhanced sensitivity results from the various properties of MXene such as elevated surface area (promoting higher analyte interaction), metallic conductivity causing effective coupling and biocompatibility. Thus, MXene coated sensor has the potential for the detection of HIV and sickle cell anaemia.

3.2.2 Influence of the thickness of gold layer

The gold layer thickness is carefully optimized to elevate the sensor performance. In view of initial study, a 40 nm thickness of gold layer and 14 nm MXene layer, a maximum sensitivity of 13,071.42 nm/RIU and 3500 nm/RIU and a loss of 4246.6 dB/cm and 652.06 dB / cm is obtained for HIV infected cell and sickle cell anaemia respectively. A slight elevation in wavelength sensitivity to 13,571.42 nm/RIU and 3666.66 nm/RIU for HIV and Sickle cell anaemia respectively is noticed when a 35 nm thick gold layer is used. However, the thinner gold layer leads to a higher loss spectrum due to the reduced excitation of the surface plasmon polaritons (SPPs). On employing a thicker gold layer (45 nm), a significant reduction in the wavelength sensitivity to 12,714.28 nm/RIU and 3500 nm/RIU is detected.

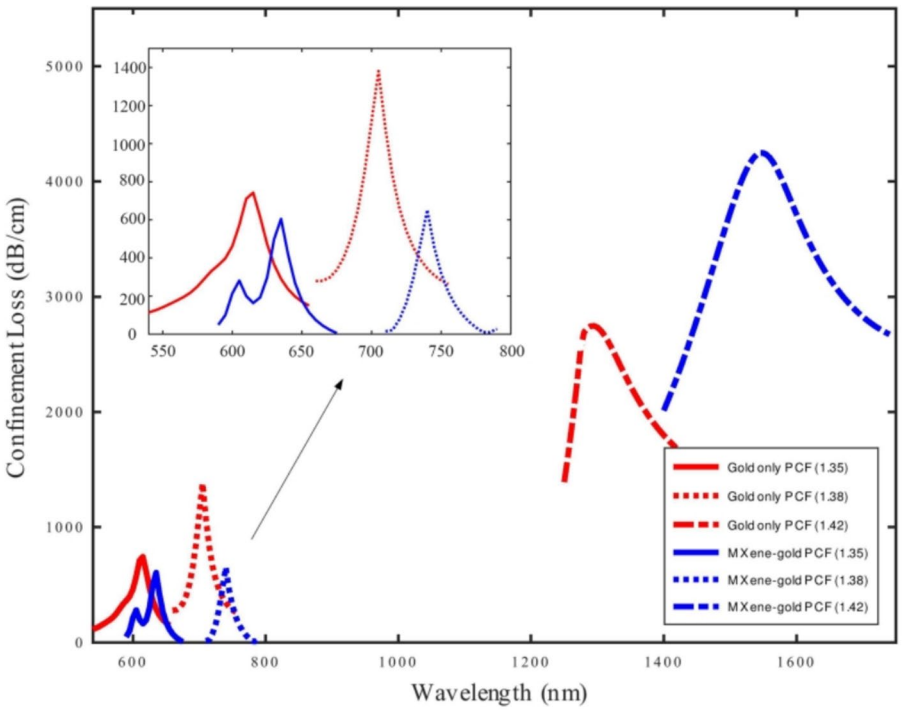


Fig. 6 Comparison of gold only PCF and MXene-Gold PCF

As the thickness of gold layer rises, the evanescent field produced by the SPPs penetrates lesser into the surrounding analyte. This reduced interaction with the bio-analyte makes it less responsive to the RI variations, resulting in lower sensitivity. In view of the moderate loss spectrum and an elevated wavelength sensitivity, optimized sensor performance is obtained at a thickness of 40 nm gold layer. The graphical summary of the gold layer thickness optimization is demonstrated in the Fig. 7.

3.2.3 Influence of the radius of smaller air holes

The air holes in vicinity to the core plays a pivotal role in ensuring efficient coupling between the core mode and SPP mode. The radius is varied and assigned three different values (200 nm, 250 nm and 300 nm) as shown in the Fig. 8. The wavelength sensitivity of HIV infected cell and sickle cell anaemia increased from 13,071.43 nm/RIU and 3500 nm/RIU to 13,357.14 nm/RIU and 3833.33 nm/RIU when the radial parameter was varied from 200 to 250 nm with the increment of 50 nm. The sensitivity of HIV cell could be further increased to 15,825.71 nm/RIU for a radius of 300 nm. However, its loss spectrum reaches a very high and undesirable value of 5144.80 dB/cm. This is due to the fact that bigger air holes near the core causes leakage of optical energy to the cladding and plasmonic metal layers leading to radiative losses. Moreover, it curtails the RI contrast between the core and cladding. This degradation in the sensor performance cannot be encouraged and hence, the airhole of radius 250 nm is opted as the optimized parameter.

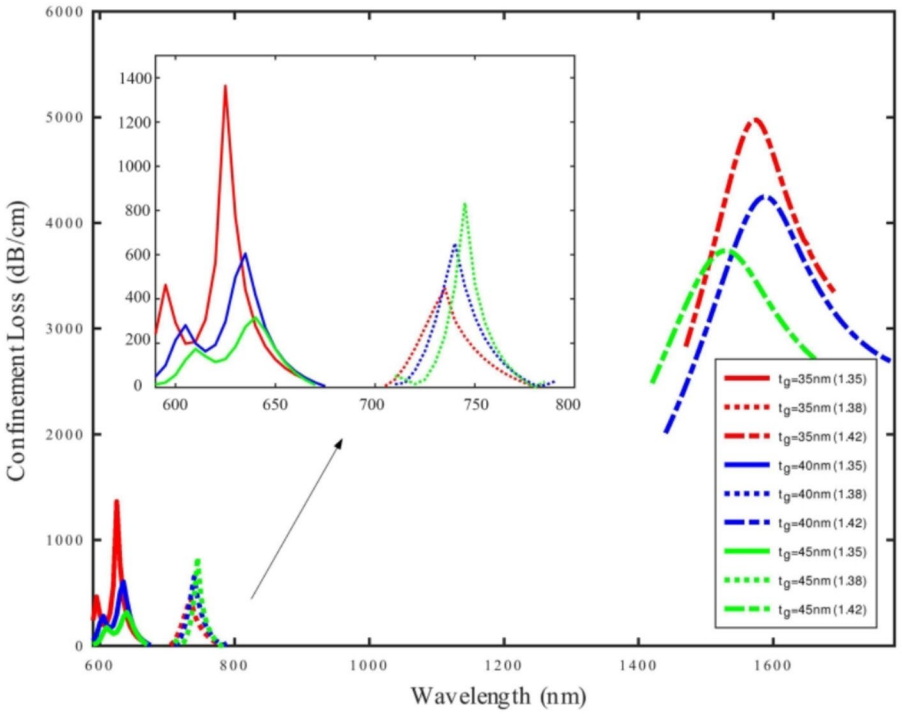


Fig. 7 Gold thickness optimization of the MXene-gold PCF

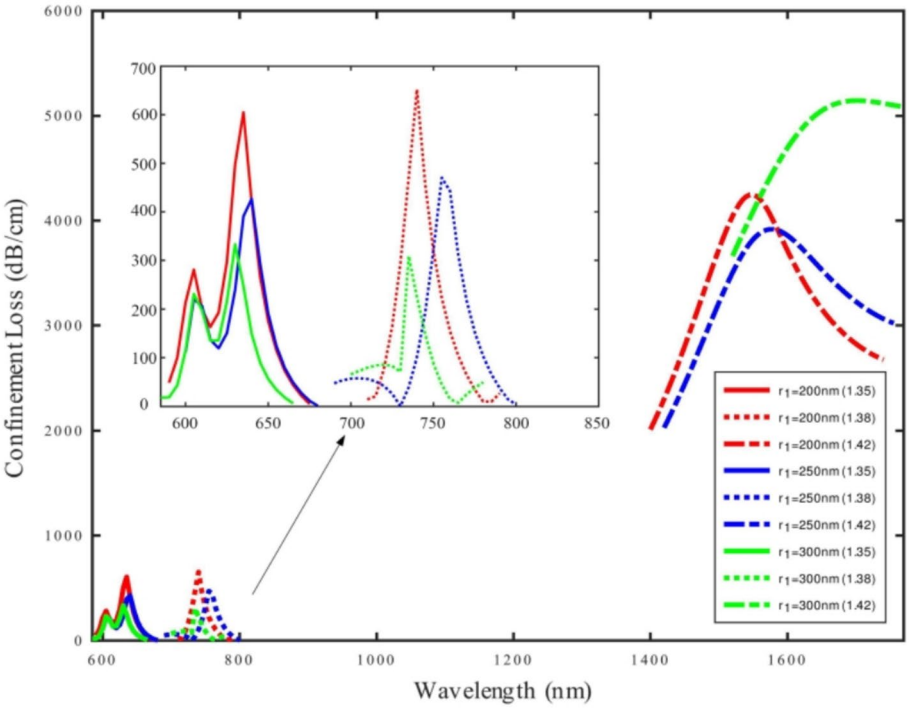


Fig. 8 Small airhole radius optimization of the MXene-gold PCF

3.2.4 Influence of the radius of bigger air holes

The size of the outer set of bigger airholes can largely influence sensor performance by altering the sensitivity and confinement loss. A boost in wavelength sensitivity to 14,142.86 nm/RIU and 4000 nm/RIU for HIV cell and sickle cell anaemia respectively, is observed at a radius of 525 nm, in comparison to both 475 nm and 575 nm. On inspecting the loss peaks given in the Fig. 9, we can conclude that moderate loss is also obtained at 525 nm radius, bearing a value of 4655.3 dB/cm and 475.37 dB/cm for HIV and sickle cell respectively. We can also observe that the smaller airhole (475 nm) gave a much higher confinement loss as they diminish the effective mode index contrast between the core and cladding. Moreover, inefficient coupling caused by smaller airholes results in inadequate energy transmission, resulting in higher loss. Therefore, optimized sensor performance is obtained at a radius of 525 nm.

The analysis using the initial parameters revealed how the utilization of MXene along with gold reflected in an increase in wavelength sensitivity of the PCF sensor. The optimized parameters are tabulated at Table 1. After optimization of thickness of gold and radii of both big and small air holes, we can observe a significant hike in both Gold only and MXene-Gold PCF. The wavelength sensitivity of gold only PCF for HIV cell and Sickle cell anaemia extracted as 13,071.42 nm/RIU and 3500 nm/RIU, on using the optimized structural parameters as shown in Fig. 10. A surge is observed in MXene-gold PCF also, wherein the sensitivity increased to 14,142.85 nm/RIU and 4000 nm/RIU for HIV and

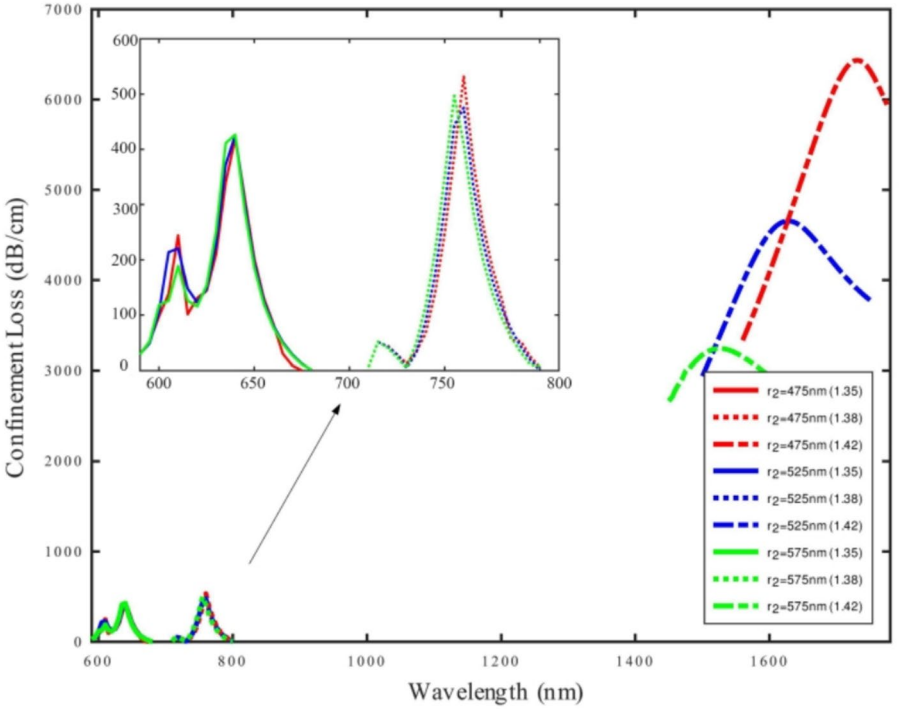


Fig. 9 Big airhole radius optimization for MXene-gold PCF

Table 1 Optimized parameters

Parameter	t_m	t_g	r_1	r_2
Value	14 nm	40 nm	250 nm	525 nm

sickle cell anaemia respectively and the analysis is displayed in Fig. 11. The comparative study is also tabulated in Table 2. The HIV and sickle cell anaemia are first detected through photonic crystal probe by britto et al. (Britto et al. 2024). In comparison to previous sensor, the proposed sensor offers an advanced sensitivity and the same is displayed in Table 3.

3.3 Tolerance study of the structural parameters

To study the tolerance of the structural parameters, the dimensions are adjusted at a range of ± 5% to analyze the significant variations in the optimized results. The first tolerance study covers the variation of gold thickness by ± 5%, shows no variation in the sensitivity with a change in confinement loss by ~ ± 10% from the optimized results. After discussing the thickness of gold, the small air holes are varied by ± 5%, observing negligible variation with the optimized sensitivity. The respective loss at 1.38 RI (Sickle cell) have suffered alterations of ~ ± 8%. At final step of tolerance study, big air holes are altered by ± 5%, again showing minor changes in the sensitivity magnitude with loss variations of ± 6%. The

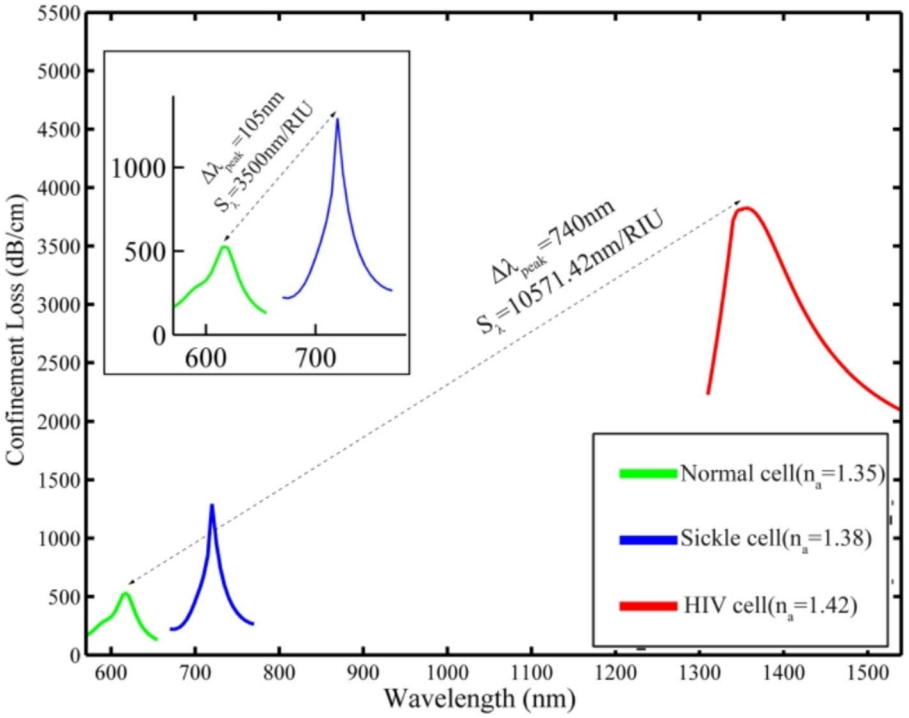


Fig. 10 Loss spectra and sensitivity of only gold PCF with optimized parameters

whole tolerance study concludes that the sensitivity is not affected by ± 5% variations of the parameters, which creates our proposed sensor highly feasible to experimental fabrication, and can tolerate minor dimensional instabilities. The performance analysis of only gold sensor and MXene-gold SPR sensor is shown through bar graph in Fig. 12.

4 Conclusion

HIV caused AIDS and Sickle cell anaemia are two of the most threatening diseases in the present scenario. Although there exist many treatment approaches, none is useful in the case of delayed detection. The aim of this work is to fabricate a biosensor reliant on the RI of the bio-analytes of HIV infected cell and Sickle cell anaemia, facilitating a much earlier detection. The proposed design put forth a PCF sensor based on SPR for the concurrent detection of HIV infected cell and Sickle cell anaemia. The two sets of semi-concentric arrangement of air holes offers an eye-shaped core, with a MXene-Gold hybrid layered coating on the microchannels on either side. The concave shaped microchannels produce a much-pronounced wavelength sensitivity as it provides optimized light guidance and

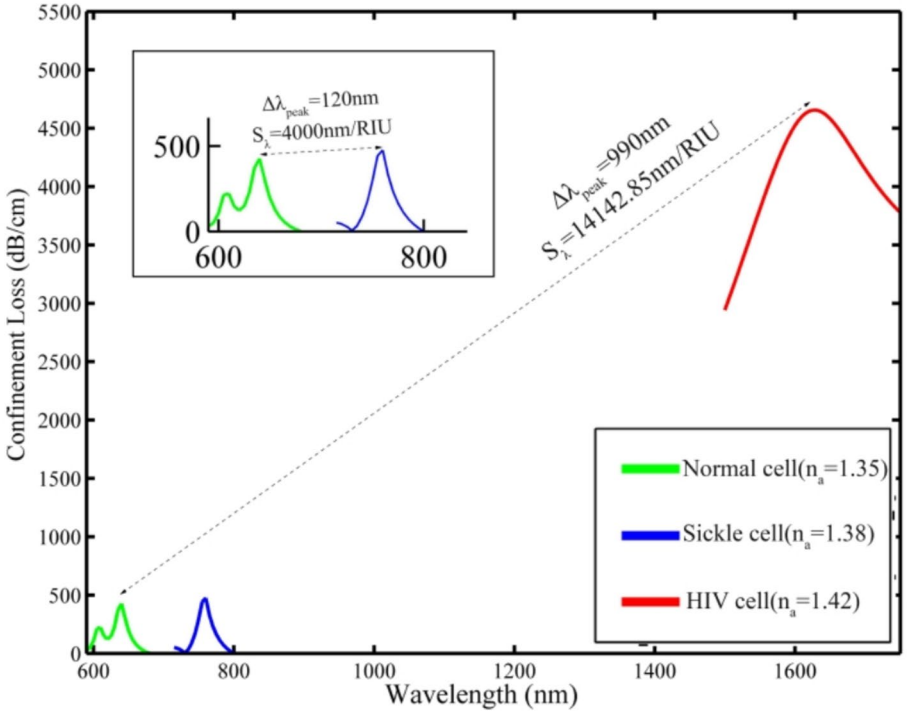


Fig. 11 Loss spectrum and sensitivity for MXene-gold PCF with Optimized parameter

Table 2 Comparative study between only gold PCF and MXene-gold PCF using optimized parameters

Biosensing application	Only Gold sensor		MXene-Gold sensor	
	Wavelength sensitivity (nm/RIU)	Confinement loss (dB/cm)	Wavelength sensitivity (nm/RIU)	Confinement loss (dB/cm)
HIV cell detection	10,571.42	3825.1	14,142.85	4655.3
Sickle cell anaemia detection	3500	1293.2	4000	475.37

enhanced surface area, resulting in increased coupling of evanescent field. The highlight of the design is the appropriate use of the bio-compatible, 2D plasmonic enhancer MXene along with a layer of gold for maximizing the wavelength sensitivity. A noticeable rise of sensitivity from 10,571.42 nm/RIU and 3500 nm/RIU (Gold only PCF) to 14,142.85 nm/RIU (+34%) and 4000 nm/RIU (+15%) (MXene-Gold PCF) can be achieved on optimization of the structural parameters. Therefore, in view of the desired sensor performance based on wavelength sensitivity and confinement loss spectra for the suggested sensor, it is appropriate for the detection of HIV infected cell and Sickle cell anaemia.

Table 3 Comparative study of the proposed SPR PCF sensor with an earlier published sensor

Biosensing application	Type of sensor	Refractive index difference	Wavelength Sensitivity (nm/RIU)	References
Cancer	SPR based PCF	1.385–1.399	11,429 nm/RIU (Breast cancer)	Bhuyan et al. (2024)
	SPR based PCF	1.376–1.390	6071 nm/RIU (Blood cancer)	Bhuyan et al. (2024)
	Split Ring Resonator	1.3333–1.4833	1778.3 nm/RIU (Brain cancer)	Khodaie et al. 2025)
	Split Ring Resonator	1.36–1.38	570.4 nm/RIU (Skin Cancer)	Khamaru and Kumar (2024c)
Malaria	SPR based PCF	1.395–1.402	26,428.57 nm/RIU	Meshginqalam and Barvestani (2022)
	SPR based PCF	1.395–1.402	38,571 nm/RIU	Khodaie and Heidarzadeh 2024)
Corona Virus	Split Ring Resonator	1.29–1.37	557 nm/RIU	Amin et al. 2024)
	SPR based PCF	1.29–1.37	13,000 nm/RIU	Khodaie and Heidarzadeh, (2024)
HIV	Photonic Crystal	1.35–1.42	1034 nm/RIU	Britto et al. (2024)
	SPR based PCF	1.35–1.42	**14,142.86 nm/RIU (Proposed sensor)**	
Sickle Cell Anaemia	Photonic Crystal	1.35–1.38	1071 nm/RIU	Britto et al. (2024)
	SPR based PCF	1.35–1.38	**4000 nm/RIU (Proposed sensor)**	

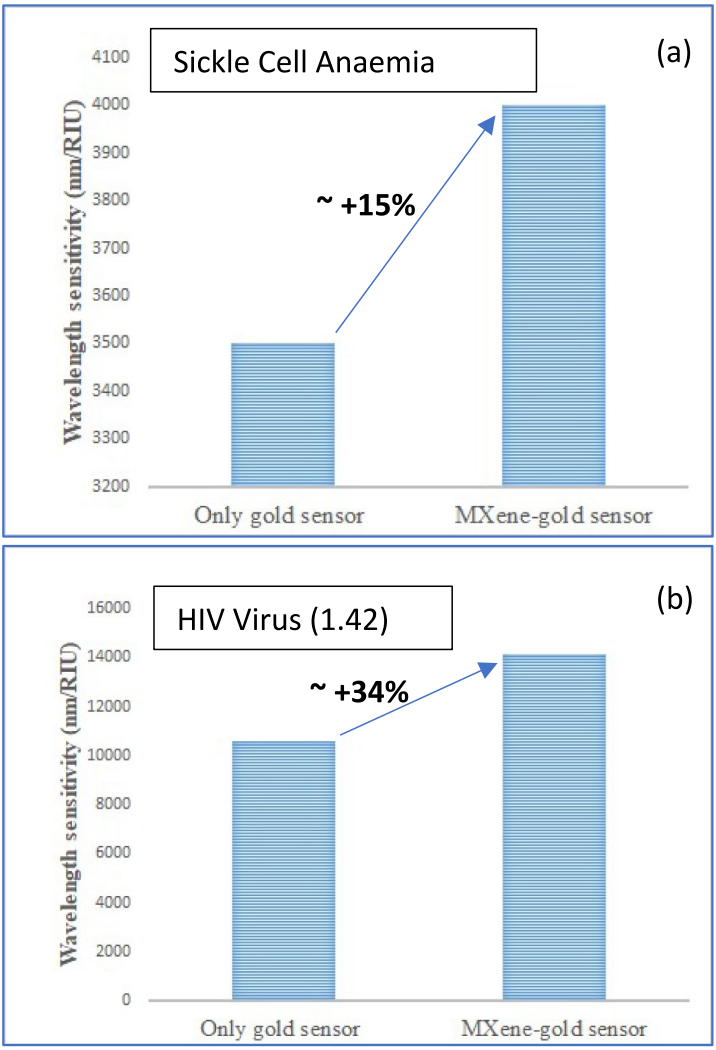


Fig. 12 Performance analysis of Only Gold and MXene-Gold PCF Sensor

Acknowledgements Preethi Krishnan, student of Masters of Science in Physics in the Department of Applied Physics at Delhi Technological University acknowledges for providing an excellent research facilities for conducting research.

Author contribution P.K. did the simulation and presented results and wrote the original manuscript, A. Khamaru helped in the explanation of results and helped in paper writing, A. K. reviewed paper and concept building.

Funding There is no funding associated with this work.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no competing interests.

References

Amin, S., Sarker, H., Jawad, A.: A hybrid sensing approach for coronavirus and alcohol detection based on PCF-SPR biosensor. Plasmonics, 1–16 (2024). https://doi.org/10.1007/s11468-024-02571-3

BS, D. K., A. Khamaru, A. Kumar. Design and Analysis of Twin-Core Side-Polished SPR PCF Sensor for Preliminary Malaria Detection. Plasmonics, 1–12, 2025

Bhargava Reddy, M.S., Kailasa, S., Marupalli, B.C., Sadasivuni, K.K., Aich, S.: A family of 2D-MXenes: synthesis, properties, and gas sensing applications. ACS Sensors **7**(8), 2132–2163 (2022)

Bhuyan, A., Khamaru, A., Kumar, A.: Black phosphorus-based slotted D-shaped PCF SPR sensor for cancer detection. Plasmonics, 1–13 (2024)

Britto, E.C., Krishnamoorthi, B., Rajasekar, R., Nizar, S.M.: Photonic crystal-based nanoscale multipurpose biosensor for detection of brain tumours, HIV, and anaemia with high sensitivity. Plasmonics, **19**(5), 2687–2703 (2024)

Capelli, D., Scognamiglio, V., Montanari, R.: Surface plasmon resonance technology: recent advances, applications and experimental cases. TrAC, Trends Anal. Chem. **163**, 117079 (2023)

Chaudhary, V.S., Kumar, D., Kumar, S.: Au-TiO_2 coated photonic crystal fiber based SPR refractometric sensor for detection of cancerous cells. IEEE Trans. Nanobiosci. **22**(3), 562–569 (2022)

Dillon, A.D., Ghidiu, M.J., Krick, A.L., Griggs, J., May, S.J., Gogotsi, Y., Fafarman, A.T.: Highly conductive optical quality solution-processed films of 2D titanium carbide. Adv. Funct. Mater. **26**(23), 4162–4168 (2016)

Divya, J., Selvendran, S., Itapu, S., Borra, V.: A novel dual-channel SPR-based PCF biosensor for simultaneous tuberculosis and urinary tract infection diagnosis toward SDG3. IEEE Access **12**, 85484–85498 (2024). https://doi.org/10.1109/ACCESS.2024.3410698

Divya, J., Selvendran, S., Sivanantha Raja, A., Borra, Vamsi: A novel plasmonic sensor based on dual-channel d-shaped photonic crystal fiber for enhanced sensitivity in simultaneous detection of different analytes. IEEE Trans. NanoBiosci. **23**(1), 127–139 (2024). https://doi.org/10.1109/TNB.2023.3294330

El-assar, M., Taha, T.E., El-Samie, F.E.A., Fayed, H.A., Aly, M.H.: Zinc selenide based dual-channel SPR optical biosensor for HIV genome DNA hybridization detection. Opt. Quant. Electron. **55**(13), 1143 (2023)

Elabdein, M.Z., Khedr, O.E., Nazmi, A., Mohammed, E.-S.M., El-Rabaie: Recent advances in photonic crystal fiber based chemical and industrial sensors: a review. J. Opt., 1–19 (2024). https://doi.org/10.1007/s12596-024-02218-w

Fei, Y., Luo, B., An, M., Hu, T., Lin, W., Jia, H.: A Highly efficient d-shaped dual-core PCF-SPR sensor coated with ITO film for refractive index detection. Plasmonics, 1–15 (2024)

Gandhi, R.T., Bedimo, R., Hoy, J.F., Landovitz, R.J., Smith, D.M., Eaton, E.F., Saag, M.S.: Antiretroviral drugs for treatment and prevention of HIV infection in adults: 2022 recommendations of the International Antiviral Society. USA Panel. Jama **329**(1), 63–84 (2023)

Garg, D., Khamaru, A., Kumar, A.: Supercontinuum generation in Ga-Sb-S chalcogenide-based PCF using optofluidic approach. Microw. Opt. Technol. Lett. **66**(9), e34316 (2024)

Gupta, P., Khamaru, A., Kumar, A.: Highly sensitive black phosphorus–layered SPR-PCF refractometric sensor. Plasmonics, 1–11 (2025). https://doi.org/10.1007/s11468-024-02723-5

Gupta, A., Singh, T., Singh, R.K., Tiwari, A.: Numerical analysis of coronavirus detection using photonic crystal fibre-based SPR sensor. Plasmonics **18**(2), 577–585 (2023)

Hossain, M.B., Hossain, M.S., Islam, S.R., Sakib, M.N., Islam, K.Z., Hossain, M.A., Cho, G.H.: Numerical development of high performance quasi D-shape PCF-SPR biosensor: an external sensing approach employing gold. Results Phys. **18**, 103281 (2020)

Jeon, J., Yang, Y., Choi, H., Park, J.H., Lee, B.H., Lee, S.: MXenes for future nanophotonic device applications. Nanophotonics **9**(7), 1831–1853 (2020)

Khamaru, A., Kumar, A.: Highly sensitive graded index PCF-based refractometric SPR sensor offering enhanced FOM. Plasmonics, 1–13 (2024a)

Khamaru, A., Kumar, A.: As38Se62 based segmented clad-graded index photonic crystal fiber for supercontinuum generation covering 3–9.5 μm with moderate peak power. Opt. Quant. Electron. **56**(7), 1246 (2024b)

Khamaru, A., Kumar, A.: Ge-Se-Te based penrose photonic quasi-crystal fiber for SCG covering 2–21 μm MIR regime. Opt. Mater. **155**, 115849 (2024c)
Khodaie, A., Heidarzadeh, H.: Evaluation of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) using a high figure-of-merit plasmonic multimode refractive index optical sensor. Sci. Rep. **14**(1), 25499 (2024)
Khodaie, A., Heidarzadeh, H.: Design and analysis of a multi-modal refractive index plasmonic biosensor based on split ring resonator for detection of the various cancer cells. Opt. Quant. Electron. **56**(9), 1439 (2024)
Khodaie, A., Heidarzadeh, H., Harzand, F.V.: Development of an advanced multimode refractive index plasmonic optical sensor utilizing split ring resonators for brain cancer cell detection. Sci. Rep. **15**(1), 433 (2025)
Kumar, V., Raghuwanshi, S.K., Kumar, S.: Highly sensitive Ag/$BaTiO_3$/MoS_2 nano composite layer based SPR sensor for detection of blood and cervical cancer. Results Opt. **14**, 100597 (2024)
Kumar, B.H., Vaibhav, A.M., Srikanth, P.C.: Si/SiO_2 based nano-cavity biosensor for detection of anemia, HIV and cholesterol using refractive index of blood sample. Indian J. Sci. Technol. **15**(18), 899–907 (2022)
Kumar, A., Verma, P., Jindal, P.: Surface plasmon resonance sensor based on MXene coated PCF for detecting the cancer cells with machine learning approach. Microelectron. Eng. **267**, 111897 (2023)
Li, W., Chen, Y., Xu, J., Jiang, M., Zou, H.: A D-shaped SPR-based PCF sensor with an extremely high-amplitude sensitivity for measuring the refractive index. Micromachines **14**(7), 1295 (2023)
Li, K., Jiao, T., Xing, R., Zou, G., Zhou, J., Zhang, L., Peng, Q.: Fabrication of tunable hierarchical MXene@ AuNPs nanocomposites constructed by self-reduction reactions with enhanced catalytic performances. Sci. China Mater. **5**(61), 728–736 (2018)
Luo, G., Zhang, X., Zhu, X.S., Shi, Y.W.: Experimental study of the copper-coated hollow fiber surface plasmon resonance sensor. Opt. Commun. **520**, 128481 (2022)
Majeed, M.F., Ahmad, A.K.: Design and analysis of a high sensitivity open microchannel PCF-based surface plasmon resonance refractometric sensor. Opt. Mater. **147**, 114617 (2024)
Mao, Y., Ren, F., Zhou, D., Li, Y.: Highly sensitive PCF-SPR RI sensor for cancer detection using gold/graphene/$Ti_3C_2T_x$-MXene hybrid layer. Plasmonics **23**, 1–12 (2024)
Melwin, G., Senthilnathan, K.: Modelling a simple arc shaped gold coated PCF-based SPR sensor. J. Opt. **53**(1), 117–126 (2024)
Meng, F., Wang, H., Fang, D.: Research on D-shape open-loop PCF temperature refractive index sensor based on SPR effect. IEEE Photonics J. **14**(3), 1–5 (2022)
Meshginqalam, B., Barvestani, J.: Highly sensitive photonic crystal fiber-based plasmonic biosensor with improved malaria detection application. Europ. Phys. J. Plus **137**(5), 1–11 (2022)
Mu, R., Wan, H., Shi, W., Liang, H., Lou, Y.: Design and theoretical analysis of high-sensitive surface plasmon resonance sensor based on Au/$Ti_3C_2T_x$-MXene hybrid layered d-shaped photonic crystal fiber. IEEE Sens. J. **23**(16), 18160–18167 (2023)
Mumtaz, F., Roman, M., Zhang, B., Abbas, L.G., Dai, Y., Ashraf, M.A., Kumar, A.: MXene (Ti_3C_2Tx) coated highly-sensitive D-shaped photonic crystal fiber based SPR-biosensor. Photon. Nanostruct. Fundam. Appl. **52**, 101090 (2022)
Opoku, G., Danlard, I., Dede, A., Akowuah, E.K.: Design and numerical analysis of a circular SPR based PCF biosensor for aqueous environments. Results Opt. **12**, 100432 (2023)
Parums, D.V.: Forty years of waiting for prevention and cure of HIV infection-ongoing challenges and hopes for vaccine development and overcoming antiretroviral drug resistance. Med. Sci. Monitor Int. Med. J. Experim. Clin. Res. **30**, e944600–e944601 (2024)
Singh, J., Khamaru, A., Kumar, A.: Spiral shaped highly sensitive rectangular PCF-based cancer cells detector in terahertz regime. Phys. Scr. **99**(11), 115546 (2024a)
Singh, N., Khamaru, A., Kumar, A.: Design and analysis of a rectangular core refractive index-based PCF sensor for bio-sensing application. Opt. Quant. Elect. **56**(7), 1133 (2024)
Singh, S., Prajapati, Y.K.: Novel bottom-side polished PCF-based plasmonic biosensor for early detection of hazardous cancerous cells. IEEE Trans. Nanobiosci. **22**(3), 647–654 (2023)
Tavakoli, A., Karbalaie Niya, M.H., Keshavarz, M., Ghaffari, H., Asoodeh, A., Monavari, S.H., Keyvani, H.: Current diagnostic methods for HIV. Future Virology **12**(3), 141–155 (2017)
Tisdale, J.F., Thein, S.L., Eaton, W.A.: Treating sickle cell anemia. Science **367**(6483), 1198–1199 (2020)
Z. Tasnim, R. Islam, R. L. Khan, & E. Moazzam, "*Design and Investigation of PCF based Highly Sensitive Surface Plasmon Resonance Biosensors* "(Doctoral dissertation, Department of Electrical and Electronic Engineering, Islamic University of Technology (IUT) The Organization of Islamic Cooperation (OIC) Board Bazar, Gazipur-1704, Bangladesh), 2021

Zhang, Z., Shen, T., Wu, H., Feng, Y., Wang, X.: A temperature sensor based on D-shape photonic crystal fiber coated with Au–TiO_2 and Ag–TiO_2. Opt. Quant. Electron. **53**, 1–11 (2021)
Zhang, Y., Xue, J., Liu, W., Zhang, Y., Li, S., Liu, Z., Yuan, L.: Cascaded dual-channel fiber SPR sensor based on $Ge_2Sb_2Te_5$. IEEE Sens. J. **22**(5), 4083–4089 (2022)
Zhang, Z., Yin, Z., Li, S.: A SPR sensor with wide RI measuring range and narrow FWHM based on PCF-silver film structure. J. Lightw. Technol. **42**(20), 7398–7404 (2024)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Materials Research Proceedings 49 (2025) 571-580 https://doi.org/10.21741/9781644903438-57

Navigating the green combustion landscape: Optimizing emissions and performance in CI engines fueled by biogas and nanoparticle-doped biodiesel

S. LALHRIATPUIA[1,a *], H. LALHMINGSANGA[1 *], Gulam MUSTAFA[2], Neeraj BUDHRAJA[2] and Kiran PAL[3,b]

[1]Department of Mechanical Engineering, National Institute of Technology Mizoram, Mizoram – 796012, India

[2]Department of Mechanical Engineering, Delhi Technological University, Delhi – 110042, India

[3]Department of Mathematics, DSEU, Okhla Campus II, New Delhi, India

[a]sactrix777@gmail.com, [b]kiranpal@dseu.ac.in

Keywords: ANN, RSM, HHO, Biodiesel, Biogas, Nanoparticles

Abstract. This study explores the use of biogas and nanoparticle-doped biodiesel as sustainable options for compression-ignition (CI) engines, driven by worries about the depletion of fossil fuels and their environmental effect. This study provides a complete overview of the interconnected impacts of engine load, nickel oxide nanoparticle doping rate, Neem biodiesel mix ratio, and biogas flow rate on the performance and emissions of a CI engine. Response Surface Methodology (RSM) provides a means to comprehend and conduct organized experiments, while Artificial Neural Network (ANN) is capable of capturing intricate non-linear interactions. Both models undergo thorough statistical analysis utilizing several assessment metrics to reveal the impact of each input parameter on important outcomes. This study also extensively explores the important field of optimization, focusing on the progress achieved in RSM and ANN models. RSM's desirability function optimized the output parameters at 67.45% engine load, 96.06 ppm NDR, 10.7 % BBR and 0.85 kg/h BFR, while ANN with Harris Hawks Optimization (HHO) optimized the output parameters at 67.01% engine load, 98.39 ppm NDR, 8.41% BBR and 0.846 kg/h BFR. This integrated approach offers a robust framework for enhancing the sustainability and efficiency of CI engines, presenting practical solutions for transitioning toward cleaner energy systems.

Introduction

Despite its reliance on fossil fuels, the internal combustion engine continues to be a prominent factor in transportation. The limited availability and environmental impacts of these resources, such as the release of greenhouse gases and air pollution, need a fundamental change in our approach towards sustainable alternatives. Biogas, an abundant and sustainable fuel produced by the anaerobic decomposition of organic material, presents a very promising option. Biogas is mostly composed of methane (CH_4) in addition to carbon dioxide (CO_2) and other minor gases. Biogas, although having a carbon-neutral combustion cycle, might result in higher levels of NO_X emissions because of its naturally high combustion temperature [1].

Biodiesel, a fuel made from vegetable oil, offers an additional sustainable alternative. Nevertheless, the elevated viscosity and lower cetane number of this substance in comparison to diesel might result in diminished engine performance and heightened emissions [2]. Nanoparticle doping is a potential approach to enhance the properties of biodiesel. Nanoparticles such as ceria (CeO_2) and ceria-zirconia ($CeZrO_2$) have shown promise in augmenting combustion, decreasing emissions, and boosting stability [3].

To achieve optimal engine performance and reduce emissions while using sustainable fuels, it is necessary to have a thorough grasp of how different elements interact with each other. The

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 license. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. Published under license by Materials Research Forum LLC.

combustion characteristics and emissions are greatly influenced by engine load. Elevated loads result in higher temperatures inside the combustion chamber, which in turn leads to an increase in the generation of nitrogen oxides (NO_X) [4].

Similarly, the pace at which nanoparticles are doped may have an impact on combustion efficiency, stability, and emissions. An ideal doping rate may optimize combustion by increasing the availability of oxygen and enhancing ignition characteristics. Nevertheless, an excessive amount of doping may result in the aggregation of particles and an increase in viscosity, which can have a detrimental effect on the performance of the engine [5].

The biodiesel mix ratio in conjunction with conventional diesel is crucial in establishing the equilibrium between sustainability and engine performance. Increasing the mix ratios may enhance lubricity and decrease smoke emissions [6]. Nevertheless, the use of biodiesel might result in reduced brake thermal efficiency (BTE) and higher brake specific fuel consumption (BSFC) owing to the lower calorific value of biodiesel in comparison to diesel [7].

To achieve optimal engine performance and reduce emissions while using sustainable fuels, it is necessary to have a thorough grasp of how different elements interact with each other. The combustion characteristics and emissions are greatly influenced by the engine load. Elevated loads result in higher temperatures inside the cylinder, which in turn leads to an increase in the generation of nitrogen oxides (NO_X) [8,9]. Similarly, the pace at which nanoparticles are doped may have an impact on combustion efficiency, stability, and emissions. An optimum doping rate may optimize combustion by increasing the availability of oxygen and enhancing ignition characteristics. Nevertheless, an excessive amount of doping may result in the aggregation of particles and an increase in viscosity, which can have a detrimental effect on the performance of the engine [10].

The biodiesel mix ratio in conjunction with conventional diesel is a critical factor in establishing the equilibrium between sustainability and engine performance. Increasing the mix ratios may enhance lubricity and decrease smoke emissions [11,12]. Nevertheless, the use of biodiesel might result in reduced brake thermal efficiency (BTE) and higher brake specific fuel consumption (BSFC) owing to the lower calorific value of biodiesel in comparison to diesel [13,14]. It is important to carefully examine the flow rate of biogas, since too much flow might result in power loss and incomplete combustion. This may lead to a rise in hydrocarbon (HC) emissions as a consequence of unburned fuel particles escaping from the combustion chamber [15,16]. Identifying the most efficient biogas flow rate is crucial in order to strike a balance between power generation and emissions.

It is necessary to go beyond developing models to achieve optimal engine performance and reduce emissions while using sustainable fuels. The Desirability Function Approach (DFA) in Response Surface Methodology (RSM) provides a robust method for addressing the trade-offs between competing goals in optimization [17-19]. Within the framework of CI engines, this might include the simultaneous reduction of NOx emissions and the maximization of BTE (Brake Thermal Efficiency). The DFA methodology enables the consolidation of distinct desire functions for each answer into a unified overall desirability function. This overall function guides the optimization process towards a solution that fulfills all objectives [20,21].

The objective of this study is to explore the combined effects of nanoparticle-doped biodiesel, biogas, and engine parameters on the performance and emissions of compression ignition (CI) engines. It aims to identify optimal conditions that enhance engine efficiency and minimize harmful emissions. The novelty lies in the use of advanced modeling techniques, such as RSM and ANN, along with Harris Hawks Optimization (HHO) to capture complex interactions between variables. This integrated approach provides valuable insights for optimizing sustainable fuel usage in CI engines.

Materials Research Proceedings 49 (2025) 571-580 https://doi.org/10.21741/9781644903438-57

Material and Methods

This section provides a comprehensive description of the experimental setup, which encompasses the specifications of the CI engine, the kinds of biogas and nanoparticle dopants used, and the procedure for preparing the nanoparticle-doped biodiesel.

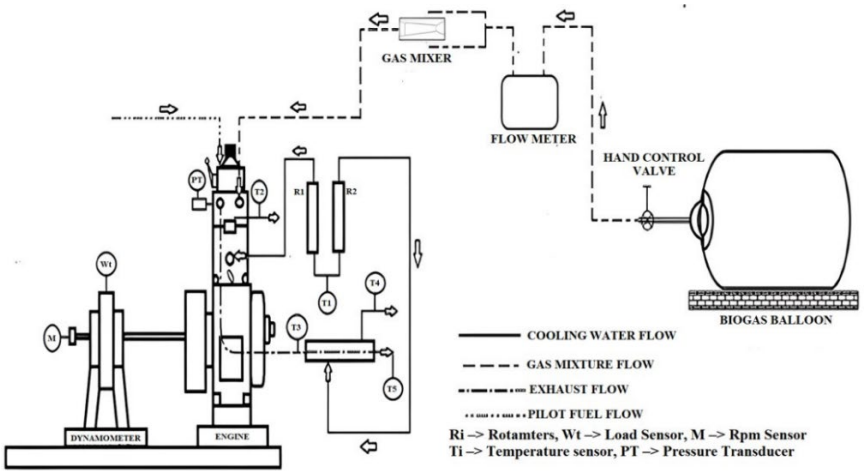


Fig. 1. *Schematic representation of test setup after modification*

The experiment was conducted using a single-cylinder, four-stroke, constant-speed compression ignition (CI) engine with a rated power of 3.5 kW. An eddy current dynamometer was used to apply engine load, which ranged from 2.4 kg (20%) to 12 kg (100%). To monitor the engine's performance, rotameters measured the flow of water for cooling, while airflow and fuel flow were tracked using a manometer and fuel flow meter, respectively. Temperature sensors (PT100) were strategically placed at various points in the engine cylinder to continuously record data. These sensors were connected to a National Instruments (NI) unit, which relayed the data to a computer via Engine Soft software. To facilitate dual-fuel operation, a gas-air mixing device was incorporated into the setup, designed according to engine specifications. A schematic of the test setup and gas mixer, based on our previous research, is shown in Fig. 1 [21].

Experimental Design

As given in Table 1, the nickel oxide Nanoparticle Doping Rate (NDR) was varied between 0 and 100 ppm, as nanoparticles are known to enhance combustion by facilitating better oxidation and improving fuel efficiency. The neem biodiesel blend ratio (BBR) ranged from 0% to 20% by volume, encompassing the transition from pure diesel to biodiesel, with the blend ratio being a critical factor in improving engine lubricity and reducing particulate emissions. Higher biodiesel content generally improves the engine's sustainability but can reduce the calorific value of the fuel, which impacts fuel efficiency. The Biogas Flow Rate (BFR) was varied from 0 to 1 kg/h, as biogas, primarily composed of methane, serves as a supplemental fuel in dual-fuel engines.

Table 1 *Input Parameters and Levels*

Input Parameters	Units	Symbol	Levels		
			-1	0	1
Engine load	%	Load	20	60	100
Nanoparticles doped rate	ppm	NDR	0	50	100
Biodiesel blend rate	%	BBR	0	10	20
Biogas Flow rate	kg/h	BFR	0.5	0.75	1

Fuel Properties

The physiochemical properties of diesel (D 100), raw neem oil (LO), and pure biodiesel (BBR100) were evaluated using various ASTM testing procedures, and the results are presented in Table 2.

Table 2 *Fuel properties*

Properties	Unit	Testing methods	D 100	NR	BBR 10	BBR 20	BBR 100
Kinematic Viscosity (40 °C)	cSt	D-445	2.98	30.6	3.15	3.36	4.21
Density (40 °C)	kg/m^3	D-1298	841	920	828.3	834.2	881
Calorific Values	MJ/kg	D-240	43.9	38.4	43.5	43.03	40.5
Flash Point	°C	D-93	60	244	90	107	169
Pour Point	°C	D-97	-14	-16	-14	-15	-18
Cetane Number	-	D-4737	51	38	50	49	42

Modeling and Optimization

RSM Modelling

The modeling and optimization of engine performance and emissions were carried out using Response Surface Methodology (RSM), which is well-suited for analyzing and optimizing complex systems with multiple variables. RSM relies on polynomial regression models that approximate the relationship between the input variables (such as NDR, BBR, BFR, and load) and output responses. These models were derived by fitting second-order polynomial equations (Equation 1) to the experimental data, which included both linear and quadratic terms to account for the direct and non-linear relationships between input parameters and responses. Interaction terms between parameters were also included in the models to explore how combinations of variables could have a synergistic or antagonistic effect on engine performance and emissions. For example, the interaction between NDR and BBR may have a combined impact on BTE or NOx emissions, which could not be captured by examining the effects of each variable in isolation.

$$Y= b_0 + \sum_{i=1}^{4} b_i x_i + \sum_{i=1}^{4} b_{ii} x_i^2 + \sum_{i\leq 1\leq j}^{4} b_{ij}\, x_i x_j \quad (1)$$

x_i is an independent input factor, whereas Y is the dependent output. This polynomial equation has an intercept coefficient denoted by b_0. The quadratic coefficients are denoted by b_{ij}, the interactive coefficients by b_{ii}, and the linear coefficients for Load, NDR, BBR, and BFR by bi. Design Expert software was used to achieve the model fitting. To validate the accuracy of the regression models, Analysis of Variance (ANOVA) was employed. This method decomposes the variation in the response variable to assess the statistical significance of the input factors and their interactions. Goodness-of-fit metrics such as R^2 were used to quantify how well the model represented the data, with high values indicating a good fit.

ANN Modelling

The fundamental design of an Artificial Neural Network (ANN) consisting of three layers. The arrows in the image depict the flow of data through the network. Data enters the input layer, propagates through the hidden layer, and reaches the output layer, where the final prediction is made [13].

A multilayer perceptron (MLP) architecture with an input layer, two hidden layers, and an output layer makes up the Artificial Neural Network (ANN) utilized in this study. Input data is represented by the input layer, and buried layers use activation functions to process it. The last forecast is produced by the output layer, where the amount of neurons is proportional to the output categories. Each output is assigned to its own network in order to train the DOE findings. Neuron

combinations ranging from 1 to 10 are used to train the network's hidden layers. After that, we follow the procedure in Table 3 to choose the neuron pairings with the lowest RMSE. For each output, a new network is constructed using the best neuron combination that was achieved. The backpropagation technique was used to train the network, with the Mean Square Error (MSE) serving as the performance metric and the Levenberg-Marquardt method (trainlm) as the training function. By experimenting with several combinations for each output, we were able to find the optimal neuron configuration for the hidden layers. The configuration with the lowest Root Mean Square Error (RMSE) was then chosen. Training accounted for 70% of the DOE's 30 runs, while testing and validation each accounted for 15%.

Table 3 architecture for engine outputs

BTE	BSEC	NO_X	HC	CO	SO
4-6-5-6	4-9-3-6	4-9-3-6	4-10-8-6	4-6-6-6	4-4-2-6

Optimization of RSM and ANN model

In RSM optimization, the desirability function is used to handle multiple response variables. It assigns a desirability score between 0 and 1 to each response, based on its optimal value and permissible range. The individual desirability scores are then combined, typically using the geometric mean, to form an overall desirability score. The goal is to find the combination of input variables that maximizes this overall desirability, optimizing conflicting response variables simultaneously. This approach allows for balancing competing objectives to identify the best solution.

Results and Discussion

RSM Model Analysis

The empirical models for analyzing key engine responses—Brake Thermal Efficiency (BTE), Brake-Specific Energy Consumption (BSEC), Nitrogen Oxides (NOx), Hydrocarbons (HC), Carbon Monoxide (CO), and Smoke Opacity (SO)—were derived using the experimental data from varying the input parameters: engine load, nanoparticle doping rate (NDR), biodiesel blend ratio (BBR), and biogas flow rate (BFR). Each model includes coefficients that determine how each parameter affects the response variable, either positively or negatively. Positive coefficients indicate a direct relationship, where an increase in that parameter will proportionally increase the response (e.g., engine load and NDR positively affect BTE). Negative coefficients suggest an inverse relationship, where an increase in the parameter results in a decrease in the response (e.g., BBR and BFR negatively affect BTE).

$$BTE = 15.9776 + 4.33371 * A + 0.79072 * B + -1.12249 * C + -1.5205 * D + 0.244867 * AB + -0.458222 * AC + -0.566295 * AD + -0.0629676 * BC + -0.0850689 * BD + 0.12252 * CD + -1.95413 * A^2 + -0.132335 * B^2 + 0.0170302 * C^2 + -0.173233 * D^2 \quad (4)$$

As given in equation 4, For BTE, engine load showed the largest positive influence, with increasing load resulting in a proportional rise in BTE due to enhanced combustion efficiency. NDR also positively impacted BTE, though its effects diminished beyond a certain threshold. In contrast, BBR and BFR reduced BTE, as higher BBR decreases the calorific value of the fuel, and excess biogas flow reduces combustion temperature, causing incomplete combustion and lower thermal efficiency. The interaction between load and NDR had a synergistic effect on BTE, meaning that increasing both load and NDR simultaneously resulted in a higher BTE. However, load with BBR or BFR exhibited antagonistic effects, meaning their combined increase led to a reduction in BTE.

$$BSEC = 21.9175 + -7.89433 * A + -1.40957 * B + 2.37464 * C + 1.70115 * D + 0.391571 * AB + -0.853927 * AC + -0.4978 * AD + -0.100708 * BC + -0.0652905 * BD + 0.126599 * CD + 5.06601 * A^2 + 0.0570941 * B^2 + 0.0420941 * C^2 + 0.440009 * D^2 \quad (5)$$

As given in equation 5, The Brake-Specific Energy Consumption (BSEC) model showed that BFR and BBR had a directly proportional relationship with BSEC, meaning that as BFR or BBR increased, so did the energy consumed per unit of work. This suggests that higher biodiesel blends and biogas flows require more energy to achieve the same power output, likely due to the lower energy density of the fuels. On the other hand, engine load and NDR had a negative impact on BSEC, with higher load and NDR leading to better fuel utilization and reduced energy consumption per unit of output. The interaction between load and NDR was synergistic, meaning that increasing both resulted in significant reductions in BSEC. Load with BBR or BFR exhibited antagonistic effects, as the increased energy demands of these factors offset the efficiency gains from higher loads.

$$NO_X = 246.545 + 136.083 * A + -17.9106 * B + 15.7414 * C + -37.9381 * D + -10.0984 * AB + 10.0213 * AC + -22.6869 * AD + -0.976769 * BC + 2.64935 * BD + -2.33765 * CD + 6.23173 * A^2 + 2.68564 * B^2 + -6.19198 * C^2 + -9.78391 * D^2 \quad (6)$$

As given in equation 6, For NOx emissions, the model showed a direct relationship with engine load and BBR, meaning that higher engine loads and increased biodiesel blending resulted in higher NOx emissions due to the higher combustion temperatures associated with these factors. Conversely, BFR and NDR exhibited a negative relationship with NO_X, meaning that increasing these variables led to a reduction in NO_X emissions. This was likely due to the oxygenation effect of biogas and nanoparticles, which improves combustion efficiency and reduces the formation of NO_X. The interaction between load and BBR, as well as between NDR and BBR, had a synergistic effect on NO_X emissions, while other combinations showed antagonistic effects. For example, compared to the BBR20-NDR100 blend, NO_X emissions decreased by 14% with BBR20-NDR0 and decreased by 27.15% for BFR1-NDR100 compared to BFR0.5-NDR100.

$$HC = 44.9637 + -6.77761 * A + -4.45204 * B + -2.79144 * C + 2.98671 * D + 0.618654 * AB + 0.382929 * AC + -0.231771 * AD + 0.249148 * BC + -0.268035 * BD + -0.16326 * CD + 0.731055 * A^2 + 0.267014 * B^2 + 1.24267 * C^2 + 1.17926 * D^2 \quad (7)$$

As given in equation 7, For HC emissions, the positive coefficients for BFR indicated that higher biogas flow leads to increased HC emissions, likely due to incomplete combustion at higher biogas flow rates. On the other hand, load, NDR, and BBR negatively affected HC emissions, meaning that higher loads, NDR, or BBR values helped reduce HC emissions. The interactions between load and NDR, load and BBR, and NDR and BBR exhibited synergistic effects on HC, indicating that optimizing these combinations could help reduce unburned hydrocarbons. At the same central load and BFR level, the BBR20-NDR100 blend showed a 17% decrease in HC emissions compared to BBR20-NDR0. In contrast, when comparing BFR1-NDR100 to BFR0.5-NDR100, there was an observed increase of 13.02% in HC emissions.

$$CO = 0.0929081 + -0.0212321 * A + -0.00945834 * B + -0.00649946 * C + 0.00726853 * D + 0.001596 * AB + 0.0004185 * AC + -0.00136361 * AD + 0.000490613 * BC + -0.00054915 * BD + -0.00037665 * CD + 0.0317948 * A^2 + 0.00100293 * B^2 + 2.9251e-06 * C^2 + 0.000302925 * D^2 \quad (8)$$

As given in equation 8, For CO emissions, the positive coefficient of BFR indicated that higher biogas flow leads to increased CO emissions, likely due to incomplete combustion at high biogas concentrations. The negative coefficients for engine load, BBR, and NDR suggest that higher load, biodiesel blend ratios, and NDR reduce CO emissions by improving combustion efficiency. The most sensitive interactions for CO emissions were observed between load and NDR, as well as load and BBR. Compared to the BBR20-NDR100 blend, CO emissions decreased by 14% with BBR20-NDR0, but increased by 11.9% when comparing BFR1-NDR100 to BFR0.5-NDR100.

Materials Research Proceedings 49 (2025) 571-580 https://doi.org/10.21741/9781644903438-57

$$SO = 28.8435 + 13.2206 * A + -2.02695 * B + -1.2147 * C + -4.41904 * D + -0.924537 * AB + -0.553756 * AC + -2.00184 * AD + 0.0855113 * BC + 0.306983 * BD + 0.184264 * CD + -0.127199 * A^2 + -0.458502 * B^2 + -0.300336 * C^2 + 1.33338 * D^2 \quad (9)$$

As given in equation 9, Finally, for Smoke Opacity (SO), the presence of positive coefficients for engine load and negative coefficients for BFR, BBR, and NDR indicated that SO increased with higher load but decreased with higher biogas flow, biodiesel blend ratio, and nanoparticle doping rate. The combination of load with BFR showed the most significant effect on SO, as increasing both parameters led to increased smoke opacity. Compared to the BBR20-NDR100 blend, SO decreased by 13% for the BBR20-NDR0 blend, but increased by 26.15% when comparing BFR1-NDR100 to BFR0.5-NDR100.

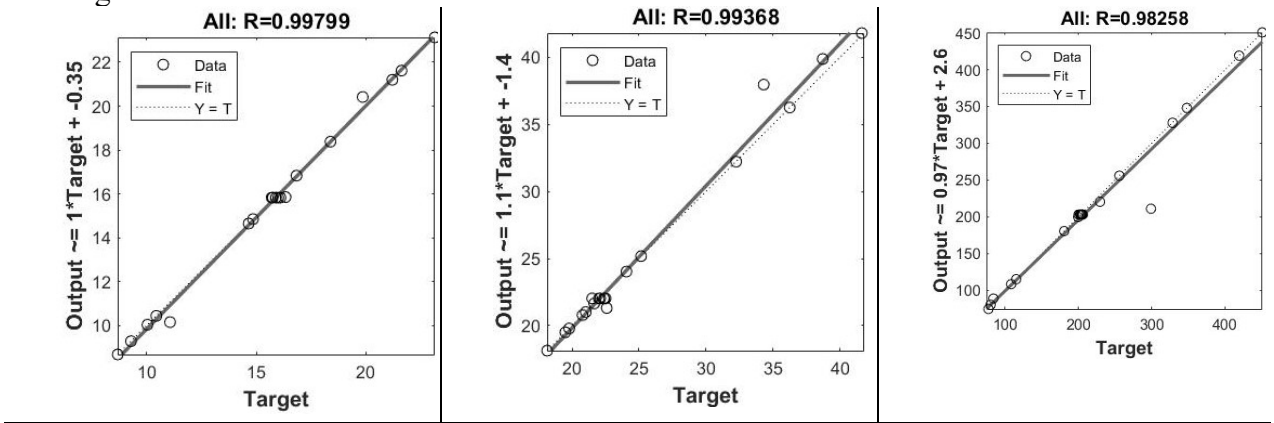
From the ANOVA analysis, high F-values (Table 4) confirm that the model effectively explains the variation in experimental data. Engine load had the highest F-value, followed by BFR, BBR, and NDR, indicating that engine load has the most significant influence on the output responses. The high R^2 values for all responses suggest the model closely matches the experimental data. ANOVA analysis also showed that the lack of fit is insignificant with P-values greater than 0.05.

Table 4 *Analysis of variance (F-value) and R^2 for performance and emission measurements*

Source	BTE	BSEC	NO_X	HC	CO	SO
	F-value	F-value	F-value	F-value	F-value	F-value
Model	2464.5	6844.83	14491.97	4697.42	19054.97	16152.16
A-Load	25502.84	69586.09	1.77E+05	34762.25	1.13E+05	1.93E+05
B- NDR	849.02	2218.54	3068.52	14999.37	22323.99	4534.23
C-BBR	1710.93	6296.3	2370.26	5896.73	10541.33	1628.36
D-BFR	3139.37	3231.31	13767.7	6750.58	13183.63	21551.19
R^2	0.9996	0.9995	0.9991	0.9998	0.9999	0.9999

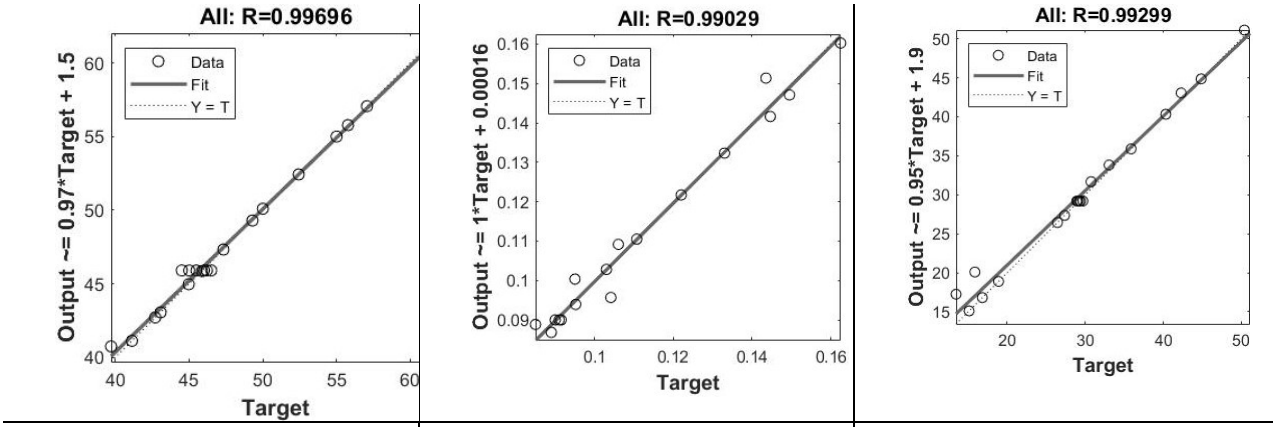
ANN Model Analysis

The network's output parameter-corresponding correlation coefficient (R) was computed for each phase of training, testing, and validation. You can see the total correlation coefficients for all the dependent responses in Figures 3 and 4. The high R values observed across all networks indicate that the data was effectively trained using the DOE, reinforcing the reliability and robustness of the regression model.



(a) BTE (b) BSEC (c) NO_X

Fig. 3 *R as a function of (a) BTE, (b) BSEC, and (c) NO_X as inputs to the trained network*



(d) HC (e) CO (f) Smoke opacity

Fig. 4 *R as a function of (d) HC, (e) CO, and (f) SO as inputs to the trained network*

Optimization of RSM and ANN Model

The RSM model exhibiting the highest desirability of 0.71 (Fig. 5) identified an optimal solution that predicts the following values: BTE at 16.58%, BSEC at 20.28 MJ/kWh, NO_X at 239.704 ppm, HC at 41.06 ppm, CO at 0.085%, SO at 27.2%. These predictions correspond to an engine load of 67.45%, NDR of 96.06 ppm, BBR of 10.7%, and BFR of 0.85 kg/h.

The ANN-HHO optimization process concluded after 102 generations, yielding optimal values for BTE, BSEC, NO_X, HC, CO, and SO of 16.87%, 19.76 MJ/kWh, 234.85 ppm, 41.47 ppm, 0.0857%, and 27.26%, respectively, at an engine load of 67.01%, NDR of 98.39 ppm, BBR of 8.41%, and BFR of 0.846 kg/h.

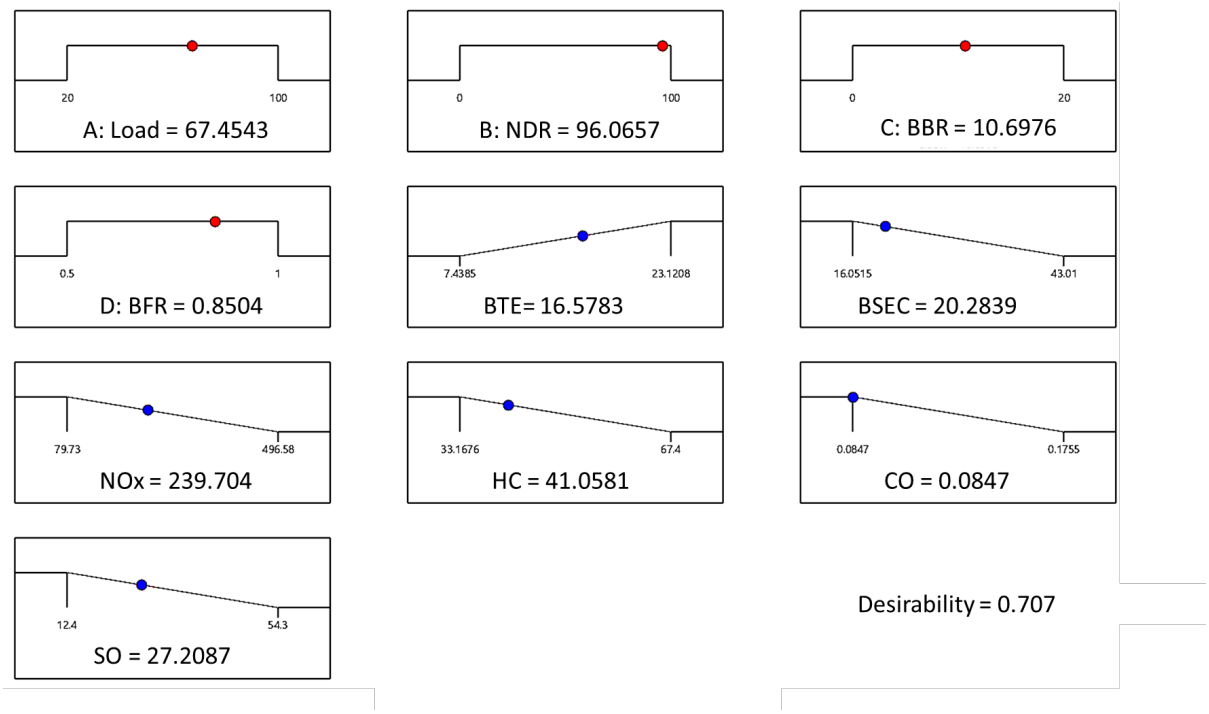


Fig. 5 *Desirability Plot*

Conclusions

This study yields the following conclusions:

1. The F-value of engine outputs derived from the RSM model indicates that, among the input variables, engine load exerts the most significant influence on the output responses, followed by BFR, NDR, and BBR in that order.

2. The evaluation metrics indicate low prediction error and high model performance for both regression analysis RSM and ANN, with a higher R2 suggesting that RSM is more accurate and reliable.
3. RSM optimization indicated an optimal engine output response at 67.45% engine load, 96.06 ppm NDR, 10.7% BBR, and 0.85 kg/h BFR. The optimization of the ANN-GA model revealed an optimal engine output response at 67.01% engine load, 98.39 ppm NDR, 8.41% BBR, and 0.846 kg/h BFR. Based on the optimized results from both the RSM and ANN models, it can be concluded that the combination of nanoparticles is effective.

The integration of biogas and biodiesel demonstrates promise in achieving reduced greenhouse gas emissions while maintaining engine performance. The application of optimization tools such as the HHO algorithm enhances the decision-making process for engine parameter tuning. This approach paves the way for adopting sustainable fuel technologies in CI engines, contributing to global efforts in reducing reliance on fossil fuels.

If you follow the “checklist” your paper will conform to the requirements of the publisher and facilitate a problem-free publication process. Thank you for your help in this matter.

References

[1] Ullah Khan I, Hafiz Dzarfan Othman M, Hashim H, Matsuura T, Ismail AF, Rezaei-DashtArzhandi M, et al. Biogas as a renewable energy fuel – A review of biogas upgrading, utilisation and storage. Energy Convers Manag 2017;150:277–94. https://doi.org/10.1016/j.enconman.2017.08.035.

[2] Huang D, Zhou H, Lin L. Biodiesel: an Alternative to Conventional Fuel. Energy Procedia 2012;16:1874–85. https://doi.org/10.1016/j.egypro.2012.01.287.

[3] Hatami M, Hasanpour M, Jing D. Recent developments of nanoparticles additives to the consumables liquids in internal combustion engines: Part I: Nano-fuels. J Mol Liq 2020;318:114250. https://doi.org/10.1016/j.molliq.2020.114250.

[4] Mei D, Zuo L, Adu-Mensah D, Li X, Yuan Y. Combustion characteristics and emissions of a common rail diesel engine using nanoparticle-diesel blends with carbon nanotube and molybdenum trioxide. Appl Therm Eng 2019;162:114238. https://doi.org/10.1016/j.applthermaleng.2019.114238.

[5] Yusof SNA, Sidik NAC, Asako Y, Japar WMohdAA, Mohamed SB, Muhammad NM. A comprehensive review of the influences of nanoparticles as a fuel additive in an internal combustion engine (ICE). Nanotechnol Rev 2020;9:1326–49. https://doi.org/10.1515/ntrev-2020-0104.

[6] Sharma A, Singh Y, Ahmad Ansari N, Pal A, Lalhriatpuia S. Experimental investigation of the behaviour of a DI diesel engine fuelled with biodiesel/diesel blends having effect of raw biogas at different operating responses. Fuel 2020;279:118460. https://doi.org/10.1016/j.fuel.2020.118460.

[7] Rajendran S, Venkatesan EP, Dhairiyasamy R, Jaganathan S, Muniyappan G, Hasan N. Enhancing Performance and Emission Characteristics of Biodiesel-Operated Compression Ignition Engines through Low Heat Rejection Mode and Antioxidant Additives: A Review. ACS Omega 2023;8:34281–98. https://doi.org/10.1021/acsomega.3c03252.

[8] Feroskhan M, Ismail S, Natarajan G, Manavalla S, Khan TMY, Khadar SDA, et al. A Comprehensive Study of the Effects of Various Operating Parameters on a Biogas-Diesel Dual Fuel Engine. Sustainability 2023;15:1232. https://doi.org/10.3390/su15021232.

[9] Veza I, Spraggon M, Fattah IMR, Idris M. Response surface methodology (RSM) for optimizing engine performance and emissions fueled with biofuel: Review of RSM for sustainability energy transition. Results in Engineering 2023;18:101213. https://doi.org/10.1016/j.rineng.2023.101213.

[10] Hadiyat MA, Sopha BM, Wibowo BS. Response Surface Methodology Using Observational Data: A Systematic Literature Review. Applied Sciences 2022;12:10663. https://doi.org/10.3390/app122010663.

[11] Agatonovic-Kustrin S, Beresford R. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. J Pharm Biomed Anal 2000;22:717–27. https://doi.org/10.1016/S0731-7085(99)00272-1.

[12] Zou J, Han Y, So S-S. Overview of Artificial Neural Networks, 2008, p. 14–22. https://doi.org/10.1007/978-1-60327-101-1_2.

[13] Lalhriatpuia S, Pal A. Computational optimization of engine performance and emission responses for dual fuel CI engine powered with biogas and Co3O4 nanoparticles doped biodiesel. Fuel 2023;344:127892. https://doi.org/10.1016/j.fuel.2023.127892.

[14] Plevris V, Solorzano G, Bakas N, Ben Seghier M. Investigation of performance metrics in regression analysis and machine learning-based prediction models. 8th European Congress on Computational Methods in Applied Sciences and Engineering, CIMNE; 2022. https://doi.org/10.23967/eccomas.2022.155.

[15] Costa NR, Lourenço J, Pereira ZL. Desirability function approach: A review and performance evaluation in adverse conditions. Chemometrics and Intelligent Laboratory Systems 2011;107:234–44. https://doi.org/10.1016/j.chemolab.2011.04.004.

[16] Lalhriatpuia S, Pal A. Computational optimization of engine emissions and performance of a CI engine powered with biogas and NiO nanoparticles doped diesel. Environ Prog Sustain Energy 2023. https://doi.org/10.1002/ep.14207.

[17] Budhraja N, Pal A, Mishra R. S. Various Treatment Methods for Enhanced Bio-Hydrogen Production, Materials Today: Proceedings, March 2023. https://doi.org/10.1016/j.matpr.2023.03.198

[18] Budhraja N, Pal A, Mishra R. S. Simulation and Optimization of Biohydrogen Production from Biomass Feed via Anaerobic Digestion, Chemical Engineering & Technology, 47. https://doi.org/10.1002/ceat.202300395.

[19] Budhraja N, Pal A, Mishra R. S. Simulation and Optimization for Biohydrogen Production Potential of Various Organic Waste via Anaerobic Digestion, Fuel, 360, https://doi.org/10.1016/j.fuel.2023.130563.

[20] Heidari AA, Mirjalili S, Faris H, Aljarah I, Mafarja M, Chen H. Harris hawks optimization: Algorithm and applications. Future Generation Computer Systems 2019;97:849–72. https://doi.org/10.1016/j.future.2019.02.028.

[21] Alabool HM, Alarabiat D, Abualigah L, Heidari AA. Harris hawks optimization: a comprehensive review of recent variants and applications. Neural Comput Appl 2021;33:8939–80. https://doi.org/10.1007/s00521-021-05720-5.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

NoCiPUF: NoC-Based Intrinsic PUF for MPSoC Authentication

Deepank Grover, *Graduate Student Member, IEEE*, Tarun Sharma, *Graduate Student Member, IEEE*, Sneha Agarwal, *Graduate Student Member, IEEE*, Sidhartha Sankar Rout, Anushka, Madhur Kumar, and Sujay Deb, *Senior Member, IEEE*

***Abstract*— Modern Multi-Processor-Systems-on-Chips (MPSoCs) use Network-on-Chips (NoCs) as a scalable and efficient communication fabric. The applications running on these devices rely on frequent communication with central database servers, which are vulnerable to impersonation attacks by adversarial clones. We propose NoCiPUF, a novel NoC-based intrinsic Physically-Unclonable-Function (PUF) framework for MPSoCs authentication. We re-use the circuit switched nature of NoC with path-pairs as challenges to obtain secret responses, collectively called challenge-response-pairs (CRPs). Due to the random nature of manufacturing variations, equal hop paths exhibit unequal delays. We leverage the delay differences of flits traversing in equal-hop paths to generate unique responses. NoCiPUF is fully-synthesizable and readily scalable as it requires changes only at the behavioral level. To counter Machine-Learning (ML)-based modeling attacks on PUFs, we provide a comprehensive technique and reduce the prediction accuracy to ∼52%. NoCiPUF framework incurs low area (0.76%) and power (1.14%) overheads and has no impact on NoC performance in normal mode due to independent authentication mode. Obtained responses have near-ideal PUF metrics and are verified against the NIST randomness test suite. This scheme offers high number of CRPs in larger NoC networks (> 0.74 million CRPs in 5 × 5 mesh), proving its scalability.**

***Index Terms*— MPSoC authentication, physically-unclonable-function (PUF), network-on-chip (NoC), challenge-response-pairs (CRPs), on-chip variations, ML-resilience.**

I. Introduction

CONVENTIONAL Systems-on-Chip (SoCs) have evolved into multi-processor SoCs (MPSoCs), incorporating a multitude of processing elements to serve a diverse set of applications on a single chip. NoC serves as an efficient and scalable intra-core shared communication fabric in state-of-the-art MPSoCs, supporting up to hundreds of peripherals. The scalability, high bandwidth and design flexibilities offered by NoC has made it the popular industrial choice over bus-based interconnects [1], [2].

The NoC supported modern MPSoCs are being deployed in various domains ranging from smartphones to automotive grade applications, often dealing with sensitive data.

Received 11 August 2024; revised 9 December 2024 and 10 March 2025; accepted 2 April 2025. This article was recommended by Associate Editor C.-T. Cheng. *(Corresponding author: Deepank Grover.)*

The authors are with the Department of Electronics and Communication Engineering, Indraprastha Institute of Information Technology Delhi, New Delhi 110020, India (e-mail: deepankg@iiitd.ac.in; taruns@iiitd.ac.in; snehaa@iiitd.ac.in; sidharthas@iiitd.ac.in; anushka21187@iiitd.ac.in; madhur19171@iiitd.ac.in; sdeb@iiitd.ac.in).

Digital Object Identifier 10.1109/TCSI.2025.3559419

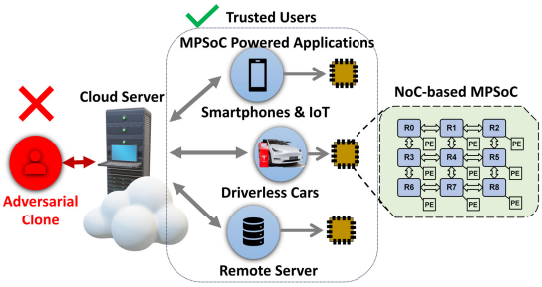


Fig. 1. Need for authentication of MPSoCs.

They primarily run applications that require heavy computations and frequent communication with central data servers to exchange and store user-critical information. Moreover, this type of communication among distributed remote servers, catering to a larger, central database is bound to rise in near future. However, these central databases are vulnerable to unauthorized access via adversarial clones impersonating the original MPSoC device, to access confidential information or flood the server with malicious data [3].

As shown in Fig. 1, in the absence of some authentication primitive, any adversarial (or malevolent) device similar to trusted devices can easily gain access to the central server. This can lead to loss of confidential user data and severely impacts the server performance. The report [4] lists multiple data breaches due to authentication vulnerabilities in systems of tech giants like Change Healthcare causing $870 million loss in first quarter of 2024 and affecting more than 100 million customers. A recent example of an attack is [5], in which an "unauthorized party" obtained access to "basic personal information" stored in NXP's online server(s). The described threat model makes the argument for the need to verify if a non-trusted or cloned MPSoC is trying to access or exchange data with the server. Even though several schemes deploy hash and ciphers to generate public/private keys, their complex nature coupled with low speed and area intensive nature makes them a costly choice. The existing MPSoCs are already complex enough. This necessitates the need for a scheme that is easily integrable in MPSoCs and has low area, power and performance overheads. One such scheme is Physically-Unclonable-Functions (PUFs).

The PUFs are popularly used by parent servers to authenticate in-field devices against adversarial clones. PUFs rely on randomly infused manufacturing (process) variations in CMOS

1549-8328 © 2025 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:49:13 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

circuits to harness secret keys (responses) corresponding to certain inputs (challenges) [6]. The absence of additional on-device on-chip storage for secret keys and security key management makes PUF more attractive compared to their conventional cryptography-based counterparts [7], [8], [9]. Extrinsic PUFs require dedicated on-chip placement, making them more susceptible to side-channel and layout recognition attacks [10], [11]. However, intrinsic PUFs re-use the existing on-chip circuitry to extract the responses. For example, the SRAM PUFs use power-on states to extract the PUF responses, requiring limited changes to existing circuit [12]. They incur lower area and power overheads as only few changes are required in original design [13]. Therefore, we focus on intrinsic PUFs in this work.

In the domain of NoCs, existing router-crossbars can be utilized to realise intrinsic PUFs [14]. But such schemes often assume and work on specific gate-level implementation of various router blocks and require transistor and layout level changes, rendering them non-generic. Such modifications are not directly synthesizable as they can not be specified at behavioral level and require intervention in standard cell flow, leading to increased design effort, flow-complexity, and turn-around-time. Synthesized gate-level netlists can also be used to filter approximately equal gate-delay paths to realise intrinsic PUFs [15]. But this approximation not only adds a default delay-based bias in PUF responses but filtering only quality CRPs makes it computationally heavy and reduces set of usable CRPs. Moreover, scaling and technology-portability challenges will persist as the netlists size grow further for modern NoC-MPSoCs. Also, these schemes are silent over ML-based modeling attacks which are quite common in the domain of PUFs [16]. A detailed qualitative comparison with these schemes has been provided in Section V. Hence, these shortcomings necessitate the need for a more generic scheme for realising intrinsic PUFs with minimal impact on the power, performance and area of NoC-MPSoCs. Such scheme should be readily compatible with conventional standard-cell flow and scalable with increasing NoC size. In current work, we solve aforementioned challenges by exploiting the fundamentally symmetric multi-hop paths in NoC to our benefit, to extract the secret responses.

We propose NoCiPUF, a novel, fully-synthesizable *NoC-based intrinsic-PUF* framework for MPSoC authentication. Our scheme requires only behavioral-RTL level architectural changes in baseline NoC-routers, which eliminates requirement of manual intervention in existing standard-cell flow for NoCs. By limiting our scheme to RTL level modifications, technology portability becomes extremely easy and the framework can be scaled to any N×N NoC size using these minor changes in the router. NoCiPUF makes use of equal nominal delays corresponding to equal hop paths in a given NoC. The comparison of flit arrival times (delays) of such equal-hop paths at a common destination in an asynchronous setting yields unpredictable outputs encoded as PUF responses, unique to a given chip. The random distribution of manufacturing process variations across NoC among router components and interlinks [17] are responsible for the variations in otherwise nominally equal delays. Asynchronous flit-transfer during circuit switching [18], [19] preserves inherent delays of selected paths. These path-pairs constitute the challenges.

However, merely suggesting the re-use of the existing circuit does not provide the comprehensive view of how the modified system will work with the added functionality of intrinsic PUF [14]. In other words, given the complex nature of NoC-MPSoCs, questions like where do the PUF-enabled authentication operation fit in with the existing system operations, how do the challenges using the existing circuit blocks get applied or what exactly constitutes the challenges have to be answered whenever an intrinsic PUF is suggested. In this paper, we answer these questions by, 1) introducing an independent authentication mode in NoCiPUF besides its normal functioning mode, 2) providing elaborate algorithms to determine CRP space for given NoC size at design time, and 3) providing a systematic strategy to define and apply challenges for generating unique responses. Therefore, to the best of our knowledge, NoCiPUF is the first one to offer this degree of completeness which provides an intrinsic PUF framework for NoC based MPSoC authentication.

Intrinsic nature of the proposed PUF in NoC offers an inbuilt protection against discovery using side-channel attacks, reverse engineering, layout-recognition or modeling attacks [10], [11], [20]. However, the existing PUF implementations have been proved to be vulnerable to ML-based modeling attacks [16], with CRP prediction accuracy up to $\sim$ 99%. Therefore, we investigate the ML-resilient nature of NoCiPUF's CRPs. We find that the proposed framework inherently offers an error in the prediction accuracy of about $8\% - 18\%$. To further reduce the prediction accuracy, we suggest an ML resilient technique for state-of-the-art ML models used in PUF domain, i.e., LR (Logistic regression), SVM (Support-vector-machine), CMA-ES (Covariance matrix adaptation evolution strategy), and DNN (Deep neural network). We reduce prediction accuracy to $\sim$ 52%, which is close to ideally accepted value of 50%. This technique results in area and power overheads of 0.76% and 1.14%, respectively. These overheads are very low, justifying that the suggested ML-resilience technique is lightweight.

The framework is applicable for any NoC-based device deployed in-field every time the parent server needs to authenticate it prior to granting access to the server database. The NoCiPUF offers a wide range of challenge path-pair combinations, which is indicated by the CRP space computed for different NoC sizes. This ensures that the framework can function even if a particular node is faulty, as many other alternative challenge paths can be utilised in such cases.

Fig. 2 highlights the general authentication flow, with steps numbered (in circles) in the figure. Upon receiving a request from an in-field NoC-based MPSoC, the server first throws a challenge (CN_M) to the MPSoC (SoC_M; step 1.) and then the flit-pairs are sent over two equal-hop circuit-switched paths (step 2.). Upon reaching the comparator, a response (RN_M) is generated (step 3.) based on the path-delay comparison, and is sent to the server (step 4.). This response is then compared against the corresponding CRP entry in the parent-server's LUT. Our work is aimed at providing the on-chip NoC framework as required to support intrinsic-PUF operation.

The key contributions of this work are summarized as follows:

- We propose NoCiPUF, a fully-synthesizable and scalable NoC-based intrinsic PUF framework for MPSoC authentication which exploits unpredictability of flit-arrival

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:49:13 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

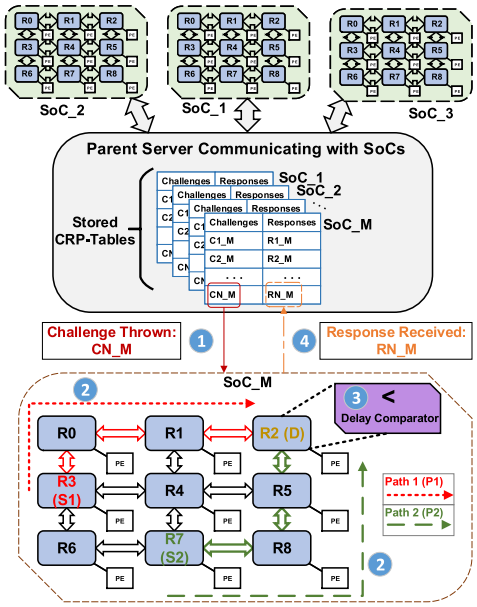


Fig. 2. A parent server authenticating the SoCs connected to it. Challenge (CN_M) is being sent to SoC_M. NoCiPUF framework is being used to generate response RN_M using P1 and P2.

delays owing to symmetric nature of multi-hop paths in NoC to encode PUF responses.

- NoCiPUF framework provides independent operation modes for authentication and normal working of NoC using circuit and packet switching strategies.
- Our framework involves all changes at behavioral level, making it directly compatible with standard ASIC design flow.
- We also investigate the ML-resilient nature of suggested framework and reduce the accuracy of response prediction to acceptable levels ($\sim$ 52%) using an on-chip lightweight post-processing technique.
- The intrinsic nature of proposed framework results in low area (0.76%) and power (1.14%) overheads.
- Monte-Carlo analysis has been performed to measure the PUF metrics, yielding 51.24% (31.26-bits) uniqueness and 53.3% uniformity. The framework achieved 96.95% and 97.34% reliabilities against temperature and voltage respectively. Responses are further validated with NIST randomness test suite.

The rest of the paper is organized as follows: Section II provides the background to PUFs, basics of NoC and switching strategies, and some authentication aspects of MPSoCs. Section III elaborates the hardware changes of proposed framework, algorithm used to find challenge path-pairs, and applied technique for ML-resilience. Section IV provides implementation details of the proposed methodology, discusses simulation results of intrinsic PUF parameters achieved using NoC, and reports the associated hardware overheads, followed by Section V which provides a detailed qualitative comparison of our work with related works. The last Section VI concludes this work.

II. Background

Before describing the proposed framework, it is important to understand various elements contributing to this work.

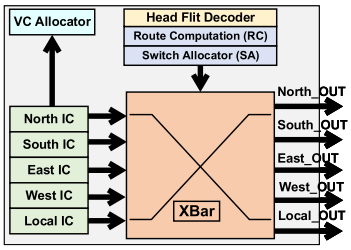


Fig. 3. Baseline NoC router architecture.

In this section, we briefly elaborate on the concept of PUFs and their intrinsic nature, followed by an overview of a baseline NoC, and packet and circuit switching strategies. We also investigate and highlight the state-of-the-art works in the domain of SoC authentication.

A. Intrinsic PUFs and Quality Metrics

PUFs use manufacturing process variations to generate a set of unique fingerprints (responses) corresponding to a set of inputs (challenges) from integrated circuits, in the form of CRPs. Intrinsic PUFs use existing circuitry available in the design to generate CRPs. The state-of-the-art intrinsic SRAM PUFs [12] have offered better reliability and energy efficiency as compared to their conventional counterparts. The intrinsic nature as a part of a larger circuit makes them harder to locate and be detected on the chip, making them a popular choice [15]. The effect of manufacturing process variations is reflected across NoC-routers (among various circuit blocks in them) and wires (interlinks), inherently affecting the multi-hop path delays.

For a given PUF, its CRPs should satisfy some measurable quality metrics at *inter*-chip and *intra*-chip level, to qualify for its practical feasibility. The *uniformity* (an intra-chip metric) indicates how varied the response bits are for different challenges. 50% uniformity indicates that the response bits are not biased towards either bit, 0 or 1. Its value is calculated by averaging the binary bits $r_{i,j}$ of the over m, n-bit responses obtained from the given chip, using the following relation:

$$Uniformity = \frac{1}{m}\sum_{i=1}^{m}\left(\frac{1}{n}\sum_{j=1}^{n} r_{i,j}\right) \times 100\% \qquad (1)$$

where $r_{i,j}$ represents j^{th} response bit corresponding to the ith challenge.

The *uniqueness* (an inter-chip metric) is used to measure the amount of variation in the response bits across different chips for the same challenge. It helps to distinguish a chip uniquely from other chips having same type of PUF implemented. The ideal value is 50% and it implies lower probability of repetition of responses among multiple chips. It is found by averaging the hamming distance (HD) of n-bit responses (R_i,R_j) among k chips using by the following relation:

$$Uniqueness = \frac{2}{k(k-1)}\sum_{i=1}^{k-1}\sum_{j=i+1}^{k}\frac{HD(R_i, R_j)}{n} \times 100\% \qquad (2)$$

The *reliability* (an intra-chip quantity) is an indicator of the bit-error-rate (BER), i.e., the variation in response bits

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:49:13 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE I
CLASSIFICATION OF RELATED WORKS

Work	{Intra-NoC/System} Level	{Pre/Post}-deployment
[21]	Intra-NoC	Post-deployment
Sentry-NoC [22]	Intra-NoC	Post-deployment
SP-PUF [15]	System (IP)	Pre-deployment
POCA [23]	System (Chip)	Pre-deployment
PROTECTS [24]	System (Chip)	Pre-deployment
[14]	System (IP)	-
This Work	System (IP)	Pre-deployment

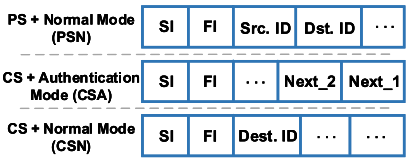


Fig. 4. Head-flits for different modes of operation in NoCiPUF.

obtained against identical challenge given to the same chip under different environmental conditions. Intra-chip HDs for responses at different environmental conditions *env* (voltage or temperature variations) are averaged using following relation to find the corresponding BER:

$$HD_{INTRA} = \frac{1}{m}\sum_{env=1}^{m}\frac{HD(R_i, R_{i,env})}{n} \times 100\% \quad (3)$$

where R_i represents the response obtained at nominal voltage and temperature, $R_{i,env}$ represents the n-bit response at a temperature or voltage value of env, for i^{th} chip, and m is the total number of samples recorded (at different voltages and temperatures). Reliability is defined to be $(100\% - HD_{INTRA})$. Ideal reliability should be 100% when the response bits to a corresponding challenge do not change across varied environmental conditions. We realize the PUF functionality and measure the above metrics to establish feasibility of proposed framework. Choice of number of hops, path pairs and size of network determine the size of CRP space (summarized in Section III.)

B. NoC: Basic Working and Switching Strategies

Baseline NoC consists of various units: switch allocator (SA), route computation (RC), cross bar (XB), virtual-channel (VC) allocator, flit decoder (as per requirement), input controller (IC) (inclusive of input buffers), as shown in Fig. 3. These components are generally modified as per the design requirement. The RC stage computes the next node or identifies if the current node is the final destination. The SA stage then configures the XB to divert the incoming flit to correct directional output port. Then the flit passes through XB stage and is received by the input buffer (inside the IC) of the next node with allocated virtual channel (decided using VC allocator). If the current router is destination of corresponding flit, the XB is configured to divert it to local output (Local_OUT in Fig. 3).

We consider commonly adopted packet-switching (PS) with XY routing for the baseline router. In the PS, the RC stage calculates the next node for the current packet based on the

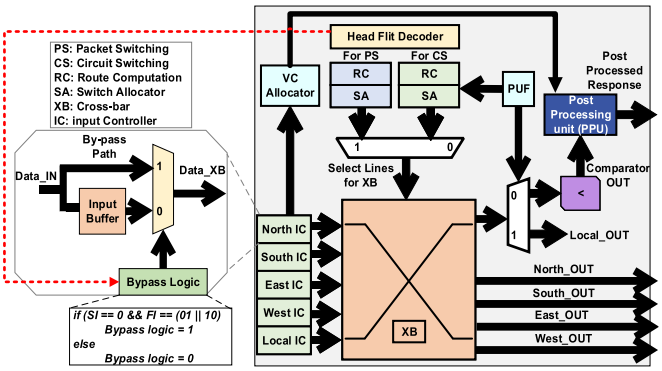


Fig. 5. Proposed router architecture.

source and destination node ID present in head-flit (HF) using XY routing scheme. HF, followed by body-flit (BF) and tail-flit (TF) of this packet propagate in a wormhole manner after XB is configured. Another switching scheme called circuit-switching (CS) is also prevalent in the domain of NoC [25]. Unlike PS, CS first reserves the path from source to destination using the HF and then the BF and TF traverse through this set-up path in one go. Some earlier works have leveraged the power and performance benefits of CS and proposed hybrid switching based NoC solutions [18], [19], [26].

C. Authentication Aspects of MPSoCs

Authentication in MPSoCs has been addressed in multiple aspects. These works target two different abstraction levels: at intra-NoC level (secure communication among the routers) and at MPSoC as a whole-system level (against third party or off-chip adversary). Some of these schemes are applicable when the SoC is being used in field (post-deployment) and some deal with authenticating the SoC still at foundry or testing facility (pre-deployment). We have classified these works accordingly in Table I.

For the context of this paper, we discuss some works which address SoC-authentication against external entities. The work [15] offers a boolean-satisfiability solver (SAT) based solution to use DFT-sensitizable combinational paths as PUFs in conventional benchmark circuits. They offer multi filter and post processing approach to select PUF-feasible paths from multitude of available gates. The work [27] uses watermarked paths inside NoC to verify if the IP with buyer has been sold by correct seller. If the buyer suspects any malicious activity in the system using this IP or there is a need to track illegal distribution, watermark identification is carried out. The works [23], [24] target chip authentication against suspicious activities at third-party level, like foundry or compromised testing facility. They rely on assumption that, the design house is able to establish a communication with the chip prior to deployment while still at foundry or testing facility to authenticate it. The work [14] attempts to construct PUFs out of individual router cross-bars in NoC. They modify NoC cross-bar and provide simulation results to validate the quality metrics.

Our work provides the authentication framework applicable to an NoC-based MPSoC deployed in-field by generating secret keys using NoC paths' inherent delays. We elaborate the corresponding routing and switching schemes, modes of

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:49:13 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

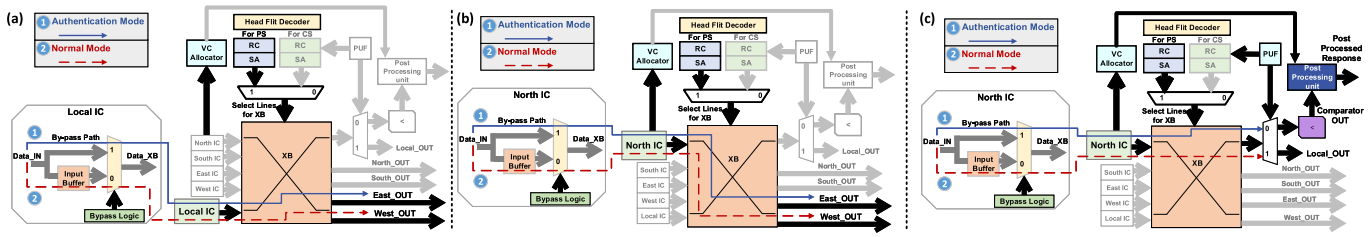


Fig. 6. Flow of BF and TF inside router during normal and authentication modes when router is either, (a) source, (b) intermediate or, (c) destination node.

operation (authentication and normal), along with minor modifications in existing routers, making the framework feasible for any given NoC and can be easily scaled up. Additionally, as both circuit-switching and packet switching are enabled, the SoC can use either type of switching as per requirement.

III. The NoCiPUF Framework

In a NoC, data packets can be sent from a source router (S) to a destination router (D) via different paths. Packets are typically split into flits as head-flit (HF), body-flit (BF) and tail-flit (TF). HF is mainly responsible for routing and other flags, BF carries the payload for the application and TF marks the end of packet. A path is a sequence of routers a flit would need to traverse starting at S and ending at D. Considering two different paths P1 (between source S1 and D) and P2 (between source S2 and D), if P1 and P2 have equal hop count (H), then ideally, flits launched from S1 and S2 should reach D after an identical delay. However, because of manufacturing variations, the delays along both paths differ. In the proposed scheme, a multi-bit delay-comparator is used at D to enable parallelism to determine which flit arrives first. This choice of a path pair PP (P1, P2) in the network forms the NoCiPUF's challenge space. We describe the behavioral changes required in the NoC router to support the PUF based authentication operation. The modified router is the fundamental unit of the NoC in our experimental setup. The multi-bit interconnects between routers are of equal-length and require no modifications. This approach ensures a physically symmetric NoC structure by the virtue of identical routers and interconnects throughout the NoC. The following subsection III-A discusses the working of NoCiPUF framework while elaborating corresponding modifications required in the router hardware for authentication. The subsection III-B elaborates the algorithms used to find NoC-paths suitable for PUF operation and reports the CRP space with respect to number of hops and NoC mesh size.

A. Authentication Scheme and the Modified Router

To support the authentication operation in addition to regular data-flow, we classify data transfer in the proposed framework among three modes of operation: 1) CS for authentication mode (CSA), 2) PS for normal mode (PSN), and 3) CS for normal mode (CSN). PS and CS refer to packet and circuit switching respectively. In order to differentiate among these modes, we define three types of header-flits as shown in Fig. 4. All the three types of HF have 3-MSBs dedicated to identify the type of switching (1st bit: switching identifier, SI) between CS (SI=0) or PS (SI=1) and type of flit (2nd and 3rd bits: flit-identifier, FI). Table II shows the significance of various possible combinations of the 3-bits.

TABLE II
Switch and Flit Indicator Bits in the Head Flit

SI bit (1)	FI bits (2)	Mode	Flit Type
0	00	CSN Mode	HF
0	01	CSN or CSA mode	BF
0	10	CSN or CSA mode	TF
0	11	CSA mode	HF
1	00	PSN mode	HF
1	01	PSN mode	BF
1	10	PSN mode	TF
1	11	NA	NA

The baseline router is modified to support these modes and flits. The modified router architecture required to support the proposed framework is shown in Fig. 5. This router architecture contains separate route-computation (RC) and switch-allocation (SA) units to support hybrid switching (PS and CS), a PUF unit to identify common destination corresponding to challenge (explained later), modified input-controller-IC with bypass path, and a multi-bit comparator unit to generate response bits parallely. The head-flit decoder is used to switch between the NoC's authentication (CSA) and normal (PSN or CSN) modes of operations and to differentiate among the flits corresponding to the three modes. The SI and FI bits in the HF (Fig. 4) govern this decision. The bypass logic inside the IC as shown in Fig. 5 receives the SI and FI bits as inputs from the head-flit decoder to switch the MUX above it. Bypass logic selects by-pass path for BF and TF during CSA and CSN whereas the buffered path is selected for the HF in all modes (CSA, PSA, CSN), for BF and TF in normal mode (PSN). The bypass logic is summarized using if-else statement in the Fig. 5 to indicate how the input controller switches between different modes of NoC operation.

The Fig. 6 illustrates the datapath taken by the BF and TF during normal (here, PSN) and authentication modes (CSA) of operations. We choose PSN as normal mode for this illustration because it uses conventional packet-switching in a typical baseline NoC. It is assumed in this illustration that the HF has already switched the XBar (XB) for the upcoming flits. For simplicity, a convention of East and West output ports for CSA and PSN modes respectively, is taken. Fig. 6a shows the flit starts from the local port if it is in source router node and crosses the XB and reaches the east output port. If it is PSN mode, then, buffered path inside the IC is chosen whereas it is bypassed during CSA mode. Flits flow similarly in any intermediate router, except they enter router through a directional node (North in Fig. 6b) unlike the source router. For destination router (Fig. 6c), the de-multiplexer transfers the flits to comparator in CSA mode and to local output node (connected to processing element) in PSN mode. The Fig. 7 further illustrates flow of the pair of flits during CSA mode

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:49:13 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

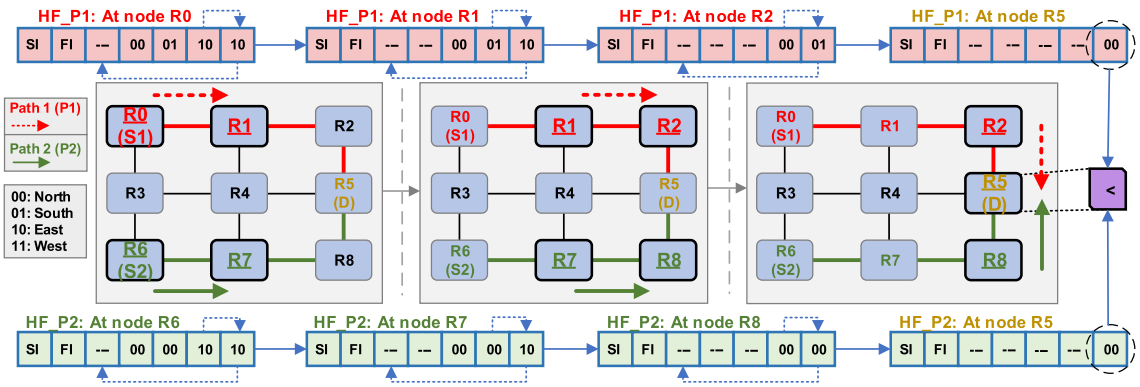


Fig. 7. Path set-up phase using simultaneous travel of two HFs along two equal hop (=3) paths. Updates in the HFs due to right shift has been shown. At destination node, both HFs point towards "00" (north) and hence XB switches to comparator output.

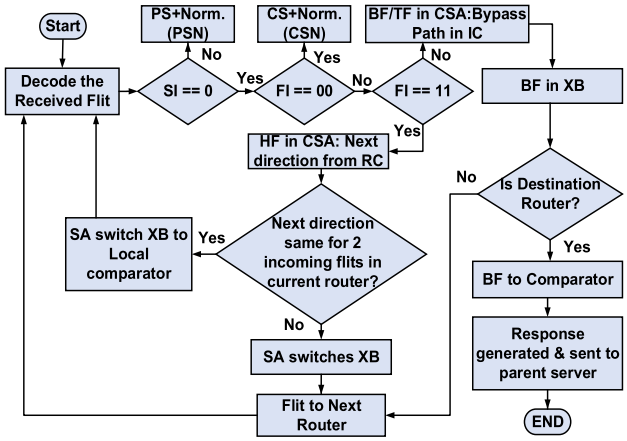


Fig. 8. NoCiPUF-router flow during authentication mode.

exclusively. The above discussion shows that the flits traveling in normal (PSN) mode of operation are not affected by the changes done for authentication (CSA) mode, implying no performance overheads during normal mode of operation.

As shown in Fig. 4 for CSA, rest of the HF contains bit-pairs (Next_1, Next_2, etc.) indicating the next directional node to be reserved to set-up the pre-encoded path. The two bits per pair, correspond to one of the four available directions. An example of directional mapping and flit-traversal in a 3-hop path-pair during authentication mode (CSA) can be observed in Fig. 7 and can be referred to visualize the framework at NoC level. To ensure design simplicity during the authentication mode, we use right-shift (while preserving SI and FI bits) to bring bit-pair for next router at LSB position in HF. The route-compute (RC) unit decodes the 2-LSBs to identify the direction of next router, and SA switches XB accordingly. This routing strategy is inspired from vector-based routing [28]. As seen in Fig. 7, the updates in HFs: HF_P1 (path, P1) and HF_P2 (path, P2) are indicated at the top and bottom of the figure, respectively.

Whenever the two HFs, launched from sources S1 and S2, arrive at a common node (D), the next bit-pairs (which were already pre-set to point to the same direction) flag this node as the destination. The dotted circles in HF_P1 and HF_P2 at node R5 in the Fig. 7 show '00' bits indicating both want north output port. PUF-unit sets select-line of DeMUX as '0' to divert upcoming flits to the comparator during CSA, and to '1' in other two modes. This is followed by BF and then TF un-reserves the path. The remaining two flits, BF and TF in path-pairs flow in the same manner, except there is no right shift involved as the IC bypasses other router components (except XB) to preserve the effect of process variations along the path.

Fig. 8 shows the proposed router's behavior during the CSA mode. After receiving the flit, the head-flit decoder identifies the flit-type and the mode. For SI = 0 and FI = 11, HF is identified and CSA mode is selected (Table II) and RC computes the next router node. The SA then configures the XB by generating select lines. The HF decoder also sets the MUX above XB to select input from CS based RC and SA stages. The last-bit-pair in the challenge HFs, is designed to ensure that both the HFs from S1 and S2 point to the same direction (in our case, north) as soon as they reach the common destination (D) node. After the destination node is identified, SA configures the XB to divert rest of the incoming flits to comparator using bypass paths in IC (right side of Fig. 8). This is followed by the BFs to flow asynchronously and concurrently in the reserved paths (Fig. 7). On arriving at the destination node, XB diverts both the BFs to the comparator to generate the final response bits. We exploit the complete BFs from the two paths as input to multi-bit comparator (of the same length as the flit) to generate the multi-bit response. This response is then sent to parent server to compare against the stored CRP entry present in it.

During the CSA mode, bypass paths in router IC are activated for the BFs and TFs. These bypass paths skip sequential elements inside the routers and provide asynchronous transfer, thereby preserving the much-needed unpredictability in delays due to manufacturing process variations (PVs). This avoids clocked latching and locking, preserving the expected PUF behavior. The comparator is implemented using parallel DFFs (D-flip-flops) equal to the required number of bits in each response. These DFFs receive the two body flits (from P1 and P2) as their clock and data inputs. Now, even though the nominal delays of equal hop paths should make BFs arrive exactly at the same time instant, PVs cause them to be delayed unpredictably. This eliminates the need to measure individual flit delays as whichever BF arrives earlier, dictates the output of the DFFs and hence the complete response. NoCiPUF framework ensures that all bits in BFs from S1 and S2 cause same transition i.e. 0 to 1 at the comparator inputs. The CSN mode only requires the destination ID to decide the next

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:49:13 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

router for path-setup, based on availability during regular data flow. For PSN, HF stores information regarding source and destination node IDs. HF goes to input buffer, followed by RC and then SA generates select line for other flits flowing in wormhole fashion. We have considered XY routing for PS mode but there is no restriction on any routing strategy with respect to our scheme.

We now describe the algorithms used for determining CRP space for the given NoC, where the challenges are defined in terms of path pairs.

B. The Challenge-Response-Pair Space

We aim to find all possible path pairs (P1, P2) for all possible triplets (S1, S2, D), in the given N×N mesh network, to calculate the available CRP space. The challenges are pre-calculated at the design-house itself, based on the size of NoC required, avoiding the need to generate them on the fly. This can be done first by finding all possible multi-hop paths using a depth-first search approach as presented in Algorithm 1, and then pairing equal hop non-overlapping paths with common destination using Algorithm 2. After all such path-pairs have been found, they are stored as challenges at the parent server, which are mapped to the PUF responses for different chips after the enrollment procedure at the chip-manufacturer's end.

In Algorithm 1, in the function *obtain_paths*, we define three data structures, $path_dict$, $visited_nodes$ and $direction_list$. The $path_dict$ stores paths sorted according to the number of hops, $visited_nodes$ stores the visited routers in the path and $direction_list$ contains the allowable directions for a router. For instance, in a mesh based network, only two directions are possible for corner nodes. When the destination router is reached (line 4), the taken path arranged by number of hops is returned (line 5 and 6). When path is being determined (lines 7-21), we consider if the node has been visited earlier (line 8-10). This is done to avoid overlapping in the path. We then consider all the possible directions a packet can take while leaving the present router. If the next router on the path has been visited already (line 13-15), we consider the router present in another direction. If the router has not been visited, it is appended to the $visited_nodes$ array and the above process is repeated till it reaches the destination router.

In Algorithm 2, we look at the flow to extract non overlapping path pairs from the set of all possible path-pairs. We aim to obtain all possible path pairs for the authentication flow as discussed in Section III-A. We consider two source routers and one destination router, such that all three are different (line 3). We calculate all possible paths between the source and destination router (line 4 and 5) using function *obtain_paths* as discussed in Algorithm 1 and traverse through them. Then, for a given source pair and destination nodes, if the number of hops are equal and no node is common amongst the two paths, (line 9), the resultant paths are marked to be used for PUF mode. Now these common destination, equal-hop path-pairs are ready to be stored at the parent server in CRP table.

The maximum number of hops (H_{max}) allowed is determined by the mesh size (N×N). Table III shows the possible number of path-pairs supported by the proposed router micro-architecture, corresponding to hops for different NoC-mesh sizes. This table is compiled using Algorithms 1 and 2. The

Algorithm 1 Finding Paths for CSA

Input: Source (S), Destination (D)
Output: Paths sorted according to hops
Function: $obtain_paths(S, D)$

1: $path_dict$ = {};
2: $visited_nodes$ = [];
3: $direction_list \leftarrow$ allowed directions for S router depending on its location in the mesh;
4: **if** $S == D$ **then**
5: $hops$ = length($visited_nodes$);
6: $path_dict$.append({$hops$, $visited_nodes$});
7: **else**
8: **if** this is the first iteration **then**
9: $visited_nodes$.append(S);
10: **end if**
11: **for** all directions in the $direction_list$ **do**
12: Calculate the $next_node$;
13: **if** $next_node$ is in $visited_nodes$ **then**
14: Go to the next iteration of this **for** loop;
15: **end if**
16: $visited_nodes$.append($next_node$);
17: $S = next_node$;
18: Call $obtain_paths(S, D)$;
19: $visited_nodes$.remove(S);
20: **end for**
21: **end if**
22: **return** $path_dict$

four possible directions for the flit traversal at every node in the chosen path are represented by two bits. If there are "H" hops in the path, it corresponds to "H×2" bit-space in the HF. In our experiment, each HF is 64-bits long. Leaving the three MSB SI and FI bits, the remaining 61 bits can support up to 30-hop paths (applicable for NoC size 6×6 and above). As seen in Table III, network 5×5, which can support up to 12 hops, has over 7.42×10^5 (> 0.74 million) possible path-pairs. This observation implies that when the proposed NoCiPUF framework is scaled up, the maximum possible hops increases and when supported with a longer HF, it leads to an even larger CRP-space. Modern-day NoCs are typically deployed in multi-node systems to seamlessly integrate hundreds of cores [29], [30]. Such systems require NoCs with mesh sizes 8×8 and above. Given the NoCiPUF's trend of exponential growth of the number of path-pairs (challenges) with increasing network size, one can easily conclude the possibility of a multi-million CRP space for NoC sizes 8×8 and above. Therefore, our intrinsic authentication enabled solution is perfectly suited for exa-scale MPSoC systems. Section IV shows various PUF-metric plots, which are classified according to the number of hops for a 4×4 NoC.

We have applied the challenges generated using algorithms 1 and 2 during simulation of the network with the proposed router architecture to obtain the results reported in next section.

C. Resilience Towards Modeling Attacks

Delay based extrinsic PUFs are inherently vulnerable to ML-based modeling attacks. For conventional Arbiter PUFs,

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:49:13 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Algorithm 2 Non Overlapping Path Pairs

Input: N (Mesh network size)
Output: Non overlapping path pairs sorted according to hops for each triplet ($S1$, $S2$, D)

```
path_dict = [0,1,2, ....., N-1];
for each combination of S1, S2, D do
  if S1 ≠ S2 ≠ D then
    P1_dict = obtain_paths(S1, D)
    P2_dict = obtain_paths(S2, D)
    path_pairs_dict = {}
    for i in P1_dict do
      for j in P2_dict do
        if i.hops == j.hops & ( ∀ nodes in
        i.visited_nodes != ∀ nodes in
        j.visited_nodes) then
          hops = i.hops
          PP = (P1, P2)
          path_pairs_dict.append(hops, PP)
        end if
      end for
    end for
  end if
end for
return path_pairs_dict
```

TABLE III
HOPS AND PATH PAIRS FOR NOCS OF DIFFERENT SIZES

Hops (H) \ Mesh Size	Number of path pairs (PP) 2x2	3x3	4x4	5x5
1	4	22	52	94
2	0	52	216	486
3	0	56	624	2076
4	0	20	1320	7260
5	0	0	2048	21368
6	0	0	1888	53804
7	0	0	616	111028
8	0	0	0	174380
9	0	0	0	191464
10	0	0	0	130596
11	0	0	0	45448
12	0	0	0	4324
TOTAL	**4**	**150**	**6764**	**742328**

such trained models are able to impersonate the trusted device by predicting the responses correctly with over 95% accuracy [16]. This observation prompts us to investigate the ML-resilience of NoCiPUF against the ML based modeling attacks and further suggest countermeasures to reduce the prediction accuracy to acceptable levels.

1) ML Attack Mechanism: For launching a ML based attack, the adversary first queries the trusted PUF device to collect a subset of CRPs as training data for models like LR, SVM, CMA-ES or DNN. To query the device means to throw a challenge and receive a response. Then the adversary trains a ML model based upon the received responses from the queries. Once the model is trained, the adversary uses this trained model to impersonate the trusted device (having PUF) by predicting new responses as and when queried by the parent server. Other attack strategies which involve tailoring ML-models to exploit the specific obfuscation mechanisms of

TABLE IV
ML-ACCURACY OF RESPONSE PREDICTION

ML Model	Without ML-Res.	After ML-Resilience
LR	78.82%	48.21%
SVM	82.99%	48.19%
CMA-ES	79.51%	56.62%
DNN	86.14%	54.81%

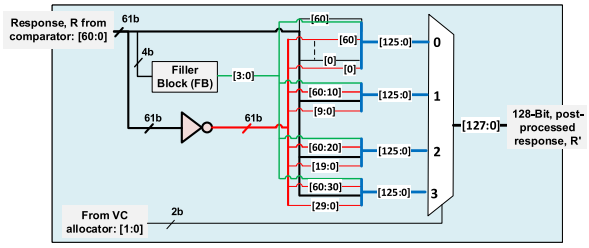


Fig. 9. Post-processing unit (PPU) inside router to counter ML attacks.

PUFs also exist [31]. Unlike the extrinsic PUFs using MUX chains or ring-oscillators, we re-use existing NoC structure to implement the PUF operation. The existing switches (routers) and interlinks are used to harness the unique responses. The already complex nature of NoC, combined with multitude of possibilities of path-pairs (hence, challenge-space) spread across the network makes it an extremely difficult task for an attacker to model the equivalent PUF hardware at their end. This provides a primitive layer of protection against the modeling attacks, by the virtue of the way NoCiPUF defines the PUF operation and provides a huge and scalable CRP space. This results in a nearly $8\% - 18\%$ decrease in the default prediction accuracy of NoCiPUF CRPs when no ML attack protection strategy is used, as shown in first column of Table IV

2) ML-Resilient Technique: To address the ML-based modeling attacks, existing works introduce more non-linearity in PUF stages using entropy sources based on advanced-encryption-standard (AES) [32] or XOR network [33], or complex of pseudo-LFSR networks [34]. Such schemes can only be applied for standalone extrinsic PUFs and incur very high area and power overheads. Also, due to their highly complex nature, these schemes are difficult to integrate with existing MPSoCs.

On the other hand, some schemes offer runtime-reconfigurability of CRPs or post-processing of challenges in PUFs to counter ML-modeling attacks [35], [36]. These schemes effectively conceal randomized post-processing steps from the attacker, and cause the adversarial model train on incorrect challenge-response mapping. Taking motivation from these reconfigurable approaches, for NoCiPUF, we deploy a post-processing hardware unit (thereafter called PPU) inside the router to modify the generated responses, as shown in Fig. 9. We assume that this post-processing hardware and its functionality are not known to the attacker. The PPU receives two sets of inputs, one is the default 61-bit response, R which is the output of comparator as explained in Section III-A and the second is the 2-bit signal received from the VC allocator of the router as shown in top and bottom left of Fig. 9, respectively. The PPU extends the 61-bit response (R) to a 128-bit post-processed response (R') using any one out of four predefined post processing rules. These rules can be defined

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:49:13 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

as: rule-1): add filler bits between every two bits; rule-2) to shift the response bits by x_1-bits, rule-3) by x_2-bits and rule-4) by x_3-bits. Rest of the bit-positions are occupied by filler bits.

In our implementation, we selected $x_1 = 10$, $x_2 = 20$, and $x_3 = 30$ and utilized the complement of response bits as filler bits. As shown in the Fig. 9, the incoming 61-bit response is inverted to generate 61 filler bits, making a total of 122-bits. The rest of the 4 filler bits are also obtained from the original response as indicated in the filler-block (FB) in the figure and the final 2 MSBs are the bits received from the VC allocator, making the rest of the 6 bits. The four rules are implemented by rearranging the response and filler bits as shown in the right side of the figure. The 4×1 MUX receives the re-arranged inputs for all four cases, and the 2 bits from the VC allocator to decide which input to select as the final, post-processed response. Therefore, the bit-extension rules implemented using the rewiring the real response and filler bits gets easily integrated with existing hardware. For example, for rule 2), where shifting by 10-bits is required, the LSB bits [9:0] are followed by next real response [60:0] bits, followed by inverted filler bits [60:0] and finally the last MSB positions are occupied by 4-bits from FB unit. The 2-MSBs from VC allocator are further appended by the 4×1 MUX in all the four cases. These 2-MSBs represent the ID of recently used VC just before CSA run, which results in completely unpredictable selection among the four rules. So, effectively, for the same challenge, the PPU can result in 4-possible responses, based on the randomly selected rule. The 2-MSBs of 128-bit processed response indicate the applied rule number and are used by the parent server to select real 61 response bits. The information about all possible rules is supplied beforehand to the server by the design-house. Now, for an adversary querying the NoCiPUF, its ML model will be trained to map every 64-bit challenge to a single 128-bit response. This response can be mapped to any of the four possible cases, unknown to the attacker and hence the trained model. So whenever the server receives a response generated by the trained adversarial model, it will use the 2-MSBs to choose the decoding scheme. Now, it is highly unlikely that the decoded-response bits picked by server match with stored, golden response. This is because the predicted response was based on training with the post-processed responses received earlier, which are subject to change with time even if the challenge remains the same. This results in dramatic reduction of model's prediction accuracy in correctly modeling the underlying NoCiPUF.

We tested this scheme and trained the model with 80% of CRPs $(5,411)$ and measured accuracy on remaining 20% CRPs $(1,352)$. We used default parameters for LR and SVM, where $C = 1.0$ with $lbfgs\ solver$ and $kernel = rbf$. For CMA-ES, all initial parameters were set to 0 with an initial sigma of 0.5. We ran 250 iterations to find the best model and have reported the results for the same. For the DNN model, we constructed a 3-layer neural network with real and sigmoid activation for the first two and the last layer, respectively. We used $Adam$ optimizer and ran 50 epochs with a batch size of 32. As shown in Table IV, suggested post-processing scheme brings down prediction accuracy of various ML models from more than $\sim 81\%$ to around 50% (average = 51.96%), which is close to ideally expected value for ML-resilience in PUFs because this means probability of guessing the correct response by trained model is almost that of coin-flip, rendering the trained model unreliable. We make sure that the PUF's randomness and the metrics of uniformity, uniqueness and reliability remain unaffected by the ML resilient post-processing operation as the originally generated response bits (61-bits) are still preserved. The associated hardware overheads are provided in the next section.

IV. Experimental Setup & Results

We have considered a 4×4 NoC with mesh topology for experiments in this paper. The proposed NoCiPUF router is shown in Fig. 5. As explained earlier, the proposed router architecture supports hybrid switching (PS and CS). The baseline router considered for comparison is shown in Fig. 3. This baseline router is assumed to have PS and XY routing implemented. The extracted intrinsic PUF's validity and correctness has been quantified in terms of uniformity, uniqueness and reliability (explained in Section II-A). For the experiments, we have assumed 64-bit wide flits supported by 64-bit wide inter-node links. The 3 MSBs of the 64-bit flit indicate the SI and FI. This leaves the remaining 61-bits which are used as challenge in head-flit (HF), and as the required trigger in body-flit (BF) to generate the 61-bit response at a given destination router. We use a 61-bit comparator at every router to obtain the 61-bit response concurrently. This approach makes sure that every challenge fetches the complete multi-bit response in one go, thus avoiding the requirement of applying multiple challenges for longer response. As can be observed from Table III, hops (H) up to 7 are supported by 4×4 mesh which occupy 14 bits of challenge HFs. The BF is configured to have all 1s so that the transition of 0 to 1 (in both the chosen paths P1 and P2) can be observed at the comparator's inputs. The responses are obtained using Monte-Carlo SPICE simulations that have been carried out on Eldo using the ST-Microelectronics SVT 65nm Technology library. Each Monte-Carlo simulation represents a separate instance (fabricated chip). We consider all possible CRPs for 4×4 NoC mesh as shown in Table III for validating the randomness of NoCiPUF's responses, using NIST suite [37]. We have synthesized the NoC using Cadence Genus to report area and power overheads for comparative analysis.

In the following sub-sections, we report the measured PUF metrics, NIST randomness results, and area and power comparisons with baseline NoC and state-of-the-art router based intrinsic PUF.

A. Uniformity, Uniqueness and Reliability

The data used to calculate the three standard PUF metrics corresponds to original responses (61-bit) obtained just after the comparator stage, independent of their post-processing for ML-resilience. This is because only these 61 bits will be selected automatically by the parent server from among the 128-bit extended response and then compared with golden LUT entry.

Uniformity is an indicator of the average bias of the resultant PUF's responses towards bit 0 or bit 1. Ideal value is 50%. Using equation (1), and taking values of $n = 61$ (response bits corresponding to comparator output) and $m = 200$ (number of responses obtained), our Monte-Carlo SPICE simulation results yield uniformity of 53.3% across all the responses

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:49:13 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

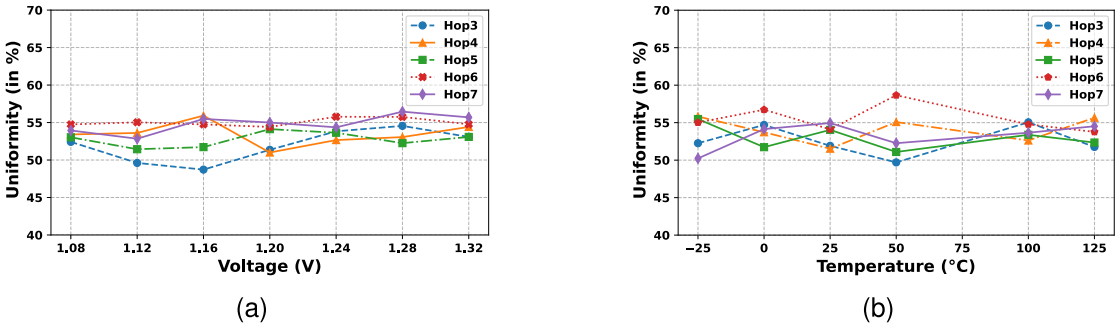


Fig. 10. Uniformity trends at different conditions: (a) as voltage is varied with steps of 40mV, (b) as temperatures are varied from −25° C to 125° C.

which is very close to the ideally expected value. This value is collected at temperature of 25° C and voltage of 1.2V. Plots in Figs. 10a, 10b show the trend of uniformity across various voltages and temperatures respectively. Uniformity varies from 49.23% (30.03-bits, HOP=3, 1.16V) to 56.27% (34.32-bits, HOP=4, 1.16V) across voltages and from 49.7% (30.32-bits, HOP=3, 50° C) to 58.66% (35.78-bits, HOP=6, 50° C), across temperatures for all the hops.

As shown in Fig. 2, fabricated chips should have different response (RN_M) to same challenge (CN_M) thrown to them during the PUF operation. Inter-chip variation, also called uniqueness, indicates how varied are the responses among the given set of chips for same challenge thrown to them. It is measured using equation (2), taking $k = 200$. In Fig. 11a, we can observe that for 200 Monte-Carlo simulations (i.e., analogically 200 chips) at hops ranging from H=3 to H=7 the average HD converges to ∼50% for all hops. Exact values vary from 50.94% (31.07-bits, HOP=7) to 51.58% (31.46-bits, HOP=3) per 61-bit response at nominal conditions of 1.2V, 25° C. The overall average uniqueness is 51.24% (31.26-bits) which is almost the ideal inter-chip variation expected. Plots in Figs. 11b and 11c show trends for uniqueness as temperatures and voltages are varied respectively. Minimum and maximum uniqueness values across all temperatures are 49.66% (30.29-bits, HOP=6, 50° C) and 51.71% (31.54-bits, HOP=3, 50° C) respectively. Minimum and maximum uniqueness values across all voltages are 50.53% (30.82-bits, HOP=4, 1.16V) and 51.70% (31.53-bits, HOP=3, 1.12V) respectively. This trend shows that the NoCiPUF-CRPs are resilient against varying temperatures and voltages in terms of inter-chip variations as the corresponding uniqueness values remain nearly ideal.

The intra-chip variation indicates the reliability of PUF responses against varying environmental conditions. Same challenge is thrown to the same chip (here, same Monte-Carlo instance) under different environmental conditions in terms of voltage and temperature variations. As per the equation (3), the average of intra-chip HDs (HD_{INTRA}) represent BER of the PUF. The plots in Figs. 12a, 12b show the distribution of intra-chip HDs with varying temperatures and voltages respectively. We vary temperatures ranging from −25° C to 125° C and record the responses, to obtain plots shown in Fig. 12a. It can be observed from the plot that, for different number of hops, the frequencies of large HD values converge to zero, indicating a very low BER. The average BER against temperature changes is found to be 1.9-bits per 61-bit response, meaning 96.95% reliability. The Fig. 12b shows the distribution of intra-chip HDs among responses for various hops as the voltage supply is varied from −10% to +10% around 1.2V. The reliability against voltage variations is found to be 97.34%. This implies an average BER of 1.7-bits. The plots in Figs. 12c, 12d show the percentage of reliably generated bits (reliability values) at different temperatures (with respect to 25 ° C) and at different voltages (with respect to 1.2V) respectively. These reliably generated bits vary from 95.23% (58.1-bits, HOP=7, −25° C) to 98% (59.8-bits, HOP=7, 0° C), per 61-bit response for different temperatures. Voltage changes result in reliably generated bits varying from 95.26% (58.11-bits, HOP=4, 1.16V) to 98.75% (60.23-bits, HOP=7, 1.24V). We have further calculated the BER for extreme conditions of high temperature (125° C) and low voltage (1.08 V) to be 5.9 bits per 61-bit response. It can be clearly observed from the plots in the Fig. 12 that the suggested intrinsic PUF framework shows good (close to ideal) resilience against environmental variations of temperature and voltage. The same observation can be extended to the two metrics of uniformity and uniqueness as NoCiPUF shows nearly ideal values to these metrics at different voltages and temperatures also. The state-of-the-art works [12] have reported the BER of as low as zero ($\sim 3.76e-08$), however our work additionally offers resilience to a variety of ML models and its NoC centric intrinsic nature makes it scalable for large MPSoCs. Furthermore, the corresponding values for uniqueness, uniformity and, reliability against voltage and temperature values for the work [14], which proposed implementation of Arbiter PUF in NoC router, are: 49.3%, 56.46% and 94.71% and 94.82% respectively.

B. NIST Randomness Analysis

The NIST [37] provides a test-suite to quantify randomness of random and pesudo-random number generators. We use this suite to evaluate the quality of randomness of responses produced from the proposed authentication mechanism implemented as intrinsic PUFs in the mesh NoC.

NIST randomness suite requires multiple (thousands or more) bit-sequences pertaining to a single TRNG instance, NoCiPUF in our case. We chose one random Monte-Carlo instance and apply challenges to generate the corresponding responses. We obtained 6, 764 61-bit responses amounting to 412, 604 bits using this method. To run the tests, we split these into 64-bit, 6, 446 bit-streams (rounded off). It can be observed from the Table V that almost all the tests fetch near ideal results (above 90%). The Table V also shows the corresponding P-values for every test. P-value greater

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:49:13 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

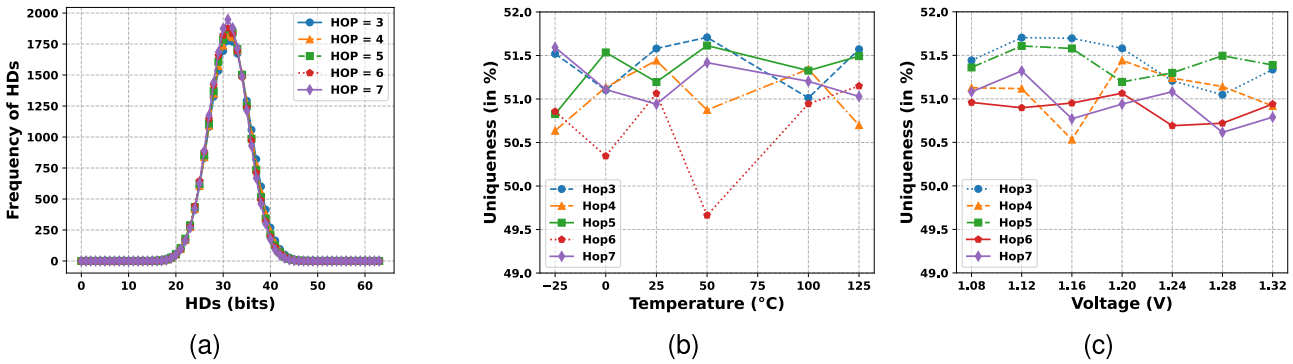


Fig. 11. (a) Hop-wise distribution of inter-chip HDs. Uniqueness values for, (b) temperatures ranging from -25° C to 125° C and, (c) values for voltages from 1.08V to 1.32V with 40mV steps.

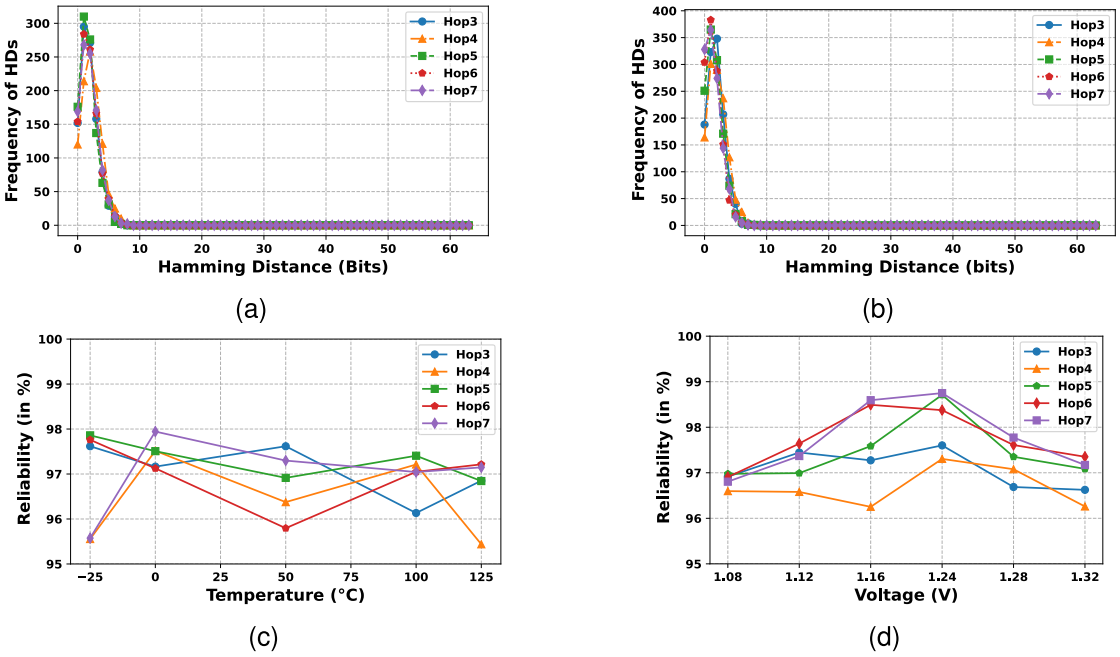


Fig. 12. Reliability (intra-chip variations) trends for hops=3 to 7. Frequencies of intra-chip HDs (BERs) vs their values for: (a) temperatures varying from -25° C to 125° C and, (b) voltages varying from -10% (1.08V) to $+10\%$ (1.32V) of nominal 1.2V with steps of 40mV. Percentage of reliably generated bits for, (c) variations with temperature (with respect to 25° C), (d) variations with voltages (with respect to 1.2V).

TABLE V
NIST TESTS RESULTS

Test	Pass Percentage	Avg. P-value
ApproximateEntropy	100%	1
Serial	94%	0.4835789725
BlockFrequency	100%	1
CumulativeSums	93%	0.4085557753
FFT	97%	0.4398678753
Frequency	92%	0.389980733
LongestRun	99%	0.4580607531
NonOverlappingTemplate	100%	0.6661900229
Rank	100%	0.1037209924
Runs	99%	0.499993895

TABLE VI
OVERHEADS COMPARISON SUMMARY OF NoCiPUF

Additional H/W / Metric	Baseline + NoCiPUF	Baseline + NoCiPUF + PPU H/W	Ref. [14]
Area overhead (mm^2)	0.0037 (0.52%)	0.0055 (0.76%)	0.009 (1.26%)
Power overhead (mW)	3.87 (0.77%)	5.73 (1.14%)	1.13 (0.22%)
Delay	1.3ns	1.4ns	5.2ns

than 0.01 is the criteria used by the NIST suite to imply if that the given bit-sequence passed the test.

C. Overheads Analysis

The differences between the baseline and proposed NoC router can be observed by comparing Figs. 3 and 5, respectively. To support hybrid switching, additional RC and SA units have been augmented in existing routers. PUF control unit has been added (Fig. 5) for CSA mode to decide how the HF will be decoded in CS mode and where should the upcoming flits travel in case of destination router. Further, bypass paths are enabled in IC. Extra multiplexers are required in IC (to bypass) after RC-SA stage (to select between CS or PS), and a demultiplexer is needed to decide where the flit should go at the destination node (to the comparator for CSA and local output for CSN/PSN). We target a performance of 1 GHz to calculate all the overheads resulting due to hardware changes.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:49:13 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE VII
QUALITATIVE COMPARISONS

	Ref. [14]	Ref. [15]	NoCiPUF
Intrinsic PUF implementation style	Comparison of delays associated with cross-bar in a single router	Comparison of delays associated with gates in a synthesized netlist	Comparison of delays associated with equal hop path-pairs in NoC
Level of abstraction (Ease of implementation)	Circuit level (moderate)	Gate level (difficult)	Behavioral level (easy)
ML Resilience in Intrinsic PUF	Not provided	Not provided	Provided, up-to ∼52% unpredictability of responses
Inherent bias in nominal delays	No	Yes	No
Detailed visibility of authentication and normal operation modes	No	No	Yes
PUF specific special layout requirement	Yes	No	No
CRP space Computational Complexity	Low	High	Moderate

These hardware changes for NoCiPUF are reflected as 0.52% area overhead and 0.77% power overhead as compared to baseline router. We have reported the delays to generate the responses once the challenge is applied. For the proposed NoCiPUF, longest possible 7-hop path for the 4×4 NoC incurs a delay of 1.3ns. Second column in Table VI summarizes these area and power overheads as compared to the baseline NoC.

Further, we also added hardware for on-chip post-processing of responses for ML-resilience (PPU) at every router. This incurs some additional power, area and delay overheads, apart from overheads from just PUF-specific changes. Cumulative overhead of NoCiPUF after addition of ML-resilient hardware becomes 0.76% in terms of area, and 1.14% in terms of power. The maximum delay to generate a single 128-bit post-processed response is measured to be 1.4ns. This delay has increased only slightly compared to previous observation. These results are summarized in the third column of Table VI. We can observe that even with additional hardware required to counter ML-attacks by adversary, the overheads are still under manageable limits, making the suggested framework seamlessly scalable. Additionally, introducing more number of post-processing rules will not only provide protection against more advanced ML models but based on the current number, will incur overheads well under practical limits.

We further compare the proposed NoCiPUF framework with an existing work [14] which used individual routers to implement Arbiter PUFs and summarize the overhead comparison in the fourth column of Table VI. It should be noted that this work does not provide any guarantee against ML resilience and hence there is no equivalent hardware used in their implementation. We can observe that as compared to baseline, their area overhead is 1.26% which is more as compared to our work, even after including ML resilience in NoCiPUF. The power overhead is 0.22% which is lower than our work, but is less significant, given the fact that our framework offers ML resilience also. However, the delay to generate response from the single router after the challenge is applied is measured to be 5.2ns. This delay is ∼4x that of the NoCiPUF response generation delay. There are several benefits of NoCiPUF over the work [14] in terms of completeness of operation, functionality and implementation worthiness, which are elaborated in the next Section V.

The following section highlights the qualitative comparisons with works in the existing literature relevant to the proposed authentication framework.

V. DISCUSSION AND QUALITATIVE COMPARISONS

As mentioned in Section II-C, several works target SoC authentication in various aspects. In this section, we provide qualitative comparisons with some works relevant to the context of this paper. These comparisons have been summarized in Table VII. The work [14] proposes to realize intrinsic PUFs using the multiplexers in the crossbar (XB) of a NoC router, limiting their scope to individual routers. This makes router-to-router communication and interaction among intra-router components ambiguous during PUF operation. In contrast, we have defined separate modes to integrate the intrinsic PUF operation in a given NoC. These modes are supported by corresponding HFs, to ensure proper communication among routers and provide intra-router strategy during authentication operation in NoCiPUF. Their scheme assumes a specific implementation of XB and modifies it along with the use of pass-transistors (PTL) between multiplexer stages. However, as we have suggested all changes at the behavioral level, they require no specific implementation of XB or of any other router components, and hence do not need any special consideration regarding the PTL or standard cell technology library. This makes NoCiPUF directly adaptable to existing standard cell flow for NoCs. Moreover, we highlight the creative use of head-flit in vector-routing-based CS in conjunction with RC, SA and XB to generate select lines for path reservation. The work [14] does not, however, address how a router should be provided with challenge information to produce select lines for their XB.

In the work [15], the authors suggest a method to obtain intrinsic PUFs in combinational netlists using path-pairs. The scheme is applicable only to gate-level netlist and requires multiple computations to select paths. Their scheme requires SAT-based PHAETON engine to find path-pairs and separate algorithms for tap-point-optimization, and these computations depend on the number and the type of gates present in synthesized netlists. However, they have tested their scheme for benchmarks circuits having few tens of thousand gates only, while modern MPSoCs can have up to millions of gates. Therefore, such multi-layered computations are bound to be heavy when applied to larger and more complex MPSoCs, making their scalability questionable. However, we have developed much simpler algorithms to find NoC path-pairs corresponding to CRP space. These can be easily reproduced and adopted for any given NoC size and topology, strengthening the generosity of our framework for NoC-based MPSoCs. Moreover, the work can only select approximately equal paths, which are subject to unequal nominal delays, thereby causing an inherent bias in delays. This biasness causes metrics such as uniformity to degrade. This phenomena necessitates an additional layer to further filter out the paths with near-ideal metrics, further reducing the set of usable CRPs. In contrast, we use all the available non-overlapping

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:49:13 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

equal-hop path-pairs amounting to equal nominal delays, ensuring maximum unpredictability due to no approximation, when flits arrive at destination node. Additionally, in our framework there is no need to take individual gate delays into account while computing the CRP space. Remarkably, none of the aforementioned intrinsic PUF schemes have talked about resilience to ML based modeling attacks, whereas, we address the ML resilience using a lightweight technique. The other works which offer ML resilience are extrinsic in nature [38]. The ML resilience offered by NoCiPUF (∼52%) is extremely close to the ideally expected value of 50%.

Other works related to authentication, such as digital IP watermarking in NoC, provide only a singular watermark per IP to verify if it is sold by a trusted seller [27]. However, using NoCiPUF facilitates multiple CRPs for the same MPSoC. So, in the event that a secret key is eavesdropped by an adversary, the next authentication phase will have a choice from the huge CRP space provided by NoCiPUF. Some schemes target chip authentication at foundry and/or testing facility [23], [24], which is different from the focus of our work.

VI. Conclusion

In this work, we propose NoCiPUF, a robust, feasible and effective authentication methodology to reuse the interconnect fabric of NoC-based MPSoCs. Our scheme is based on intrinsic PUF derived from equal-hop path-pairs in NoC. We also provide a comprehensive technique to counter most common ML-based modeling attacks on strong PUFs and successfully reduce the prediction accuracy from over 86−% to around 52%. The results promise near-ideal intrinsic-strong-PUF behavior with a uniqueness of 51.24% (31.26-bits), reliability against temperature and voltage to be 96.95% and 97.34% respectively, and uniformity of 53.3%. We also showcase robustness of PUF responses against NIST randomness test suite where tests yield more than 90% pass rate among the input response bit-streams. We make efficient use of hybrid switching, which not only re-uses circuit-switched paths for authentication but enables choice between packet and circuit switched paths for regular data flow. The resultant area and power overheads are 0.76% and 1.14% respectively with respect to equivalent baseline NoC. The algorithmically calculated usable paths-pairs clearly indicate that the CRP-space rises exponentially, up to > 0.74 million with network size of just 5×5 NoC, which is remarkably suitable for securing advanced and complex MPSoCs which require larger NoCs.

References

[1] S. Pasricha and N. Dutt, "Networks-on-chip," in *On-Chip Communication Architectures*, S. Pasricha and N. Dutt, Eds., San Mateo, CA, USA: Morgan Kaufmann, 2008, ch. 1, pp. 439–471. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780123738929000128

[2] P. D. K. A. Kumar and V. K. Reddy, "Design of NoC-based high-speed processor to suit IoT for a hyper-connected smart world," in *Harnessing the Internet of Things (IoT) for a Hyper-Connected Smart World*, 1st ed., New York, NY, USA: Apple Academic Press, 2022, ch. 1.

[3] G. S. Pandi, S. Shah, and K. H. Wandra, "Exploration of vulnerabilities, threats and forensic issues and its impact on the distributed environment of cloud and its mitigation," *Proc. Comput. Sci.*, vol. 167, pp. 163–173, Aug. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050920306591

[4] Z. Whittaker. (2024). *The Biggest Data Breaches in 2024: 1 Billion Stolen Records and Rising*. [Online]. Available: https://techcrunch.com/2024/10/14/2024-in-data-breaches-1-billion-stolen-records-and-rising/

[5] *Chipmaker NXP Confirms Data Breach Involving Customers' Information*. Accessed: Feb. 19, 2025. [Online]. Available: https://techcrunch.com/2023/09/05/chipmaker-nxp-confirms-data-breach-involving-customers-information/

[6] C. Herder, M.-D. Yu, F. Koushanfar, and S. Devadas, "Physical unclonable functions and applications: A tutorial," *Proc. IEEE*, vol. 102, no. 8, pp. 1126–1141, Aug. 2014.

[7] J. Zhang, L. Ding, Z. Chen, W. Li, and G. Qu, "DA PUF: Dual-state analog PUF," in *Proc. 59th ACM/IEEE Design Autom. Conf.* New York, NY, USA: ACM, Jul. 2022, pp. 73–78, doi: 10.1145/3489517.3530412.

[8] I. Baturone, R. Román, and Á. Corbacho, "A unified multibit PUF and TRNG based on ring oscillators for secure IoT devices," *IEEE Internet Things J.*, vol. 10, no. 7, pp. 6182–6192, Apr. 2023.

[9] G. Edward Suh and S. Devadas, "Physical unclonable functions for device authentication and secret key generation," in *Proc. 44th ACM/IEEE Design Autom. Conf.*, Jun. 2007, pp. 9–14. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1278480.1278484

[10] S. Tajik et al., "Physical characterization of arbiter pufs," in *Proceedings of the Hardware and Embedded Systems—CHES 2014*, L. Batina and M. Robshaw, Eds., Berlin, Heidelberg: Springer, 2014, pp. 493–509.

[11] A. Mahmoud, U. Rührmair, M. Majzoobi, and F. Koushanfar, "Combined modeling and side channel attacks on strong pufs," *Cryptology ePrint Archive*, Oct. 2013, Art. no. 211. [Online]. Available: https://eprint.iacr.org/2013/632

[12] L. Ni and J. Zhang, "S2RAM PUF: An ultra-low power subthreshold SRAM PUF with zero bit error rate," in *Proc. 61st ACM/IEEE Design Autom. Conf.* New York, NY, USA: ACM, Jun. 2024, pp. 1–6, doi: 10.1145/3649329.3658246.

[13] L. Ni, P. Wang, Y. Zhang, G. Li, L. Ding, and J. Zhang, "PI PUF: A processor-intrinsic PUF for IoT," *Comput. Electr. Eng.*, vol. 105, Jan. 2023, Art. no. 108540, doi: 10.1016/J.COMPELECENG.2022.108540.

[14] P. Nagabhushanamgari, V. Sehwag, I. Chakrabarti, and S. Chattopadhyay, "Embedding delay-based physical unclonable functions in networks-on-chip," *IET Circuits, Devices Syst.*, vol. 15, no. 1, pp. 27–41, Jan. 2021.

[15] M. Sauer, P. Raiola, L. Feiten, B. Becker, U. Rührmair, and I. Polian, "Sensitized path PUF: A lightweight embedded physical unclonable function," in *Proc. Design, Autom. Test Eur. Conf. Exhib. (DATE)*, Mar. 2017, pp. 680–685.

[16] U. Rührmair et al., "PUF modeling attacks on simulated and silicon data," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 11, pp. 1876–1891, Nov. 2013.

[17] C. Nicopoulos et al., "On the effects of process variation in network-on-chip architectures," *IEEE Trans. Dependable Secure Comput.*, vol. 7, no. 3, pp. 240–254, Sep. 2010.

[18] S. H. Gade, H. K. Mondal, and S. Deb, "High bandwidth off-chip memory access through hybrid switching and inter-chip wireless links," in *Proc. 2018 IEEE Comput. Soc. Annu. Symp. VLSI (ISVLSI)*. IEEE, 2018, pp. 100–105.

[19] A. Mazloumi and M. Modarressi, "A hybrid packet/circuit-switched router to accelerate memory access in NoC-based chip multiprocessors," in *Proc. Design, Autom. Test Eur. Conf. Exhib. (DATE)*, Mar. 2015, pp. 908–911. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7092516

[20] W. Xu, L. Feng, J. J. Rajendran, and J. Hu, "Layout recognition attacks on split manufacturing," in *Proc. 24th Asia South Pacific Design Autom. Conf.* New York, NY, USA: ACM, Jan. 2019, pp. 45–50, doi: 10.1145/3287624.3287698.

[21] S. Charles, V. Bindschaedler, and P. Mishra, "Digital watermarking for detecting malicious intellectual property cores in NoC architectures," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 30, no. 7, pp. 952–965, Jul. 2022.

[22] A. Shalaby, Y. Tavva, T. E. Carlson, and L.-S. Peh, "Sentry-NoC: A statically-scheduled NoC for secure SoCs," in *Proc. 15th IEEE/ACM Int. Symp. Netw. Chip (NOCS)*, Oct. 2021, pp. 67–74, doi: 10.1145/3479876.3481595.

[23] M. S. U. I. Sami et al., "POCA: First power-on chip authentication in untrusted foundry and assembly," in *Proc. IEEE Int. Symp. Hardw. Oriented Secur. Trust (HOST)*, Dec. 2021, pp. 124–135.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:49:13 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

[24] P. Slpsk, J. Cruz, S. Ray, and S. Bhunia, "PROTECTS: Secure provisioning of system-on-chip assets in untrusted testing facility," in *Proc. IEEE Int. Test Conf. India (ITC India)*, Jul. 2023, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/10235499/

[25] K.-C. Chang, J.-S. Shen, and T.-F. Chen, "Evaluation and design trade-offs between circuit-switched and packet-switched NOCs for application-specific SOCs," in *Proc. 43rd ACM/IEEE Design Autom. Conf.*, Jul. 2006, pp. 143–148.

[26] M. Modarressi, H. Sarbazi-Azad, and M. Arjomand, "A hybrid packet-circuit switched on-chip network based on SDM," in *Proc. Design, Autom. Test Eur. Conf. Exhib.*, Apr. 2009, pp. 566–569. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/5090728

[27] Q. Liu, W. Ji, Q. Chen, and T. Mak, "IP protection of mesh NoCs using square spiral routing," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 24, no. 4, pp. 1560–1573, Apr. 2016, doi: 10.1109/TVLSI.2015.2462842.

[28] Y. Yang, A. Funahashi, A. Jouraku, H. Nishi, H. Amano, and T. Sueyoshi, "Recursive diagonal torus: An interconnection network for massively parallel computers," *IEEE Trans. Parallel Distrib. Syst.*, vol. 12, no. 7, pp. 701–715, Jul. 2001.

[29] A. Olofsson, "Epiphany-V: A 1024 processor 64-bit RISC system-on-chip," 2016, *arXiv:1610.01832*.

[30] AMD. (2023). *Amd EpycT 9754s*. Accessed: Jul. 12, 2024. [Online]. Available: https://amperecomputing.com/briefs/ampereone-family-product-brief

[31] N. Hassan and U. Chatterjee, "Machine learning attacks on challenge-response obfuscations in strong PUFs," in *Proc. IEEE Int. Symp. Hardw. Oriented Secur. Trust (HOST)*, May 2024, pp. 361–372.

[32] V. Suresh, R. Kumar, M. Anders, H. Kaul, V. De, and S. Mathew, "A 0.26% BER, 1028 challenge-response machine-learning resistant strong-PUF in 14 nm CMOS featuring stability-aware adversarial challenge selection," in *Proc. IEEE Symp. VLSI Circuits*, Jun. 2020, pp. 1–2.

[33] Y. He, Q. Yu, and K. Yang, "A lossless and modeling attack-resistant strong PUF with <4E-8 bit error rate," in *Proc. IEEE Custom Integr. Circuits Conf. (CICC)*, Apr. 2022, pp. 1–2.

[34] H. Nassar, L. Bauer, and J. Henkel, "CaPUF: Cascaded PUF structure for machine learning resiliency," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 41, no. 11, pp. 4349–4360, Nov. 2022.

[35] Y. Cao, X. Zhao, W. Ye, Q. Han, and X. Pan, "A compact and low power RO PUF with high resilience to the EM side-channel attack and the SVM modelling attack of wireless sensor networks," *Sensors*, vol. 18, no. 2, p. 322, Jan. 2018. [Online]. Available: https://www.mdpi.com/1424-8220/18/2/322

[36] N. Wisiol, "Towards attack resilient delay-based strong PUFs," in *Proc. IEEE Int. Symp. Hardw. Oriented Secur. Trust (HOST)*, McLean, VA, USA, Jun. 2022, pp. 5–8.

[37] L. E. Bassham III et al., *A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications*, document SP 800-22, NIST, 2010.

[38] Z. Chen, T. Sato, and H. Shinohara, "SpongePUF: A modeling attack resilient strong PUF with scalable challenge response pair," in *Proc. IEEE Int. Symp. Hardw. Oriented Secur. Trust (HOST)*, May 2024, pp. 244–253.



Deepank Grover (Graduate Student Member, IEEE) received the B.Tech. degree in ECE from Delhi Technological University, Delhi, India, in 2019. He is currently pursuing the Ph.D. degree in ECE with Indraprastha Institute of Information Technology Delhi, Delhi, India. He worked with Synopsys, India, from 2019 to 2021. His research interests include hardware authentication and network-on-chips.



Tarun Sharma (Graduate Student Member, IEEE) received the M.E. degree from the Birla Institute of Technology and Science, Hyderabad Campus, India, in 2021. He is currently pursuing the Ph.D. degree in ECE with Indraprastha Institute of Information Technology Delhi, New Delhi, India. His research interests include approximate computing, network-on-chip, and hardware security.



Sneha Agarwal (Graduate Student Member, IEEE) received the B.Tech. degree in ECE from the Jaypee Institute of Information Technology, Noida, India, in 2020. She is currently pursuing the Ph.D. degree in ECE with Indraprastha Institute of Information Technology Delhi, India. Her research interests include NoC traffic characterization, run-time control, and hardware security.



Sidhartha Sankar Rout received the M.Tech. and Ph.D. degrees in ECE from Indraprastha Institute of Information Technology Delhi, New Delhi, India, in 2014 and 2023, respectively. His research interests include design-for-debug, system validation, and hardware security.



Anushka received the B.Tech. degree in ECE from the SRM Institute of Science and Technology, Tamil Nadu, India, in 2019, and the M.Tech. degree in ECE with specialization in VLSI and embedded systems from Indraprastha Institute of Information Technology Delhi, Delhi, India, in 2023.



Madhur Kumar received the B.Tech. degree from Indraprastha Institute of Information Technology Delhi, Delhi, India, in 2023. His research interests include CPU microarchitecture and memory systems.



Sujay Deb (Senior Member, IEEE) received the M.S. degree from Indian Institute of Technology, Kharagpur, India, in 2007, and the Ph.D. degree from Washington State University, USA, in 2012. He is currently a Professor with the Department of Electronics and Communication Engineering, Indraprastha Institute of Information Technology Delhi, India. His research interests include the design of novel interconnect architectures for multi-core chips, heterogeneous system architectures.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 06:49:13 UTC from IEEE Xplore. Restrictions apply.

Indian Journal of Pure & Applied Physics
Vol. 63, April 2025, pp. 281-300
DOI:10.56042/ijpap.v63i4.14388

Numerical Simulation on Sensitivity Modulation in Nanowire MOSFEB Detector for Biomedical Applications

Amit Das[a], Anjana Bhardwaj[a,*], Kaushik Das[b], Shivani Yadav[c], Aapurva Kaul[d], Preeti Goyal[e], Swati Sharma[e], Sonam Rewari[d], Binod Kumar Kanaujia[f] & R S Gupta[e]

[a]Department of Electronics and Communication Engineering, GLA University, Mathura, Uttar Pradesh 281 406, India

[b]Department of Electronics and Communication Engineering, National Institute of Technology Jamshedpur, Adityapur 831 014, India

[c]Department of Electronics and Communication Engineering, KIET Group of Institutions, Ghaziabad 201 206, India

[d]Department of Electronics and Communication Engineering, Delhi Technological University, New Delhi 110 042, India

[e]Department of Electronics and Communication Engineering, Maharaja Agrasen Institute of Technology, New Delhi 110 086, India

[f]Department of Electronics and Communication Engineering, Dr. B. R. Ambedkar National Institute of Technology Jalandhar 144 027, India

Received: 8 November 2024; accepted: 31 March 2025

The work aims to investigate and evaluate the biodetection capability of a nanowire MOSFEB detector for label-free biomedical applications. The study adopts a numerical computational-cum-simulation methodology to modulate, control and optimize sensitivity through dielectric modulation. Doping, a controllable factor prior to the fabrication process, plays a crucial role in enhancing the sensitivity of similar biosensors, thus making this work valuable. It is evident in the study that lower doping levels in the source and drain result in a greater percentage change in sensing metrics, whereas higher doping levels lead to a more significant relative change in sensing metrics. Additionally, the study comprehensively investigates the effect of drain voltage, material engineering, cavity dimensions/location, temperature, and device parameters on sensitivity. Sensitivity improvements of 18.573%, 13.979%, and 13.459% are achieved through gate, oxide, and channel engineering, respectively in the proposed biosensor. The study also addresses effect of trap charges on device reliability, alongside the calculation of the limit of detection. The results indicate that the silicon source exhibits better reliability and fewer short channel effects compared to other materials. Practical limiting factors have been taken into consideration in terms of non-unity fill-in factor and different filling profiles. The detector has also been compared with its junctionless variant, demonstrating a sensitivity enhancement of 95.716 mV for K_{bio}=5. The nanowire detector has been benchmarked against reported works in literature in terms of sensitivity.

Keywords: Biodetector, Cylindrical gate MOSFEB, Detector, FEB, MOSFEB, Nanosensor, Nanowire, Nanowire MOSFEB, NWM-BD, Subthreshold slope, SILVACO TCAD, Transistors

1 Introduction

Field effect transistors (FETs) have made substantial impact in the realm of electronics, offering a variety of specialized types that cater to a wide array of applications[1,2]. The versatility in FET is extensively utilized in the biomedical field, where it is used for different clinical tasks. One of the extensively researched applications is biosensing[3]. An exemplary biosensor should be highly sensitive to all the biomolecules, indicated by the difference in the value of different sensing metrics, demonstrating a high sensitiveness (response). The insertion of biomolecules into cavities primarily alters the profile of the channel's electric field and potential. Researchers have investigated different FET architectures for biomedical applications, such as extended-gate FETs, ion-sensitive FETs (ISFETs), and nanowire FETs, each offering specific benefits regarding sensitivity, selectivity, and fabrication simplicity[3-6]. The selection of semiconductor materials is also significant; silicon is widely used due to its simplified (common) processing methods, while new materials like graphene and molybdenum disulfide are promising for improved analog performance.

A biosensor is an analytical device used for the rapid detection of bio-chemical species (biomolecules). Biosensors are either label-based, using unique labels to identify targets, or label-free, suitable for untagged

*Corresponding author: (E-mail: anjana.0889@gmail.com)

analytes. The choice between these approaches is contingent upon the specific analytical requirements and the biochemical properties of the target analytes[7]. Label-free detectors offer particular utility in scenarios where molecular tagging is challenging or undesirable, and for the direct detection of unmodified target entities (biochemical species). Conversely, label-based detectors leverage the unique properties of molecular tags to achieve highly specific target identification (improving specificity). Label-less detection is preferred for its simplicity, affordability, and suitability for real-time analysis. Due to their exceptional sensitivity, MOSFET based label-free biosensors (MOSFEB)[4] have garnered significant popularity in the field of bioelectronics[8]. Harnessing the remarkable properties of semiconductors, MOSFEB stands at the forefront of biotechnology, along with its integration with nanotechnology innovation. By capitalizing on the subtle modulation of electrical properties upon biomolecular interactions, these biosensors enable label-free, rapid, and ultrasensitive detection, underscoring their role in revolutionizing fields such as personalized medicine and bioprocess monitoring[9]. These biosensors use the molecular properties (dielectric constant and charge density) of different biochemical species for label-free detection, which means that the device can directly detect the test biomolecule without any tagged entity[4]. The primary principle behind the detection ability of every MOSFEB is dielectric modulation of the sensing parameter. Figure 1 (a-b) demonstrate the fundamental device principle and basic detection principle in nanowire MOSFEB. These detectors leverage selective biorecognition elements to detect and quantify target analytes such as biomolecules, pathogens, or chemicals. The threshold voltage is the commonly used sensing parameter, but apart from this, various characteristics like subthreshold slope, off-current, I_{ON}/I_{OFF} ratio, and transconductance can be used to analyze the sensitivity. The detector derives its name from the fact that it uses a transistor whose charge carrier transportation mechanism is modulated by electric field (FET), and it incorporates dielectric engineering (MOS) for label-free detection of biochemical species. Hence, the name Nanowire MOSFEB Detector (Nanowire MOSFET + Biomolecules Detector: MOSFET for biomolecules detection), abbreviated as NWM-BD. The term "nanowire" in nanowire MOSFEB refers to its surrounding gate (SG) architecture, which enhances electrostatic control over the channel and improves charge carrier transport. The versatility of MOSFEB extends beyond their basic operational principle, encompassing a wide array of design variations and sensing mechanisms.

Several researchers have explored the concept of biomolecule sensing in transistors through gate oxide modulation. But not much-related work (doping dependent sensitivity analysis) has been done in the past, a very few authors have partially shown the effect of doping on sensitivity[10,11]. The effect of channel doping on the sensitivity of a junctionless biosensor was recently discussed[8]. Recently, a few authors have reported nanotube FET-based biosensors for detecting the biomolecule species using different sensing metrics[4,12]. Authors, such as Yadav *et al.*[13], Sharma *et al.*[9], Kumari *et al.*[14], Goel *et al.*[15] and many more, have recently reported different types of FET-

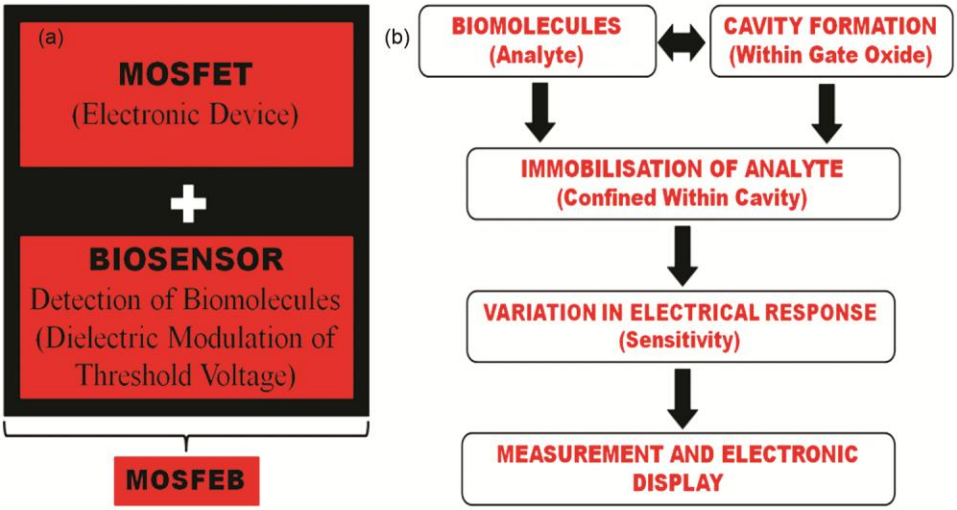


Fig. 1 — (a) Device principle, and (b) flowchart of the generalized biosensing phenomenon in NWM-BD[8,9]

based biosensors based on dielectric engineering. Upon reviewing various literature sources, it becomes apparent that researchers have investigated various potential variants for BioFETs. However, detailed exploration of doping-dependent sensitivity modulation remains largely unexplored in the existing research. Hence, the novelty of this paper as well as main goal of this work is to examine the sensitivity modulation in the nanowire MOSFEB because of doping. Doping becomes critically important in controlling and modulating sensitivity, as it is a parameter that can be adjusted before fabrication[16], highlighting the significance of this research. The second objective is to explore the other side of the nanowire MOSFEB detector, which includes studying the effect of temperature, structural parameters, trap charges, drain bias, source/channel/drain engineering techniques as well as steric hindrance on the sensitivity.

2 Device Structure and Computational Framework

2.1 MOSFEB Specifications

3D and 2D view of the nanowire MOSFEB detector is shown in Fig. 2 (a-b), respectively. The proposed detector, employing dielectric engineering by immobilizing biomolecules within the embedded cavity, is specifically designed for detecting various charged and neutral biochemical species. The measurable response of the detector is defined by changes in threshold voltage or subthreshold slope, although several other metrics have also been taken into consideration. A poly-silicon gate (work function: 4.96 eV) has been used, which has the advantage regarding fabrication simplicity and increased operating temperature range over the metallic gate. A symmetric and two-sided cavity of dimensions 14*2 nm^2 is embedded inside SiO_2 layer[8,12]. Figure 2(c) shows the detailed fabrication steps[17-19]of the proposed biosensor in the form of a flowchart. A precise cavity with the specified dimensions can be created using multiple rounds of dry etching, without compromising the quality of the gate oxide[20]. The biosensor features a dual cavity of 14×2 nm², enabling immobilization of biomolecules smaller than 14 nm, with the maximum detectable size depending on orientation (up to 14 nm horizontally and 2 nm vertically). This study primarily focuses on the upper detection limit, assuming full cavity occupancy while accounting for steric hindrance effects.

2.2 Simulator Setup and Specifications

The proposed biosensor's behavior is examined using parameters specified in Table 1 through the SILVACO TCAD simulator[21]. Various models are integrated into the simulation: the SRH model manages carrier generation and recombination, the CONMOB model correlates impurity profiles with low-field mobility at room temperature, and the AUGER model accommodates recombination effects at high doping densities. Additionally, the CVT model encompasses dependencies of mobility on doping, temperature, and field, while the FLDMOB model addresses field-dependent mobility. The BGN model handles bandgap narrowing attributed to degenerate as wellas high doping levels, and NEWTON

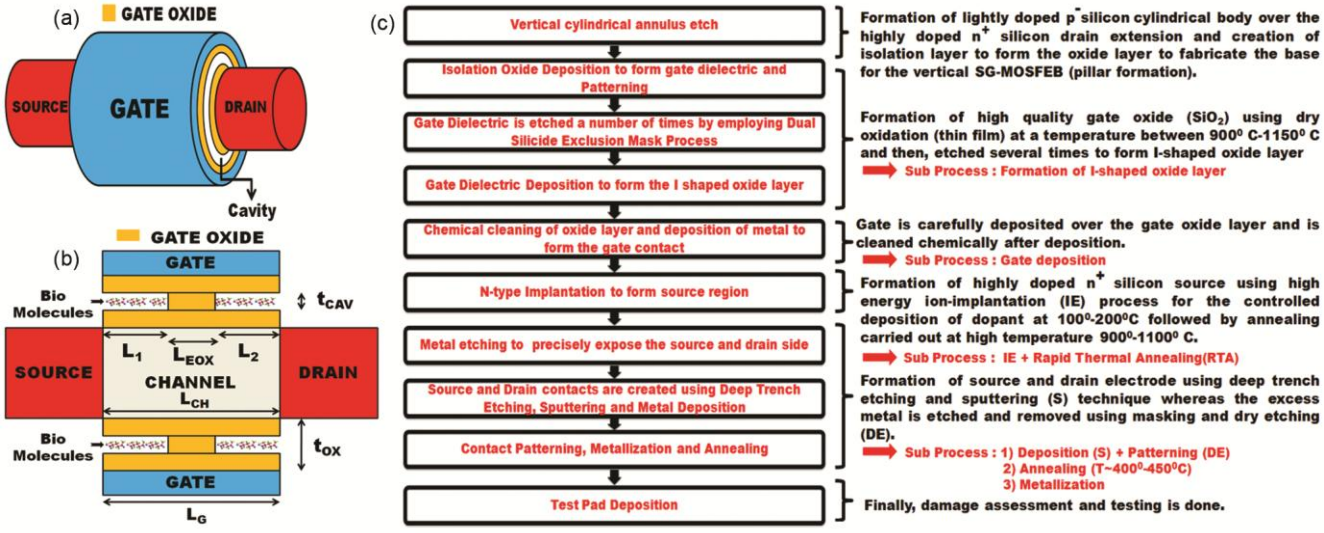


Fig. 2 — (a) 3D view, (b) 2D view, and (c) fabrication process flow-chart illustrating the basic fabrication steps of NWM-BD[20]

GUMMEL method is implemented to resolve the intricate, nonlinear coupled equations.

2.3 Methodology

Charge density (ρ) as well as dielectric constant (K_{bio}) is used to characterize various biomolecules. Biomolecules are either neutral (ρ=0) or charged (ρ≠0) with some finite dielectric constant. The different test biomolecules used in this study are mentioned in Table 1. Here, ρ of the biomolecule is kept constant, while keeping constant K_{bio}=5 to cover a large number of biomolecules (positively/neutral/negatively charged biomolecules). It is assumed that the cavity is occupied by air (dielectric constant of air is taken as K_{air}=1) in the absence of biomolecules[9,12]. When the test biomolecule (K_{bio}≠1) gets immobilized inside the cavity, threshold voltage (along with other device characteristics) gets changed because lateral field and gate oxide capacitance undergo changes when biomolecules are localized inside the cavity. Thus, the potential distribution gets modified across the channel[8, 24]. The relative difference in the sensing parameter's value is used as the qualitative indicator in the biosensing investigation of a nanowire MOSFEB detector. The numerical computational methodology used for the sensitivity analysis is shown in Fig. 3 as a flowchart. The relative change in the sensing metric (first formula) is preferred over fractional or percentage change for evaluating the biosensing performance. It provides a more accurate measure of sensitivity by directly comparing baseline and sensing values, whereas fractional or percentage change is tied to relative values, which may misrepresent true sensitivity.

Table 1 — Structural parameters of NWM-BD

Parameters		Value
Cavity thickness (t_{cav})		2-8 nm
Gate length (L_G)		30 nm
Source doping (N_S)		$1*10^{13}$ to $5*10^{20}$/cm^3
Channel length (L_{CH})		30 nm
Material	Source	n-type
	Channel	p-type Silicon
	Drain	n-type
Channel doping (N_C)		10^{10}/cm^3
Drain doping (N_D)		$1*10^{13}$ to $5*10^{20}$/cm^3
Drain/Source length		10 nm
Total oxide layer thickness (t_{ox})		4-10 nm
Biasing		V_{DS}=(0-1) V V_{GS}=(0-1) V
Drain/Channel/Source radius		10 nm
Cavity length		L_1= 14nm L_2= 14nm
Gate oxide		SiO_2
Gate work-function		4.96 eV
Oxide length (L_{OX})		L_{OX}= 30nm
BIOMOLECULES		
Hydroprotein(K_{bio}=5)[15]		
DNA (K_{bio}=5 & ρ= -1*10^{11}/cm^2)[22]		
Amino acids (K_{bio}=5 & ρ= +1*10^{11}/cm^2)[23]		

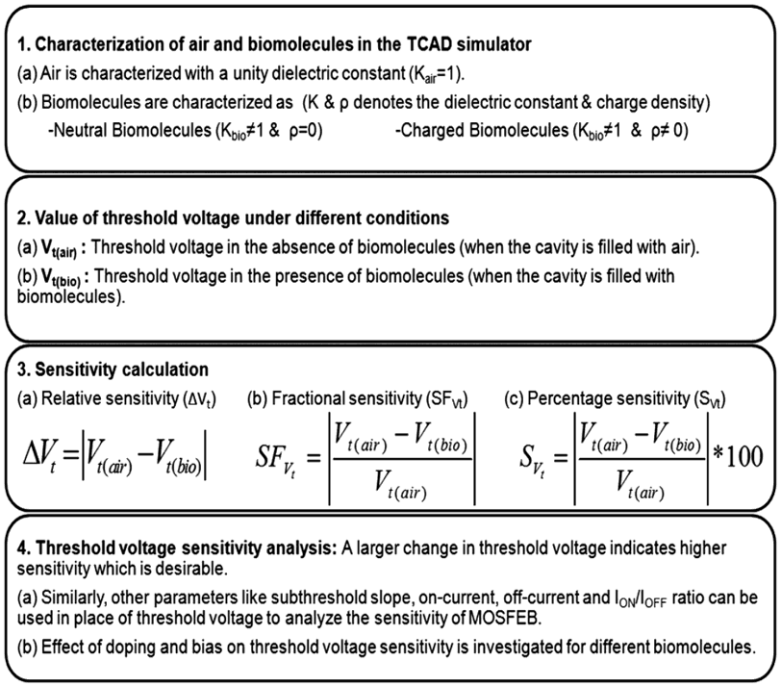


Fig. 3 — Computational methodology flowchart for sensitivity analysis

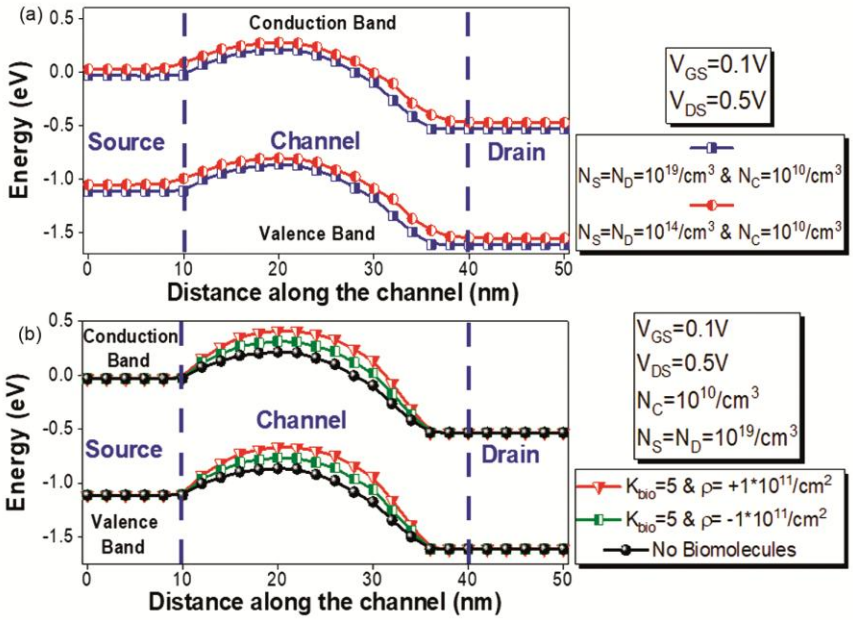


Fig. 4 — (a) Energy band diagram at various drain/source doping in the absence, and (b) presence of biomolecules in NWM-BD

3 Results and Discussion

3.1 Energy Band Diagram and Potential

Figure 4 illustrates the shift across energy bands within the nanowire MOSFEB detector across various operational conditions. In Fig. 4 (a), changes in drain/source doping levels lead to shifts in energy band positions; higher doping levels in the source increase electron density, pulling the bands downwards (Fermi level tends to moves inside the conduction band in n-type drain/source). Meanwhile, Fig. 4 (b) demonstrates how different biomolecules inside the cavity affect the positions of energy bands, influencing both the lateral electric field and gate capacitance in the channel. This interaction results in slight adjustments to the energies of both valence band as well as conduction band[8]. Contour plots in Fig. 5(a-c) depicts electron concentration profiles for negatively, neutral, and positively charged bio-chemical species.

Figure 6 presents the electron concentration contour plot across the channel in the nanowire MOSFEB detector, depicting different doping conditions. In Fig. 6(a), the contour plot shows the density of electrons at $N_C=10^{10}/cm^3$ and $N_D=N_S=10^{14}/cm^3$. In contrast, Fig. 6(b) shows the electron concentration contour plot at $N_S=N_D=10^{19}/cm^3$ and $N_C=10^{10}/cm^3$. It is evident that the peak electron concentration is localized near the channel surface. The localization of peak electron concentration near the channel surface underscores

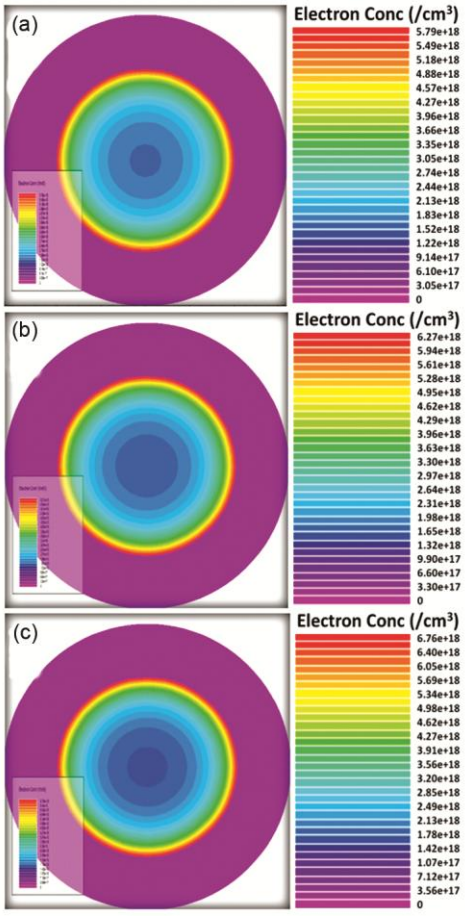


Fig. 5 — Contour plot of density of electrons (K_{bio}=5, $N_C=10^{10}/cm^3$,$N_D=N_S=10^{19}/cm^3$& $V_{DS}=V_{GS}$=1V) at different charge density (a) ρ=-1*10^11/cm^2, (b) ρ=0, and (c) ρ=+1*10^11/cm^2 of biomolecules in NWM-BD

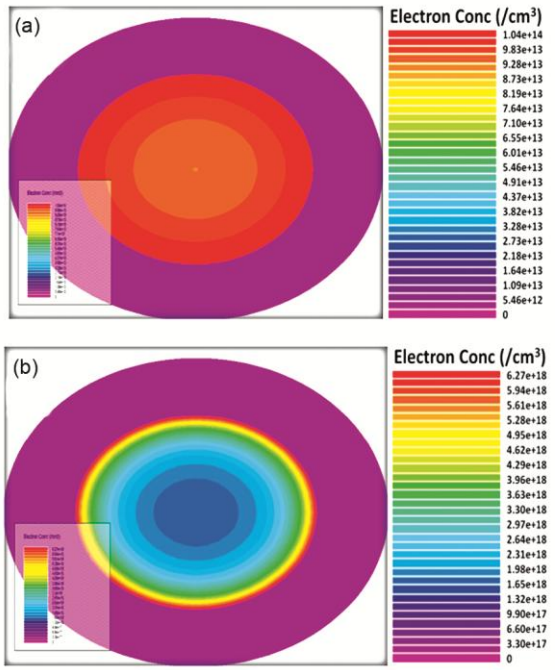


Fig. 6 — Contour plot of density of electrons (K_{bio}=5 & V_{GS}=V_{DS}=1V) at (a) N_C=10^{10}/cm^3, N_S=N_D=10^{14}/cm^3, and (b) N_C=10^{10}/cm^3, N_D=N_S=10^{19}/cm^3 in NWM-BD

the importance of surface phenomena in MOSFEB. Increasing the source doping enhances the electron concentration, thereby influencing the drain current[25]. The presence of bio-chemical species changes the electron concentration near the Si-SiO_2 interface by increasing the gate capacitance. The influence of biochemical species on the gate capacitance, as evidenced by the electron concentration changes, underscores the device's capability for label-free detection[8,26,27].

Figure 7 of the nanowire MOSFEB detector, the potential profile across the channel is illustrated. The main plot illustrates variations in the potential profile due to different doping levels of the source and drain. Meanwhile, the inset focuses on the potential profile in the presence of different biomolecules. Modifying the doping levels of the drain and source adjusts the potential barriers at the drain-channel and source-channel junctions, influencing the overall potential distribution across the channel[25,28]. The presence of biomolecules changes the gate capacitance and field profile, which in turn modifies the channel potential. Increasing the doping levels of the drain and source results in a notable rise in the channel's potential.

3.2 Effect of Equal Drain and Source Doping

Increasing the doping levels of both the source and drain results in two primary effects[16,29]. As the doping

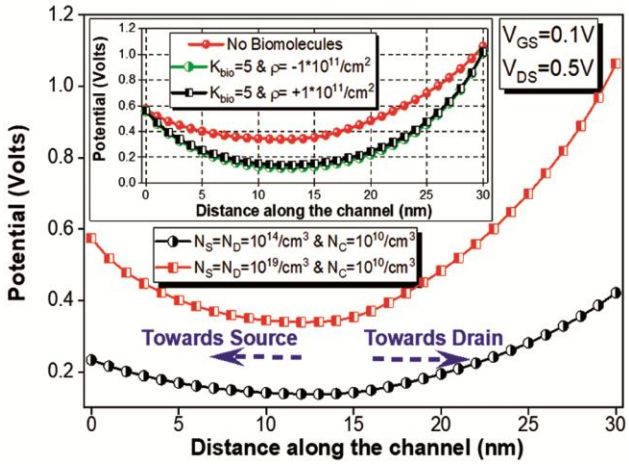


Fig. 7 — Potential profile across the channel at different source/drain doping in NWM-BD(*[INSET] Shows the potential profile across the channel in the presence of different biomolecules*)

levels of the drain and source are raised uniformly from low to high values, 'Effect 1' is observed, where a significant portion of the depletion region extends into the channel. 'Effect 2', on the other hand, involves a gradual reduction in the overall width of the depletion layer with increasing doping. Thus, both 'Effect 1' and 'Effect 2' become evident as the doping level of the drain and source increases. 'Effect 2' predominates at lower doping levels, reducing the depletion layer width and effectively increasing the channel length. Conversely, at higher doping levels, 'Effect 1' becomes more prominent, causing the depletion region to penetrate deeper into the channel and thereby decreasing the effective channel length. Figure 8 visually illustrates both 'Effect 2' and 'Effect 1'. The notation a_b_c denotes the doping levels of the source ($1*10^a$ cm^{-3}), channel ($1*10^b$ cm^{-3}), and drain ($1*10^c$ cm^{-3}), respectively, while S_C_D refers to source_channel_drain.

Table 2 presents data on the threshold voltage and its corresponding sensitivity across various doping levels (with same drain and source doping), considering different biomolecules filling the cavity. Initially, as doping increases, 'Effect 2' predominates by narrowing the depletion width. This reduces the depletion region coverage at the channel ends, thereby increasing the effective channel length. Consequently, a higher gate voltage is required to activate the device, resulting in an increase in the threshold voltage. As doping levels further increase, 'Effect 1' becomes more dominant. With higher source and drain doping compared to the channel, the depletion region extends significantly into the channel at both

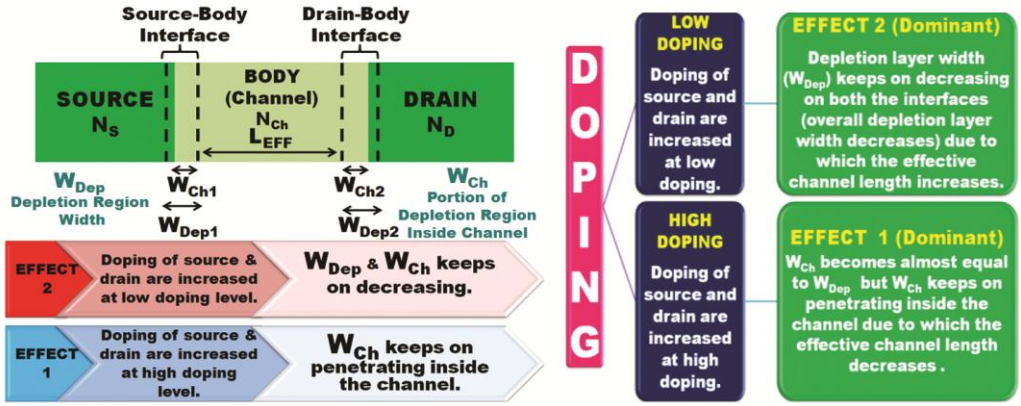


Fig. 8 — Visualization of the effect of doping on effective channel length in MOSFEB

Table 2 — Threshold voltage and threshold voltage sensitivity at different doping

	Doping (cm^{-3}) Source_Channel_Drain	Threshold Voltage (V_t : Volt)			Threshold Voltage Sensitivity (S_{Vt} : Percentage)		
		k_{bio} = 5					
		ρ $-1*10^{11}$ cm^{-2}	ρ 0	P $+1*10^{11}$ cm^{-2}	ρ $-1*10^{11}$ cm^{-2}	ρ 0	P $+1*10^{11}$ cm^{-2}
D_a	10^{13}_10^{10}_10^{13}	0.0459207	0.045366	0.0429468	453.0462	446.3645	417.2300
D_b	10^{14}_10^{10}_10^{14}	0.141106	0.134749	0.121702	264.6235	248.1967	214.4828
D_c	10^{15}_10^{10}_10^{15}	0.208939	0.202324	0.200550	65.7943	60.5452	59.1376
D_d	10^{16}_10^{10}_10^{16}	0.244491	0.241664	0.236106	49.4078	47.6803	44.2838
D_e	10^{17}_10^{10}_10^{17}	0.330133	0.315766	0.286625	44.9446	38.6368	25.8424
D_f	10^{18}_10^{10}_10^{18}	0.388328	0.371414	0.349767	40.7057	34.5771	26.7336
D_g	10^{19}_10^{10}_10^{19}	0.430309	0.416805	0.402911	74.1571	68.6916	63.0684
D_h	10^{20}_10^{10}_10^{20}	0.411948	0.399756	0.387387	133.8607	126.9394	119.9176

junctions. This shortens the effective channel length, requiring less gate voltage to activate the device and causing a reduction in the threshold voltage[16].Hence, the threshold voltage initially rises with doping, reaching a maximum value, but beyond a certain doping level, it begins to decrease, as depicted in Fig. 9(a). As seen in Fig. 9(b), although the sensitivity obtained is quite high at low doping, the conductivity of the device is quite low at such low doping, which limits the selection of low doping for MOSFET. The doping values D_a-D_h are specified in Table 2. In the current study, the threshold voltage sensitivity modulus is considered and given in Eq. 1. Since sensitivity is measured in terms of magnitude (i.e., the change in the sensing metric) rather than the sign of this change, taking the modulus ensures a consistent and comparative assessment of biosensor performance across different operational conditions.

$$S_{V_t} = \left| \left(\frac{V_{t(air)} - V_{t(bio)}}{V_{t(air)}} \right) * 100 \right| \quad \ldots (1)$$

At lower doping levels, 'Effect 2' tends to be predominant[30]. Increasing the doping of both the source and drain reduces the depletion width at the junctions of the drain-channel and source-channel, resulting in an increase in the depletion capacitance (C_{dep}). The subthreshold swing, inversely related to the subthreshold slope and directly proportional to C_{dep}, consequently increases as the depletion width decreases. This leads to a decrease in the subthreshold slope at lower doping levels[29,31]. At higher doping levels, 'Effect 1' becomes dominant, causing an increase in the subthreshold slope. A local minimum in the subthreshold slope can be seen in Fig. 10.

Table 3 illustrates the subthreshold slope observed at various symmetric doping levels for the source and drain. This metric is influenced by the doping levels of the drain, channel, and source, as well as by the presence of biomolecules within the cavity. Initially, the subthreshold slope sensitivity (S_{SS}) decreases until it reaches a minimum value, reflecting the relative difference $|SS_{(air)} - SS_{(bio)}|$ (where $SS_{(air)}$ and $SS_{(bio)}$

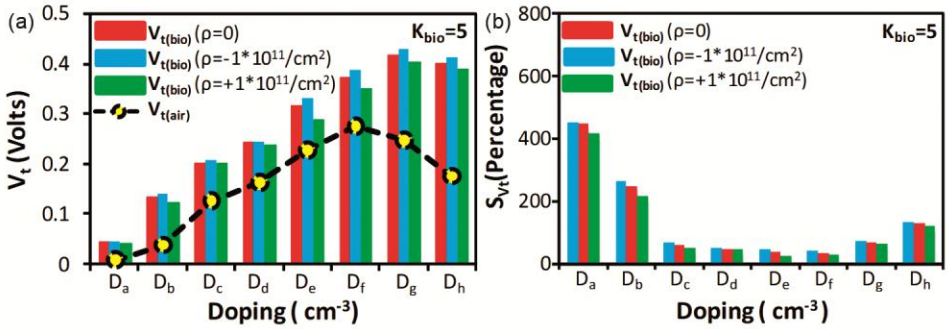


Fig. 9 — (a) V_t, and (b) S_{Vt} at different doping (N_C=10^{10}/cm^3, V_{DS}=0.5V & K_{bio}=5) in NWM-BD

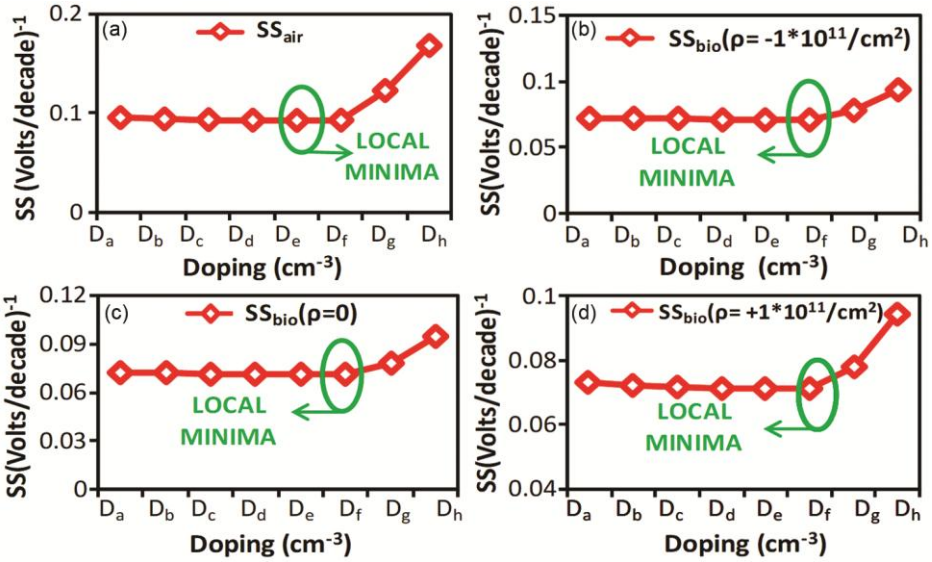


Fig. 10 — SS variation in the NWM-BD for (a) air, (b) negatively charged (ρ=-1*10^{11}/cm^2), (c) neutral (ρ=0), and (d) positively charged (ρ= +1*10^{11}/cm^2) biomolecules at different doping (K_{bio}=5, V_{DS}=0.5V, K_{bio}=5 & N_C=10^{10}/cm^3)

Table 3 — Subthreshold slope and subthreshold slope sensitivity at different doping

	Doping (cm^{-3}) S_C_D	Subthreshold Slope (SS : (Volts/decade)$^{-1}$)			Subthreshold Slope Sensitivity (S_{SS} : Percentage)		
		k_{bio} = 5					
		ρ -1*10^{11} cm^{-2}	ρ 0	P +1*10^{11} cm^{-2}	ρ -1*10^{11} cm^{-2}	ρ 0	P +1*10^{11} cm^{-2}
D_a	10^{13}_10^{10}_10^{13}	0.07194	0.0720	0.0729	25.1176	25.0421	24.0700
D_b	10^{14}_10^{10}_10^{14}	0.07193	0.07196	0.07196	24.0830	24.0579	24.0545
D_c	10^{15}_10^{10}_10^{15}	0.07148	0.07142	0.07141	23.9825	24.0456	24.0495
D_d	10^{16}_10^{10}_10^{16}	0.07128	0.07128	0.07128	23.8258	23.8267	23.8257
D_e	10^{17}_10^{10}_10^{17}	0.07121	0.07121	0.07121	23.6722	23.6692	23.6660
D_f	10^{18}_10^{10}_10^{18}	0.07114	0.07115	0.07115	23.9660	23.9640	23.9606
D_g	10^{19}_10^{10}_10^{19}	0.07747	0.07761	0.07776	37.1150	36.9970	36.8762
D_h	10^{20}_10^{10}_10^{20}	0.09395	0.09407	0.09423	44.1769	44.1020	44.0106

indicate the subthreshold slope with the cavity filled with air and biomolecules, respectively). Further equal increase in the doping rate leads to an enhancement in the S_{SS}. This trend is depicted in Fig. 11(a), 11(b), and 11(c), which shows the S_{SS} with negatively charged, neutral, and positively charged biomolecules in the cavity, respectively. The presence of charged biomolecules alters the flat-band voltage, thereby modifying the channel's field and potential profile[8, 15]. Table 3 also provides details on the

subthreshold slope sensitivity at different doping levels when the cavity immobilizes various biomolecules. Equation (2) outlines the formula for subthreshold slope sensitivity, emphasizing the absolute value to underscore the biosensor's sensitivity to relative changes rather than absolute metrics.

$$S_{SS} = \left| \left(\frac{SS_{(air)} - SS_{(bio)}}{SS_{(air)}} \right) * 100 \right| \quad \ldots(2)$$

3.3 Effect of Unequal Source and Drain Doping

To explore the effects of differing doping levels, the source and drain doping are independently adjusted (varied). A reference doping level (RD) is set with the channel doping equals to $1*10^{10}/cm^3$ and drain/source doping equals to $1*10^{19}/cm^3$. When the drain doping is increased beyond the reference level, channel as well as source doping remains unchanged. Conversely, while increasing the source doping above the reference level, channel as well as drain doping are held constant. Table 4 outlines the specific doping levels for D_{D1}-D_{D3}, D_{RD}, and D_{S1}-D_{S3}.

Increasing N_D above the reference level reduces the effective channel length, requiring less gate voltage to activate the device and resulting in a lower threshold voltage. Figure 12 displays how the threshold voltage varies without and with biomolecules. Biomolecules increase the threshold voltage due to the formation of a strong lateral electric field across the channel. Consequently, the fractional change $|V_{t(air)} - V_{t(bio)}|/V_{t(air)}$ increases as drain doping levels rise. This relationship is visually represented in Fig. 13, which illustrates the sensitivity of threshold voltage variation with respect to N_D.

As N_D increases beyond the reference level, the depletion region extends further into the channel at the drain junction[29,31]. This extension results in a reduction of the depletion capacitance (C_{dep}), leading to a decrease in the subthreshold swing while increasing the subthreshold slope. Figure 14(a) illustrates this trend. Additionally, Fig. 14 (b) shows that the fractional change $|SS_{(air)} - SS_{(bio)}|/SS_{(air)}$ increases as the drain doping level exceeds the reference doping.

As previously mentioned, when N_S is increased, Fig. 12 illustrates a decrease in the threshold voltage. The fractional change $|V_{t(air)} - V_{t(bio)}|/V_{t(air)}$ increases with higher source doping levels, as shown in Fig. 13. Furthermore, high N_S lead to a deeper penetration of the depletion region at the source-channel junction into the channel. This diminishes the C_{dep}, causing an increase in the subthreshold slope, as depicted in Fig. 14(a). Additionally, Fig. 14(b) illustrates that the subthreshold slope sensitivity also rises with

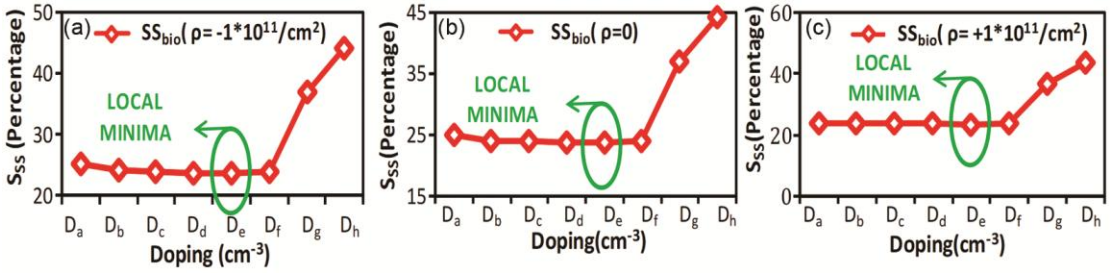


Fig. 11 — S_{SS} variation in the NWM-BD for (a) negatively charged (ρ=-1*10^{11}/cm^2), (b) neutral (ρ=0), and (c) positively charged (ρ=+1*10^{11}/cm^2) biomolecules at different doping (V_{DS}=0.5V, N_C=10^{10}/cm^3& K_{bio}=5)

Table 4 — Threshold voltage sensitivity & subthreshold slope sensitivity at different doping

	Doping (cm^{-3}) Source_Channel_Drain		Threshold Voltage Sensitivity (S_{Vt} : Percentage) — k_{bio} = 5			Subthreshold Slope Sensitivity (S_{SS} : Percentage) — k_{bio} = 5		
			ρ -1*10^{11} cm^{-2}	ρ 0	P +1*10^{11} cm^{-2}	ρ -1*10^{11} cm^{-2}	ρ 0	P +1*10^{11} cm^{-2}
D_{S3}	5*10^{20}_1*10^{10}_1*10^{19}		96.4704	90.7895	85.0216	40.5620	40.4209	40.2651
D_{S2}	1*10^{20}_1*10^{10}_1*10^{19}	N_S ↑	90.9764	85.4632	79.8665	39.5316	39.4166	39.3125
D_{S1}	5*10^{19}_1*10^{10}_1*10^{19}		87.5628	81.7503	76.6425	39.1903	39.1074	39.0235
D_{RD}	1*10^{19}_1*10^{10}_1*10^{19}	RD	74.1571	68.6916	63.0684	37.1150	36.9970	36.8762
D_{D1}	1*10^{19}_1*10^{10}_5*10^{19}		95.5378	89.2265	82.7258	40.8841	40.7662	40.6425
D_{D2}	1*10^{19}_1*10^{10}_1*10^{20}	N_D ↓	101.5939	95.0392	88.2835	41.4526	41.3113	41.1651
D_{D3}	1*10^{19}_1*10^{10}_5*10^{20}		110.4973	103.582	96.4543	42.1210	42.0127	41.9018

increasing N_S compared to the reference doping. In MOSFEBs, high doping in the source and low doping in the channel enhance charge carrier transport due to a reduced energy band gap. This doping configuration also increases the tunneling probability, altering the potential barrier at the source-channel junction, which is dependent on the electric field and gate/drain voltage. As a result, there is a larger current variation for the same range of applied voltage. The sensitivity of nanowire MOSFEB detector to biomolecules is attributed to the dependence of drain current, threshold voltage, and subthreshold swing on the source-to-channel potential barrier.

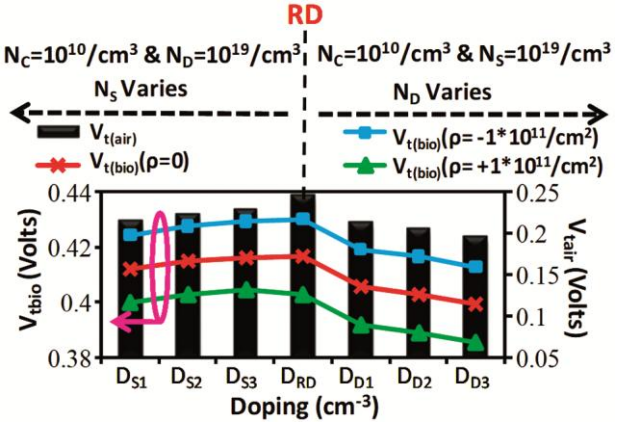


Fig. 12 — V_t variation in the NWM-BD when the cavity is filled with air and biomolecules (k_{bio}=5, N_C=10^{10}/cm^3& V_{DS}=0.5V)

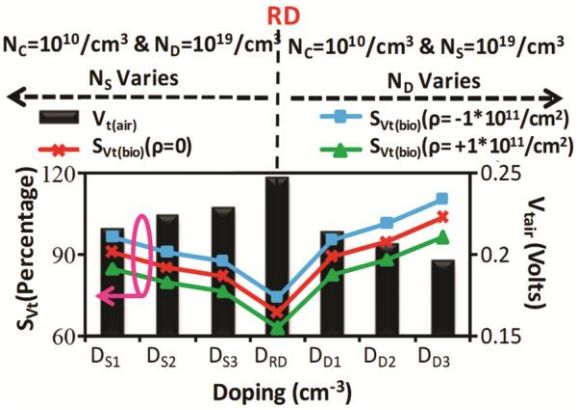


Fig. 13 — S_{Vt} in the NWM-BD for asymmetric doping variation (k_{bio}=5, N_C=10^{10}/cm^3& V_{DS}=0.5V)

Table 4 details the sensitivity of subthreshold slope and threshold voltage to different bio-chemical species under different drain/source doping conditions. Figure 15 illustrates the relative changes in subthreshold slope (ΔSS) as well as threshold voltage (ΔV_t) calculated at various doping levels. The findings highlight that negatively charged biomolecules exhibit pronounced ΔSS as well as ΔV_t absolute values, particularly at higher doping levels. As a sensitivity measure, threshold voltage yields a greater percentage sensitivity and relative change compared to subthreshold slope, rendering it more suitable for biosensing applications. Figure 16 depicts the variation in drain current (I_{DS}) at different doping, indicating enhanced current generation with increased doping due to the additional charge carriers generation.

Figure 17 illustrates the impact of V_{DS} on S_{SS} and S_{Vt} under symmetric and asymmetric doping conditions. For symmetric doping (SD), low source/drain doping (N_D=N_S=10^{19}/cm^3, N_C=10^{10}/cm^3) and high source/drain doping (N_D=N_S=10^{14}/cm^3, N_C=10^{10}/cm^3) are analyzed to observe the effects of drain biasing on sensitivity. Symmetric case (SD) features high doping (N_D=N_S=10^{19}/cm^3) and low doping (N_D=N_S=10^{14}/cm^3) levels, while asymmetric case (ASD) involves differing levels for the source and drain (N_D=10^{19}/cm^3, N_S=10^{20}/cm^3). Increasing the drain voltage correlates with increased S_{SS} and S_{Vt}. The observed increase in S_{SS} and S_{Vt} with rising V_{DS} can be attributed to the enhanced lateral electric field, which amplifies the impact of biomolecule-induced surface charge modulation on the channel

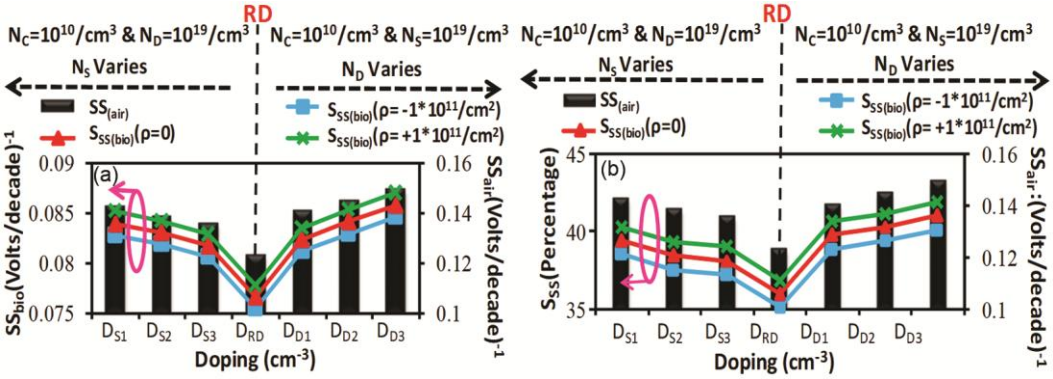


Fig. 14 — (a) SS, and (b) S_{SS} in the NWM-BD when the cavity is filled with biomolecules (k_{bio}=5, N_C=10^{10}/cm^3& V_{DS}=0.5V)

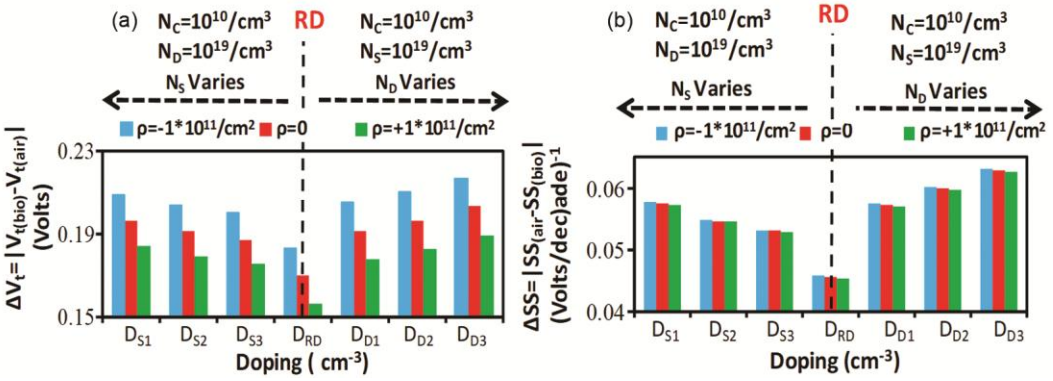


Fig. 15 — (a) Threshold voltage change, and (b) subthreshold slope change at different doping

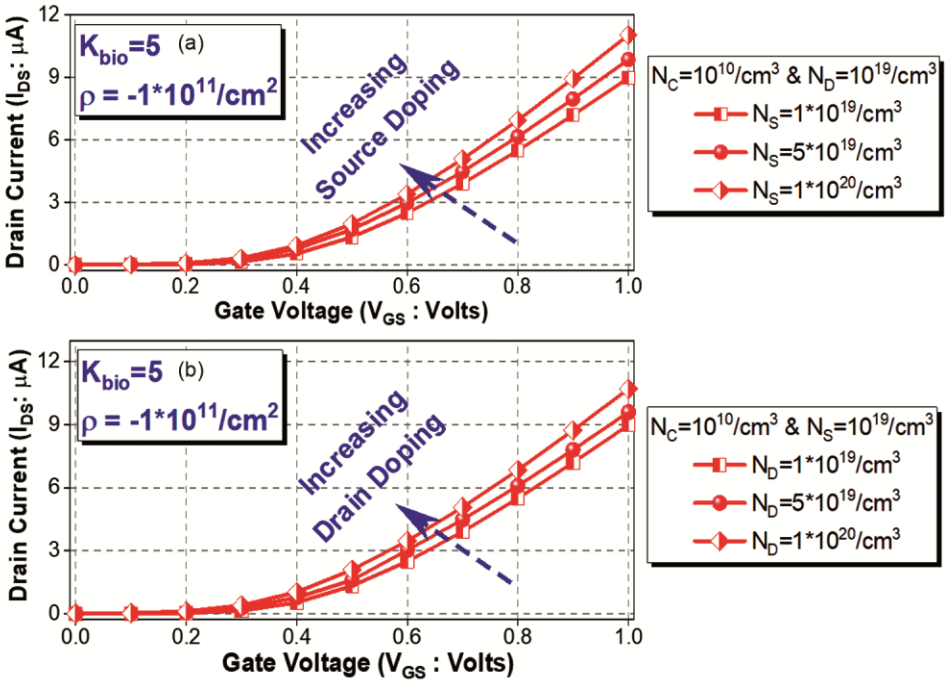


Fig. 16 — Drain current variation at different (a) N_S, and (b) N_D at V_{DS}=0.5V in NWM-BD

characteristics[24,32]. However, it is crucial to avoid exceeding a specific drain voltage limit to prevent device damage and ensure operational integrity. While higher V_{DS} generally improves sensitivity, it's important to note that this comes at the cost of increased power consumption and potential thermal effects, necessitating a careful balance in practical biosensing applications.

Table 5 shows the shift in the values of SS and V_t, which shows a higher relative change at high doping. It is quite evident from the above discussion that higher percentage change is obtained at low doping, whereas higher relative change is obtained at high doping. But, it should be kept in mind that conductivity is quite less at low doping and tunneling effect becomes prevalent at very high doping. Hence, an optimized doping should be considered for MOSFEB before fabrication. It is also interesting to note that the relative change or shift in the value of metric is a more reliable parameter in assessing the sensitivity of a MOSFEB detector. Relative change indicates the true value whereas percentage change indicates the relative fraction (modulated value) of the sensitivity.

3.4 Effect of Steric Hindrance and Cavity Dimensions

Steric hindrance is an important limiting factor in practical biosensors, resulting in a non-unity fill-in factor. The fill-in factor refers to the fraction of the total volume inside the cavity that is filled by the

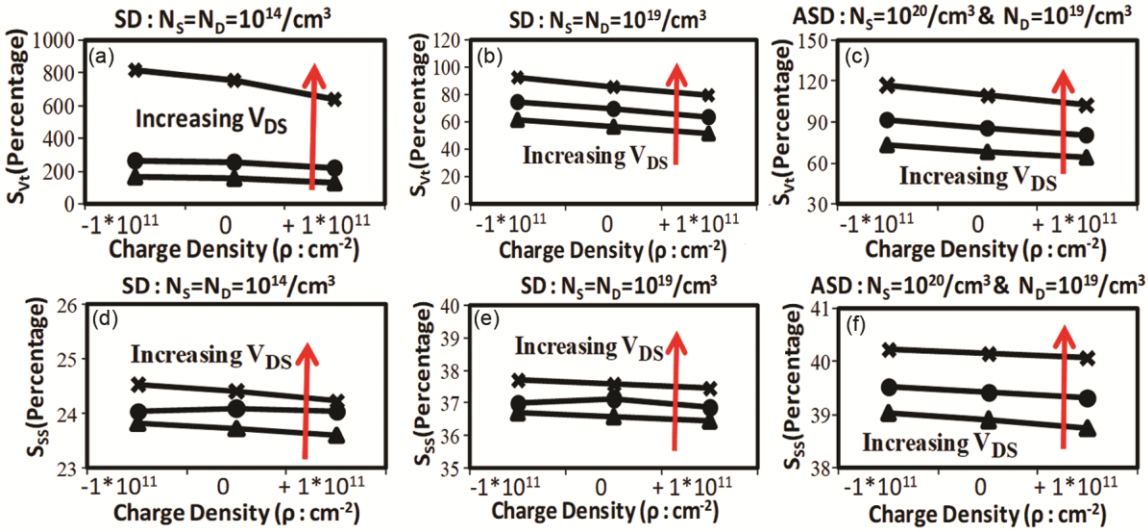


Fig. 17 — Impact of V_{DS} on (a-c) S_{Vt}, and (d-f) S_{SS} at different doping (k_{bio}=5, N_C=10^{10}/cm^3& V_{DS}=0.5V) in NWM-BD

Table 5 — Shift in sensing metrics in NWM-BD

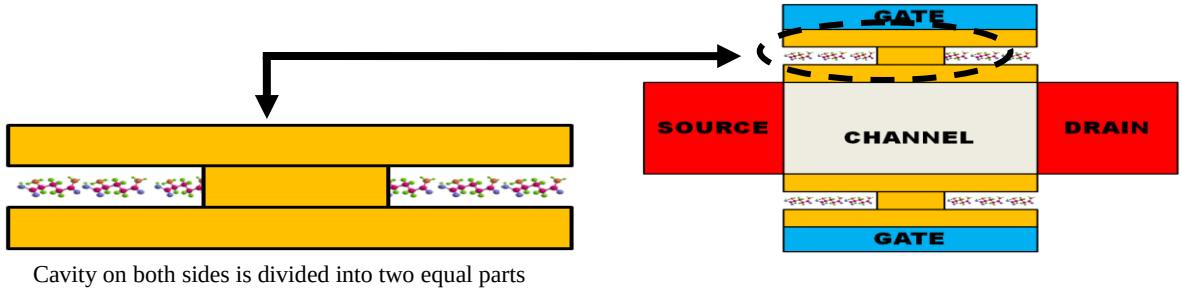
Doping	Change in V_t (mV)		
	A	B	C
D_a	37.61747	37.06267	34.64357
D_b	52.4069	46.0499	37.0029
D_c	78.916	76.301	49.527
D_d	80.851	78.024	52.466
D_e	102.368	88.001	58.86
D_f	112.342	95.428	73.781
D_g	183.228	169.724	155.83
D_h	235.797	223.605	211.236
A : K_{bio}=5 & ρ= -1*10^{11}/cm^2	B : K_{bio}=5 & ρ= 0	C : K_{bio}=5 & ρ= +1*10^{11}/cm^2	

biomolecules. The effect of steric hindrance is usually studied by considering different fill-in factors (such as horizontal or lateral filling) or fill-in profiles. Steric hindrance is a practical problem that arises from the partial hybridization of biomolecules. As a result, biomolecules only occupy a fraction of the cavity, impeding further filling of the remaining portion of the cavity. Hence, its effect must be thoroughly studied. Table 6 illustrates the impact of biomolecule location at different fill-in factors (horizontal filling) by dividing the cavity on both sides of the source and drain into four parts. Although a high fill-in factor generally results in high sensitivity[4] due to the immobilization of a large number of biomolecules, the location of the biomolecules significantly influences sensitivity, especially in partially filled cavities. For instance, although the fill-in factor of case 5 is more significant (higher) than that of case 4, the sensitivity in case 4 is higher, and the same observation can be made for other cases, such as case 10 and case 11. The influence of fill-in factor and profile on sensitivity underscores the importance of considering nano-scale phenomena in macroscopic device performance.

The fill-in factor of case 1 and case 4 are the same, but the location of the biomolecules is different in both cases. In case 1, the biomolecules are located near the source, resulting in higher sensitivity. This is because the threshold voltage is mainly affected by the potential hill appearing at the source-channel junction[28], and the biomolecules near the source affect this potential hill more than when they are near the drain. Similar observations can be made for other cases, such as case 5 and case 10, case 11 and case 14. The observed variations in sensitivity across different fill-in profiles offer valuable insights for developing computational models to predict and enhance biosensor responses. In the era of machine learning, these insights are crucial, as artificial intelligence can be harnessed to accurately predict and improve the sensitivity of biodetectors.

Figure 18 shows the effect of the fill-in factor (lateral filling), and Fig. 19 depicts how different fill-in profiles affect the threshold voltage sensitivity. These findings underscore the need for precise control over biomolecule immobilization techniques to achieve desired fill-in patterns and optimize sensor performance. A higher fill-in factor directly correlates with the increased sensitivity. The fill-in profile refers to the pattern filled by biomolecules within the cavity, with a symmetrical fill-in profile considered on both sides of the cavity. The concave fill-in profile has demonstrated the highest sensitivity compared to

Table 6 — Effect of fill-in factor and location of biomolecules on sensitivity



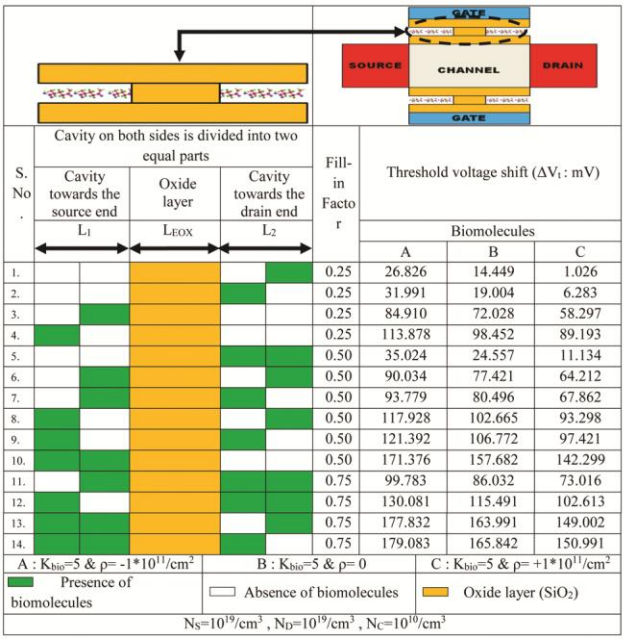
Cavity on both sides is divided into two equal parts

S. No.	Cavity towards the source end L_1	Oxide layer L_{EOX}	Cavity towards the drain end L_2	Fill-in Factor	Threshold voltage shift (ΔV_t : mV) Biomolecules A	B	C
1.				0.25	26.826	14.449	1.026
2.				0.25	31.991	19.004	6.283
3.				0.25	84.910	72.028	58.297
4.				0.25	113.878	98.452	89.193
5.				0.50	35.024	24.557	11.134
6.				0.50	90.034	77.421	64.212
7.				0.50	93.779	80.496	67.862
8.				0.50	117.928	102.665	93.298
9.				0.50	121.392	106.772	97.421
10.				0.50	171.376	157.682	142.299
11.				0.75	99.783	86.032	73.016
12.				0.75	130.081	115.491	102.613
13.				0.75	177.832	163.991	149.002
14.				0.75	179.083	165.842	150.991

A : K_{bio}=5 & ρ= -1*10^{11}/cm^2 B : K_{bio}=5 & ρ= 0 C : K_{bio}=5 & ρ= +1*10^{11}/cm^2

Presence of biomolecules Absence of biomolecules Oxide layer (SiO_2)

N_S=10^{19}/cm^3 , N_D=10^{19}/cm^3 , N_C=10^{10}/cm^3



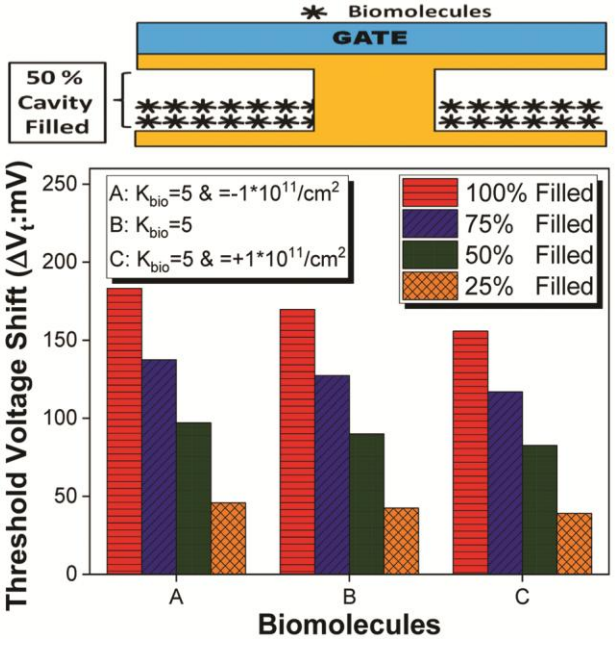


Fig. 18 — Impact of fill-in factor on sensitivity in NWM-BD

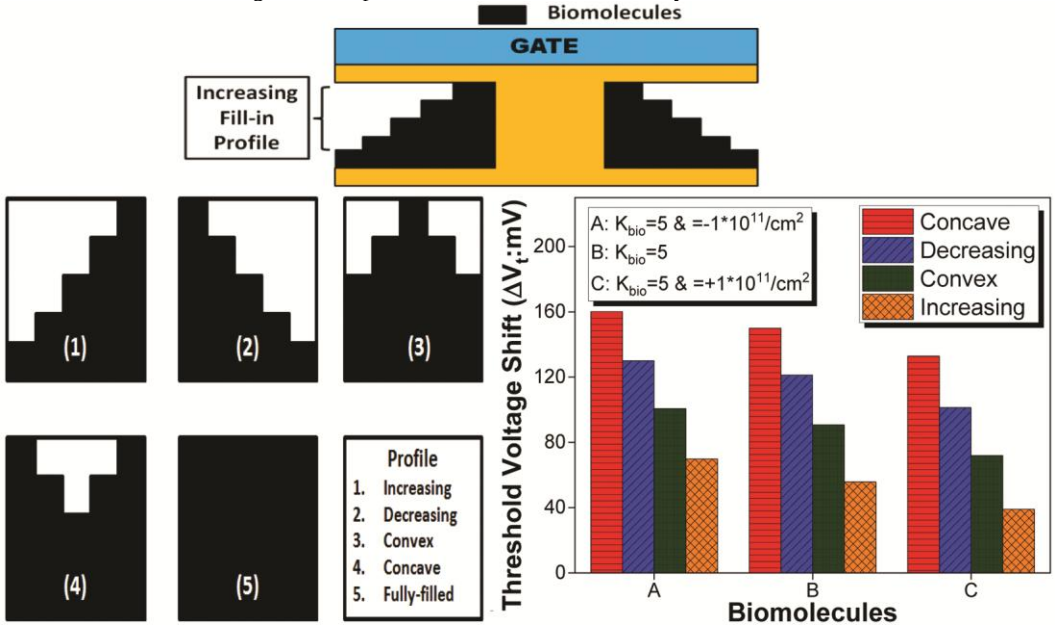


Fig. 19 — Impact of different fill-in profiles on sensitivity

Table 7 — Effect of cavity dimensions on sensitivity in NWM-BD

Parameter		Threshold voltage shift (ΔV_t : mV) A	B	C
Cavity Length (L_1=L_2)	14 nm	183.228	169.724	155.838
	12 nm	165.892	151.921	137.296
	10 nm	142.769	127.992	113.877
	08 nm	105.921	86.256	74.523
Cavity Thickness (t_{CAV})	02 nm	183.228	169.724	155.838
	04 nm	196.772	181.997	167.826
	06 nm	212.911	196.285	182.719
	08 nm	235.616	218.983	205.912

A : K_{bio}=5 & ρ= -1*10^{11}/cm^2 C : K_{bio}=5 & ρ= +1*10^{11}/cm^2
B : K_{bio}=5 & ρ= 0

other fill-in profiles. This sensitivity is due to the accumulation of biomolecules near the source-channel junction, which enhances the coupling between the gate and channel. As a result, the biomolecules near the source exerts stronger control over the flow of charge carriers, thereby increasing sensitivity. Further investigation into the kinetics of biomolecule immobilization and their spatial distribution within the cavity could lead to innovative strategies for improving sensor sensitivity and specificity.

Table 7 displays the impact of cavity thickness/length on sensitivity. The relationship between cavity dimensions and sensitivity underscores the importance of optimizing the sensing area in MOSFET-based biosensors. An increase in

Table 8 — Impact of parameters on sensitivity

Parameters		Threshold voltage shift (ΔV_t : mV) A	B	C
Radius of Channel (a: nm)	10	183.228	169.724	155.838
	11	169.491	155.669	141.287
	12	155.002	140.928	126.882
Operating Temperature (T: K)	250	178.982	164.509	151.008
	300	183.228	169.724	155.838
	350	189.945	175.672	160.429

A : K_{bio}=5 & ρ= -1*10^{11}/cm^2 C : K_{bio}=5 & ρ= +1*10^{11}/cm^2
B : K_{bio}=5 & ρ= 0

cavity size results in elevated sensitivity levels for both threshold voltage and subthreshold slope. This is due to the greater number of charge carriers immobilized within the cavity, which enhances the device's sensitivity to biomolecules. Further investigation could explore whether there's an optimal cavity size beyond which sensitivity gain diminishes, balancing performance with practical fabrication constraints.

Table 8 presents the influence of channel radius and temperature on the sensitivity. It is observed that as the channel radius increases, the sensitivity decreases. This reduction in sensitivity is attributed to the diminished control of the gate over the channel, primarily caused by the decrease in capacitive coupling. This trade-off between channel size and

sensitivity suggests that optimizing device geometry is crucial for balancing performance with manufacturability. However, it is important to note that the number of intrinsic charge carriers also increases with the increasing temperature, which leads to an increase in sensitivity. Therefore, while increasing the channel radius may have a negative impact on sensitivity due to reduced gate control, the increase in intrinsic charge carriers (due to an increase in temperature) counters this effect and contributes to an overall increase in sensitivity. The temperature dependent sensitivity increases underscores the need for precise temperature control in biosensing applications to ensure consistent and reliable measurements

3.5 Effect of Material Engineering

Material engineering also enhances the sensitivity of nanowire MOSFEB detectors[4]. Basically, material engineering refers to the i) use of different source/drain materials, ii) use of multi-gate architecture, iii) use of graded channel engineering, and iv) use of gate oxide stacking.

3.5.1 Source/Drain Engineering

Table 9 presents the threshold voltage sensitivity (threshold voltage shift) resulting from the use of different drain/source materials. This analysis examines how the conventional material, silicon, compares to alternative materials. The selection of drain/source materials also influences the device's resistance to biofouling and chemical degradation, factors that are essential for long-term stability in biological environments. This factor is crucial, as altering the drain and source materials affects the potential barrier at both the drain and source junctions. These junctions are pivotal since the charge carriers contributing to the total current are significantly influenced by these barriers at the channel-drain and channel-source interfaces. Among all the materials tested, silicon shows the highest sensitivity. Generally, materials with lower bandgaps are more sensitive to biomolecular interactions, whereas those with larger bandgaps exhibit more significant relative changes in sensing metrics[4,6,33,34]. However, each has its disadvantages: lower bandgap materials tend to increase leakage current (leading to false turn-on of the detector) and have lower breakdown voltages, while higher bandgap materials pose fabrication challenges due to high defect density and low carrier mobility, limiting their use in high-speed applications[4,34]. The interface between the chosen drain/source material and channel plays a significant role in determining the overall device performance, highlighting the need for careful material selection and interface engineering.

Table 9 — Impact of different materials on sensitivity in NWM-BD

Material			Biomolecules		
Source	Drain	Bandgap (300 K)	A	B	C
Si	Si	1.1 eV	183.228	169.724	155.838
Ge	Ge	0.7 eV	138.007	112.942	93.721
GaAs	GaAs	1.42 eV	114.408	94.589	77.298
InP	InP	1.351 eV	119.287	103.669	81.870
InSb	InSb	0.18 eV	72.115	57.346	40.379
GaSb	GaSb	0.726 eV	172.390	145.498	126.571
InAs	InAs	0.36 eV	160.284	139.330	115.690

A : K_{bio}=5 & ρ= -1*10^{11}/cm^2 C : K_{bio}=5 & ρ= +1*10^{11}/cm^2

B : K_{bio}=5 & ρ= 0

It is noteworthy that sensitivity is not solely dependent on the bandgap of the semiconductor material. Other factors, such as electron mobility, electron affinity, charge carrier density and doping profiles, also play crucial roles. Therefore, all these factors must be carefully analyzed before selecting the material for the source and drain. In general, materials with lower bandgaps have lower energy bandgaps but higher electron and hole mobility, whereas materials with higher bandgaps have higher energy bandgaps but lower mobility of electrons and holes. Consequently, silicon's sensitivity remains the highest. In the following next sections, the focus shifts primarily to three materials, specifically those with lower energy bandgaps: silicon, germanium, and indium arsenide (InAs). Germanium is considered one of the best alternatives to silicon, despite its disadvantages[35], while InAs is another promising compound semiconductor material with high potential for use in biosensor detectors[36,37].

3.5.2*Oxide Engineering*

Figure 20 demonstrates the influence of gate oxide stacking on threshold voltage sensitivity at t_1=1nm and t_2=3nm. It can be observed that sensitivity is increased when SiO_2 and high-K dielectric are stacked. The observed increase in sensitivity with gate oxide stacking can be attributed to the improved capacitive coupling between the gate and the channel, allowing for more effective modulation of the channel conductivity by biomolecular interactions[38]. The choice of high-K material and its thickness ratio with

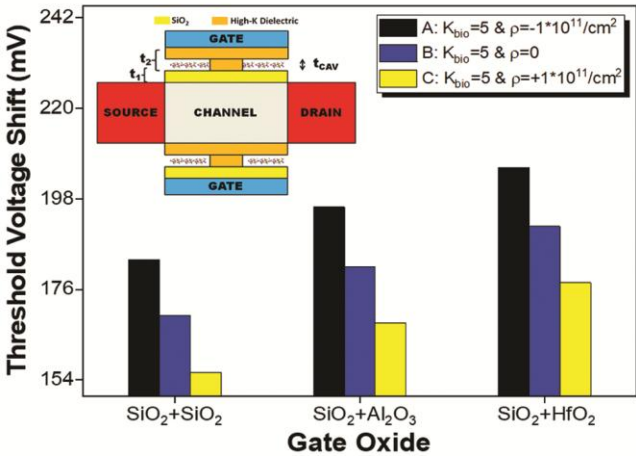


Fig. 20 — Effect of gate-oxide stacking on sensitivity in NWM-BD

SiO_2 can be further optimized to balance sensitivity gains with overall device performance and reliability. However, it is essential to remember that this approach also increases fabrication complexity, as high-k dielectric deposition is subject to certain limitations[8]. The increased fabrication complexity associated with high-K deposition might be mitigated by emerging deposition techniques or alternative high-K materials that are more compatible with standard CMOS processes.

3.5.3 *Gate and Channel Engineering*

Apart from doping and gate oxide stacking, sensitivity also increases due to dual gate architecture[39] and graded channel due to the enhanced electric field[40].To study the impact of gate engineering on sensitivity, dual gate architecture has been considered in nanowire MOSFEB with equal gate length for both gates (rest, all the parameters have been considered the same). The work function of gate-1 and gate-2 are ϕ_{G1} and ϕ_{G2} respectively. The work function of gate-1 (near to source) is greater than that of gate-2 (near to drain) so as to obtain high sensitivity[28]. The work function difference between the gates creates a built-in electric field along the channel, which can be optimized to enhance the device's response to biomolecular interactions. Reversing the value of the work function decreases the sensitivity of the biosensor since a high work-function gate effectively modulates the flow of charge carriers by controlling the source-channel potential barrier. As discussed in the previous section, the source-channel barrier highly affects the sensitivity as compared to the drain-channel potential barrier. The dual gate architecture not only improves sensitivity but also offers better control over the channel, potentially reducing short-channel effects in nanoscale devices. While dual gate architecture offers improved sensitivity, it's important to consider the trade-offs in terms of fabrication complexity and potential increase in parasitic capacitances[14,41,43].

For considering the effect of channel engineering, a dual-step graded doping profile has been considered for the channel. The channel has been divided into two equal parts with doping of N_{CH1} (part of the channel near to source) and N_{CH2} (part of the channel near to source). Lateral step-graded doping in the channel enhances the sensitivity by creating a doping gradient that improves charge carrier mobility. This gradient facilitates faster charge transport and increases the effective carrier concentration, resulting in enhanced current and voltage modulation[40]. Consequently, sensitivity increases due to the step-graded doping profile. Keeping $N_{CH2}>N_{CH1}$, the potential barrier at the source junction is lower (compared to the drain junction), which facilitates smooth transportation of charge carriers, and hence, high sensitivity is obtained when compared to the case where $N_{CH1}>N_{CH2}$[40]. The dual-step graded doping profile not only improves sensitivity but also helps in mitigating short-channel effects, which is crucial for scaling down device dimensions. While step-graded doping improves sensitivity, it's important to consider the challenges in precisely controlling doping profiles at nanoscale dimensions and their potential impact on device-to-device variability[44]. Table 10 shows the percentage improvement in sensitivity due to gate and channel engineering. Future research could explore more complex graded doping profiles, such as continuous gradients or multi-step profiles, to further enhance sensitivity and device performance.

3.6 Impact of Trap Charges, Cavity Location, and LOD

The device reliability is affected mainly by various factors, including trap charges (can adversely impact device performance by altering the electrical characteristics leading to inaccurate biosensor readings), off-current, and hot carrier injection.Trap charges in the gate oxide can lead to shifts in the threshold voltage and degradation in device performance over time. Trap charges may occur in the device as a result of defects during the manufacturing process. These interface trap charges decrease the device's reliability and accelerate the aging mechanism. Different source materials may exhibit varying susceptibilities to trap charges formation,

Table 10 — Impact of gate and channel engineering on sensitivity

Bio Molecules	ΔV_t (mV) With GE	ΔV_t (mV) Without GE	% Improvement
GE: Gate-Engineering ($L_G=L_{G1}+L_{G2}$ & $L_{G1}=L_{G2}$) ϕ_{G1}=5.1 eV[45] and ϕ_{G2}=4.1 eV[46]			
A	212.291	183.228	15.861
B	198.811	169.724	17.137
C	184.782	155.838	18.573
GE: Gate-Engineering ($L_G=L_{G1}+L_{G2}$ & $L_{G1}=L_{G2}$) ϕ_{G1}=4.1 eV[46] and ϕ_{G2}=5.1 eV[45]			
A	191.910	183.228	4.738
B	178.204	169.724	4.996
C	162.995	155.838	5.234

Bio Molecules	ΔV_t (mV) With CE	ΔV_t (mV) Without CE	% Improvement
CE: Channel-Engineering ($L_{CH}=L_{CH1}+L_{CH2}$ & $L_{CH1}=L_{CH2}$) $N_{CH1}=10^{10}/cm^3$ and $N_{CH2}=10^{15}/cm^3$			
A	204.577	183.228	11.651
B	190.004	169.724	11.948
C	176.813	155.838	13.459
CE: Channel-Engineering ($L_{CH}=L_{CH1}+L_{CH2}$ & $L_{CH1}=L_{CH2}$) $N_{CH1}=10^{15}/cm^3$ and $N_{CH2}=10^{10}/cm^3$			
A	189.001	183.228	3.150
B	175.618	169.724	3.472
C	161.279	155.838	3.812

A : K_{bio}=5 & ρ= -1*10^{11}/cm^2 C : K_{bio}=5 & ρ= +1*10^{11}/cm^2

B : K_{bio}=5 & ρ= 0

influencing the choice of materials for specific biosensing application. The impact of different interface trap charges (assumed across the length L_{EOX} across the oxide channel interface at doping $N_D=N_S=10^{19}/cm^3$ and $N_C=10^{10}/cm^3$) has been examined across various source materials for K_{bio}=5 in Table 11.

Mitigation strategies, such as optimized annealing processes or the use of high-quality interfacial layers, can help reduce the impact of trap charges on device reliability. Both positive and negative interface trap charges ($\pm$ 1*10^{12}/cm^2) have been considered on the channel-oxide interface[47]. Notably, the silicon source material has demonstrated superior reliability (lowest change in threshold voltage due to trap charges). The gate-all-around architecture also proven effective in mitigating several short-channel effects compared to the planar architecture. Hot carrier injection (causes oxide breakdown) is an undesired effect that affects long-term reliability. This is mainly due to the high electric field near the drain end. Table 12 displays the

Table 11 — Impact of trap charges on sensitivity

Source Material	Trap Charges	ΔV_t (mV)	Change in ΔV_t (mV)
Si	WPITC	166.992	2.732
	WITC	169.724	-
	WNITC	171.289	1.565
Ge	WPITC	168.529	5.447
	WITC	173.976	-
	WNITC	178.825	4.849
InAs	WPITC	170.091	6.791
	WITC	176.882	-
	WNITC	182.239	5.357

WPITC: With Positive Interface Trap Charges
WITC: Without Interface Trap Charges
WNITC: With Negative Interface Trap Charges

Table 12 — Peak electric field and off-current variation for different source materials

Source Material	Peak-Field (V/m)	I_{OFF} (pA)
Si	1.411*10^6	75.552
Ge	1.936*10^6	131.029
InAs	2.891*10^6	152.998

peak electric field near the drain end and off-current at V_{GS}=0V and V_{DS}=1V for K_{bio}=5. Excessive off-current contributes to power wastage and reduced overall device reliability.It can be seen that silicon has a lower off-current (due to its high band gap) and low electric field compared to other materials. Hence, the silicon is more reliable than other materials.

The limit of detection (LOD) refers to the minimum concentration of a biochemical species that a biosensor can detect. The analytical assessment of LOD involves the application of curve fitting and regression analysis techniques[48,49]. LOD is determined by analyzing the drain current sensitivity response for K_{bio}=5. In Fig. 21, the plot demonstrates the relationship between the fractional change in drain current ($\Delta I_{DS}/I_{DS}$) and the varying concentration range of the biomolecule within the cavity, spanning from 0.0001 to 1000 mg/ml. The inset compares the LOD of nanowire MOSFEB detector using different source materials. It is observed that the LOD obtained for the silicon source is approximately 11.61 / 27.24% higher than the LOD obtained for Ge/InAs source. Moreover, the stability and repeatability of the sensor's performance over multiple uses and different environmental conditions are critical aspects that ensure consistent and dependable detection capabilities. Continuous advancements in material science and nanotechnology contribute to the development of more efficient and precise biosensors,

further pushing the boundaries of LOD and enhancing their practical applications in various fields.

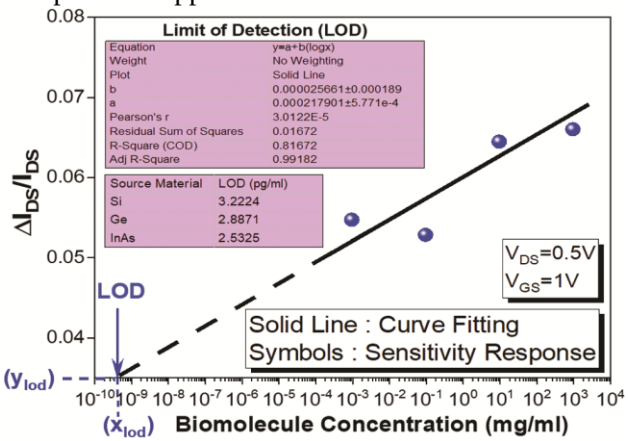


Fig. 21 — Current sensitivity response curve in NWM-BD

Table 13 — Impact of cavity location on sensitivity

Threshold voltage shift (ΔV_t : mV)					
Cavity Length			Biomolecules		
L_1 (nm)	L_{EOX} (nm)	L_2 (nm)	A	B	C
0	2	28	130.008	118.291	102.226
14	2	14	183.228	169.724	155.838
28	2	0	225.667	209.791	196.009

A : K_{bio}=5 & ρ= -1*10^{11}/cm^2 C : K_{bio}=5 & ρ= +1*10^{11}/cm^2
B : K_{bio}=5 & ρ= 0

Cavity location also plays a crucial role in affecting the sensitivity of nanowire MOSFEB detector. Table 13 shows the effect of cavity location on sensitivity (the aspect ratio is kept the same by keeping the total length L_1+L_{EOX}+L_2 constant). The biosensor is highly sensitive to biomolecules when the cavity is embedded near the source. As discussed in the previous section, biomolecules inside the cavity (near the source junction) effectively controls the flow of charge carriers via capacitive coupling[25, 50-51]. Hence, the biosensor becomes more sensitive to biomolecules. Advanced simulation and modeling techniques can aid in precisely determining the optimal cavity location and configuration to achieve the highest sensitivity and selectivity. These considerations are vital for the development of next-generation biosensors with superior performance metrics.

3.7 Comparison with Junctionless Variant

Junctionless variants have fabrication advantage over the inversion type variants[52]. Here, in Fig. 22, the sensitivity of the nanowire MOSFEB has been compared with its junctionless variant. The proposed MOSFEB shows over ~95mV improvement over the

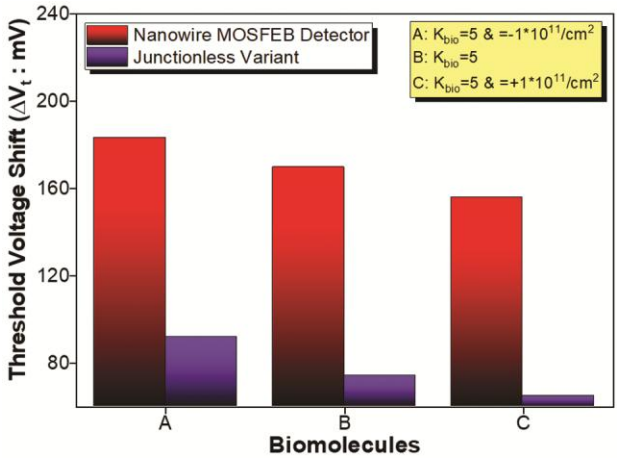


Fig. 22 — Comparison of NWM-BD with its junctionless variant

junctionless variant for K_{bio}=5. And the same trend can be seen for the charged biomolecules also. For junctionless variant, the doping and type of the semiconductor material used in this analysis is same for source/channel/drain (doping considered here is 10^{19}/cm^3).

Junctionless FET based-biosensors typically exhibit lower sensitivity in terms of threshold voltage changes compared to conventional inversion-mode FETs due to their distinct operating principles[53,54]. While inversion-mode MOSFETs rely on surface conduction in a thin inversion layer, junctionless MOSFETs operate through bulk conduction in a heavily doped channel. This bulk conduction mode makes junctionless devices less sensitive to surface charge changes induced by biomolecule interactions or localization inside the cavity. Additionally, the weaker gate control and reduced electrostatic coupling in junctionless devices result in a smaller impact on the threshold voltage when biomolecules interact with the detector[15]. Consequently, the threshold voltage in junctionless devices is less responsive to the presence of bio-chemical species, leading to lower sensitivity compared to their inversion-mode counterparts. Despite their lower threshold voltage sensitivity, junctionless MOSFET biosensor offers potential advantages in terms of simpler fabrication processes, reduced short-channel effects, and potentially lower noise levels, which can be beneficial in certain biosensing applications where these factors are prioritized over absolute sensitivity[1].

3.8 Sensitivity Analysis for Different Biomolecules

Different biomolecules have significance in the biomedical field. Biomolecules such as Uricase (K_{bio}=1.54), Streptavidin (K_{bio}=2.1), Biotin

Table 14 — Threshold voltage shift for different neutral biomolecules

Biomolecules	Reference	Sensitivity ΔV_t (mV)	Sensitivity S_{Vt} (Percentage)
Uricase	12	65.54873	26.529
Streptavidin	12	97.0815	39.291
Biotin	40	118.595	47.998
APTES	12	140.6225	56.913
Hydroprotein	15	169.724	68.691
Keratin	9	207.2831	83.892
Gelatin	24	254.7775	103.114

Table 15 — Benchmarking the performance of NWM-BD detector with similar biosensors

Structure	K_{bio}	S_{Vt} (V)	Reference
NWM-BD (Our Work)	12	0.2547	-
GaN-GME-DE-SNW-FET	8	0.1050	9
TGRC-MOSFET	8	0.0072	55
QG-MC-MOSFET	12	0.1610	56
TGAA-NWFET	2.1	0.0172	57

(K_{bio}=2.63), ChOx (K_{bio}=3.3), APTES (K_{bio}=3.57), Keratin (K_{bio}=8) and Gelatin (K_{bio}=12) are some of the commonly used neutral biomolecules.The sensitivity of the biosensor increases with the increasing dielectric constant (K_{bio}). High K_{bio} increases the coupling between the charge carriers and gate via capacitive coupling. Hence, the sensitivity increases with increasing K_{bio}. Table 14 presents the sensitivity comparison for these biomolecules. Table 15 is a benchmark for comparing the performance of nanowire MOSFEB detector to that of similar biosensors.

Label-free biosensors, including NWM-BD, inherently face challenges in achieving high selectivity. While our biosensor exhibits high sensitivity to the target analyte, selectivity can be enhanced by incorporating selective recognition elements, such as a bioreceptor layer, though this may increase fabrication complexity and reduce sensitivity. As selectivity remains a fundamental challenge in label-free biosensors, sensitivity is often the primary benchmark. The nanosensor (NWM-BD) functions by immobilizing biomolecules within a nanogap cavity, requiring advanced fabrication techniques for precise nanoscale dimensions, adding complexity to the process. However, the proposed nanosensor offers high sensitivity and cost-effectiveness compared to conventional biosensors, making it promising for label-free biosensing. Despite its high sensitivity, its selectivity is relatively low but can be improved with an additional bioreceptor layer at the cost of increased fabrication complexity. Additionally, repeated usage may cause wear and tear, potentially reducing the sensor's operational lifetime and reusability.

4 Conclusion

This paper presents a comprehensive sensitivity analysis of a nanowire-based MOSFEB detector, focusing on the impact of doping on sensitivity. The uniqueness of this study lies in its methodology, which investigates sensitivity enhancement through doping modulation and the effect of partially filled cavities (steric hindrance) at different fill-in factors and profiles. Both relative and percentage sensitivity are considered in the sensitivity exploration. The findings indicate that sensitivity exhibits a greater percentage change at lower doping levels, while the relative change peaks at higher doping levels. Various parameters, including cavity dimensions, material engineering, drain voltage, operating temperature, oxide stacking, source material, and channel radius, significantly influence sensitivity. The analysis of steric hindrance reveals that the location of biomolecules impacts sensitivity more than their fill-in factor. The detector shows a sensitivity enhancement of approximately 95.716 mV over its junctionless variant (for K_{bio}=5), with oxide and gate engineering techniques contributing more significantly to sensitivity enhancement than other engineering methods. The proposed biosensor achieves a limit of detection of 3.2224 pg/ml. Doping, bias, and biomolecule location are also the critical factors in determining sensitivity revealed through various simulated results. This work provides a valuable framework for studying and enhancing the sensitivity of similar biosensors, offering a thorough investigation into biosensing performance.

References

1 Sirohi A, Sahu C & Singh J, *IEEE Access*, 7 (2019) 141810.
2 Das A, Kanaujia B K, Nath V, Rewari S & Gupta R S, *IEEE 17th India Counc Int Conf INDICON*, (2020) 1.
3 Kaisti M, *Biosens Bioelectron*, 98 (2017) 437.
4 Das A, Rewari S & Kanaujia B K & Gupta R S, *Silicon*, 14 (2022) 5133.
5 Mendiratta N, Tripathi S L, *J Semicond*, 41 (2020) 061401.
6 Jianrong C, Yuqing M, Nongyue H, Xiaohua W & Sijiao L, *Biotechnol Adv*, 22 (2004) 505.
7 Syahir A, Usui K, Tomizaki K, Kajikawa K & Mihara H, *Microarrays*, 4 (2015) 228.
8 Chakraborty A & Sarkar A, *J Comput Electron*, 16 (2017) 556.

9 Sharma S, Nath V, Deswal S S & Gupta R S, *Microelectronics J*, 129 (2022) 105599.
10 Li J, Zhang Y, To S, You L & Sun Y, *ACS Nano*, 5 (2011) 6661.
11 Chiang M H, Lin C N & Lin G S, *Semicond Sci Technol*, 21 (2006) 190.
12 Pratap Y, Kumar M, Kabra S, Haldar S, Gupta R S & Gupta M, *J Comput Electron*, 17 (2018) 288
13 Yadav S & Rewari S, *Microsyst Technol*, 30 (2024) 565.
14 Kumari M, Singh N K & Sahoo M, *Sci Rep*, 12 (2022) 21115.
15 Goel A, Rewari S, Verma S, Deswal S S & Gupta R S, *IEEE Sens J*, 21(2021) 16731.
16 Das A, Kanaujia B K, Deswal S S, Rewari S & Gupta R S,*2022 IEEE Int Conf Electron Devices Soc Kolkata Chapter (EDKCON)*, (2022) 1.
17 Zhang Y, Han K & Li J, *Micromachines*, 11 (2020) 1.
18 Ganesh A, Goel K, Mayall J S & Rewari S, *Silicon*, 14 (2022) 4063.
19 Li C, Liu F, Han R & Zhuang Y, *IEEE Access*, 9 (2021) 63602.
20 Sze S M, Physics of semiconductor devices, (Wiley EasternLimited, New Delhi) 2nd Edn, (1983) 71.
21 Das A, Rewari S, Kanaujia B K, Deswal S S & Gupta R S, *Micro and Nanostructures*, 204 (2025) 208152.
22 Lodhi A, Rajan C, Kumar A, Dip B, Samajdar P & Soni D, *Appl Phys A*, 126 (2020) 837.
23 Rosini E, D'antona P & Pollegioni L, *Int J Mol Sci*, 21 (2020) 4574.
24 Yadav S & Rewari S, *Proc. 2022 IEEE Int. Conf. Electron Devices Soc. Kolkata Chapter (EDKCON), (2022)* 23.
25 Das A, Rewari S, Kanaujia B K, Deswal S S & Gupta R S, *Microelectronics J*, 138 (2023) 105832.
26 Yadav S & Rewari S, *Microelectronics J*, 142 (2023)106011.
27 Yadav S, Das A & Rewari S, *ECS J Solid State Sci Technol*, 13 (2024) 047001.
28 Kumari M, Singh N K, Sahoo M & Rahaman H, *Appl Phys A Mater Sci Process*, 127 (2021)1.
29 Adel SS, C.Smith K. 2020. *Microelectronics Circuit*. Oxford. 1296 pp. 8th ed.
30 Beckers A, Jazaeri F & Enz C, *IEEE Electron Device Lett*, 41 (2019) 276.
31 Kaul A, Rewari S & Nand D, *Microsyst Technol*, 30 (2024) 599.
32 Kaul A, Yadav S, Rewari S & Nand D, *IEEE Trans Dielectr Electr Insul*, (2025).
33 Bhat A M, Varghese A, Shafi N & Periasamy C, *IETE J Res*, 69 (2021) 1419.
34 Wadhera T, Kakkar D, Wadhwa G & Raj B, 2019. *J. Electron. Mater*, 48 (2019) 7635.
35 Lee H, Park J D & Shin C, *IEEE Trans Electron Devices*, 63 (2016) 1827.
36 Dey A W, *et al*., *IEEE Electron Device Lett*, 33 (2012) 791.
37 Swati, Kaur J & Singh A K, *Appl Phys A Mater Sci Process*, 129 (2023) 365.
38 Das A, Rewari S, Kanaujia B K, Deswal S S & Gupta R S, *J Comput Electron*, 22 (2023) 742.
39 Goel A, Rewari S, Verma S & Gupta R S, *Proc. Int. Conf. 2018 IEEE Electron Device Kolkata Conf EDKCON*, (2018) 337.
40 Das A, Rewari S, Kanaujia B K, Deswal S S & Gupta R S, *Int J Numer Model*, 36 (2023) e3106.
41 Darwin S & Arun Samuel T S, *Silicon*, 12 (2020) 393.
42 Kaul A, Rewari S & Nand D, *ECS J Solid State Sci Technol*, 13 (2024) 103010.
43 Barik M A, Deka R & Dutta J C, *IEEE Sens J*, 16 (2016) 280.
44 Singh K N & Dutta P K, *Superlattices Microstruct*, 155 (2021) 106897.
45 Das R, Chanda M & Sarkar C K, *IEEE Trans Electron Devices*, 65 (2018) 5487.
46 Kumar M, Haldar S, Gupta M & Gupta R S, *Superlattices Microstruct*, 111 (2017) 10.
47 Madan J & Chaujar R, *IEEE Trans Device Mater Reliab*, 16 (2016) 227.
48 Nuzaihan MMN, et al., *Biosens Bioelectron*, 83 (2016) 106.
49 Kaul A, Yadav S, Rewari S &Nand D, *Micro and Nanostructures*, 191 (2024) 207851.
50 Das A, Rewari S, Kanaujia B K, Deswal S S & Gupta R S, *Phys Scr*, 98 (2023) 115013.
51 Das A, Rewari S, Kanaujia B K, Deswal S S & Gupta R S, *Phys Scr*, 98 (2023) 074005.
52 Pratap Y, Kumar M, Kabra S, Haldar S, Gupta R S & Gupta M, *J Comput Electron*, 17 (2018) 288.
53 Parihar M S, Ghosh D, Armstrong GA, Yu R, Razavi P & Kranti A, *Appl Phys Lett*, 101 (2012) 093507.
54 Parihar M S &Kranti A, *Nanotechnology*, 26 (2015) 145201.
55 Kumar A, Tripathi M M & Chaujar R, *J Comput Electron*, 17 (2018) 1807.
56 Maiti S, De A & Sarkar S K, *Silicon*, 14 (2022) 6921.
57 Getnet M & Chaujar R, *J Electron Mater*, 51 (2022) 2236.

Int J Syst Assur Eng Manag
https://doi.org/10.1007/s13198-025-02778-5

ORIGINAL ARTICLE

P4.0B-FAHP: prioritizing industry 4.0 barriers using fuzzy AHP—case of an Indian construction company

Ankur Tayal[1] · **Saurabh Agrawal**[1] · **Rajan Yadav**[1]

Received: 19 June 2024 / Revised: 20 September 2024 / Accepted: 22 March 2025
© The Author(s) under exclusive licence to The Society for Reliability Engineering, Quality and Operations Management (SREQOM), India and The Division of Operation and Maintenance, Lulea University of Technology, Sweden 2025

Abstract Industry 4.0 is revolutionizing the construction industry by integrating advanced technologies to enhance efficiency and productivity. This article offers insights into implementing Industry 4.0 in the construction industry of developing nations like India. However, its implementation faces significant challenges due to diverse barriers. Hence, this research explores and prioritizes these barriers for the effective implementation of Industry 4.0 within the construction industry. This study delved into 20 barriers across five primary criteria, utilizing an extensive literature review and expert input. A Fuzzy Analytical Hierarchy Process is applied to prioritize the barriers. The utilized Fuzzy framework significantly deals with unreliability and ambiguity. A case study of the Indian Construction Company illustrates the suggested model. A sensitivity study was conducted to verify the model's resilience. The findings reveal that high initial investment, lack of funding, data management, and security are notable barriers to the implementation of Industry 4.0 within the construction industry of developing nations. The insights of this study serve as valuable resources for managers and policymakers in strategizing effective implementation of Industry 4.0.

Keywords Barriers · Industry 4.0 · Fuzzy analytical hierarchy process (FAHP) · Construction industry

1 Introduction

The construction industry is undergoing a profound transformation with the advent of Industry 4.0, integrating digital technologies into physical processes and reshaping traditional practices. Newman et al. (2021) argued that I4.0 technologies such as "Building Information Modeling (BIM), Internet of Things (IoT), and robotics," etc., are modernizing construction projects, increasing productivity, and improving safety standards. BIM, specifically, has emerged as a cornerstone of digital transformation in construction, facilitating collaborations among stakeholders for effective decision-making (Liu et al. 2021). According to Omrany et al. (2023), the application of IoT sensors and devices on construction sites allows for live tracking of equipment, environmental conditions, and materials, leading to proactive maintenance and optimal resource utilization. Moreover, robots and automated technologies have simplified construction operations by automating repetitive jobs and boosting precision (Javaid et al. 2021).

Despite the potential benefits, implementation of I4.0 in the construction industry faces distinct barriers such as implementation cost, poor standards, legal issues, cyber security, and workforce skills gaps "(Singh et al. 2023; Yap et al. 2022; Oesterreich and Teuteberg 2016)." Analysis of these barriers is crucial for effectively implementing I4.0 in the construction industry.

This study examined 20 specific barriers categorized into five major groups, incorporating insights from literature reviews and expert opinions. "The aim was to offer a comprehensive understanding of these barriers, primarily focusing on the Indian construction industry." "Fuzzy Analytical Hierarchy Process (FAHP) is utilized to assign weights to the identified barriers and prioritize them. Moreover, the proposed model is examined for sensitivity

✉ Ankur Tayal
ankurtayal_2K21phddsm12@dtu.ac.in

[1] Delhi School of Management, Delhi Technological University (DTU), New Delhi 110042, India

Published online: 10 April 2025

analysis to establish its consistency. The relative ranking can assist decision-makers, engineers, stakeholders, and the Government in strategizing the implementation of I4.0 in the Indian construction industry."

This study is classified into further sections, such as a review of I4.0 barriers, a focus on problem definition, outlining the two-phase model and applying the proposed model, presenting results and discussions, followed by drawing conclusions, discussing managerial implications, and addressing study limitations.

2 Examining existing literature

Barriers to Implementing I4.0 in Developing Countries.

In recent years, academicians and researchers have focused on integrating I4.0 practices in the construction industry (Ribeiro et al. 2024; Mansour et al. 2023; Sajjad et al. 2023; Newman et al. 2021; You and Feng 2020). However, implementing I4.0 remains in its infancy within developing nations, owing to numerous barriers (Demirkesen and Tezel 2022). The literature seeks to address various barriers that hinder the adoption of I4.0 in the construction industry of developing countries (Tayal et al. 2023; Sajjad et al. 2023; Singh and Misra 2021; Maskuriy et al. 2019; Osunsanmi et al. 2018). "This article categorizes barriers into five primary classifications, each with its own sub-criteria. This classification is drawn through a literature review and multiple rounds of discussions with experts. Table 1 illustrates the final investigated barriers of I4.0 in the construction Industry."

2.1 Financial barriers

Most of the literature outlines that in developing nations, high initial investments for purchasing and installing machinery, software, etc., act as a prime barrier in implementing I4.0 (Demirkesen and Tezel 2022; Singh et al. 2023). Organizations encounter financial challenges due to a lack of funding from financial institutions and reluctance from investors to contribute to risky projects (Olsson et al. 2021; Newman et al. 2021). According to Watts et al. (2023), the construction industry faces uncertainty in investment returns due to project complexity, market volatility, and technological obsolescence. Moreover, cost–benefit analysis for I4.0 investments can be challenging due to project complexity and long-term benefits (Kumar et al. 2021a, b; Hong et al. 2018). Therefore, factors such as intangible benefits, unpredictable project outcomes, and dynamic technical landscapes complicate evaluating potential returns, making decision-making difficult.

2.2 Technological barriers

Past research (Demirkesen and Tezel 2022; Aripin et al. 2019) "illustrates that the complexity of technologies like

Table 1 Investigated barriers of I4.0 in construction industry

Major barriers/criteria	Barriers code	Sub-criterias	References
Financial barriers	FI1	High initial investment	Demirkesen and Tezel (2022), Singh et al. (2023), Watts et al. (2023), Olsson et al. (2021), Newman et al. (2021), Kumar et al. (2021a, b), Hong et al. (2018)
	FI2	Uncertain return on investment	
	FI3	Lack of funding	
	FI4	Cost–benefit analysis challenges	
Technological barriers	TE1	Complexity of technologies	Tayal et al. (2024), Aripin et al. (2019), Okonkwo et al. (2023), Lau et al. (2021), Siriwardhana and Moehler (2023)
	TE2	Poor standards and protocols	
	TE3	Data management and security	
	TE4	Skill gaps and training needs	
Regulatory barriers	RE1	Privacy regulations	Almatari et al. (2023), Yap et al. (2022), Spaan and Abraham (2023), Umeokafor and Teizer (2023), Hanuar et al. (2023)
	RE2	Safety and liability regulations	
	RE3	Ethical and social implications	
	RE4	Poor tax rebates	
Organizational barriers	OR1	Lack of leadership and vision	Olsson et al. (2021), Zulu et al. (2023), Govindan and Arampatzis (2023), Demirkesen and Tezel (2022), Orzes et al. (2018), Singh et al. (2023)
	OR2	Resistant to change	
	OR3	Short term focus and risk aversion	
	OR4	Lack of performance measures	
Market barriers	MR1	Market fragmentation	Nagy et al. (2021), Demirkesen and Tezel (2022), Balkhy et al. (2021), Martinez and Pfister (2023)
	MR2	Limited customer awareness	
	MR3	Vendor lock-in	
	MR4	Market volatility	

Building Information Modeling (BIM) is a significant barrier to I4.0 implementation in the construction industry." Okonkwo et al. (2023) also emphasize the challenges of poor standardization and protocols in I4.0 technologies, leading to compatibility issues and escalated construction budgets. Without standardized interoperability frameworks, businesses may incur significant costs when customizing and connecting disparate systems. Lau et al. (2021) argued that complexities in storing and securing vast data with data integrity, confidentiality, and privacy compliances are further critical barriers. According to Siriwardhana and Moehler (2023), inadequate skill gaps and training needs among construction workplaces hinders the integration of I4.0 in the construction industry. A lack of digital literacy and technical competencies might limit the proper installation and use of innovative technology.

2.3 Regulatory barriers

"Strict privacy regulations, such as General Data Protection Regulation (GDPR), mandate data handling and permission rules, which may complicate data collecting and sharing in construction projects"(Almatari et al. 2023; Yap et al. 2022). Advanced automation systems and robotics raise new safety concerns and potential liabilities, needing strict adherence to industry norms and regulations to mitigate risks successfully (Spaan and Abraham 2023). Technological developments may result in changes in job roles and skill needs, creating concerns about labor displacement and fair access to opportunities, which can create ethical and social implications as substantial barriers to adoption (Umeokafor and Teizer 2023). According to Hanuar et al. (2023), limited or poor tax rebates and incentives for investing in I4.0 technologies in construction can serve as a technological barrier. Construction firms may be hesitant to invest in technological upgrades without adequate financial incentives or tax rebates, affecting the industry's pace of adoption and innovation.

2.4 Organizational barriers

Literature reveals that a lack of visionary leadership within construction organizations can impede the implementation of I4.0 technology (Olsson et al. 2021; Zulu et al. 2023). Without strong leadership and a clear vision for digital transformation, organizations may struggle to set strategic goals, manage resources effectively, and rally support for technical projects. Organizational resistance such as cultural norms, fear of job displacement, or reluctance to deviate from traditional practices are significant barriers to implementing I4.0 technology in construction (Govindan and Arampatzis (2023). Studies (Demirkesen and Tezel 2022; Orzes et al. 2018) illustrate that Short-term focus and risk aversion can hinder the implementation of I4.0 in construction. Singh et al. (2023) argued that the lack of relevant performance measures designed for I4.0 initiatives affects organizational efforts to monitor the effectiveness and impact of technological investments. Organizations may struggle to justify investments, manage resources appropriately, and track progress toward digital transformation objectives without defined metrics for evaluating outcomes and demonstrating ROI.

2.5 Market barriers

Nagy et al. (2021) "indicate that the construction industry is fragmented, with various small and medium-sized firms (SMEs) functioning independently. This fragmentation affects the adoption of I4.0 technologies since businesses may lack the scale and resources to invest in and implement advanced solutions successfully." Market constraints stem from a lack of customer awareness and comprehension of the potential benefits of I4.0 technologies in construction, resulting in poor adoption rates of I4.0 in the construction industry (Demirkesen and Tezel 2022). According to Balkhy et al. (2021), Vendor lock-in, where construction firms become dependent on specific technology providers or platforms, can limit flexibility, raise prices, and reduce interoperability with alternative solutions, stifling market competition and innovation. The construction industry is susceptible to market volatility due to factors such as economic fluctuations, regulatory changes, and geopolitical uncertainties that interrupt investment cycles, cause project delays, and undermine confidence in long-term technology investments, posing barriers to the widespread adoption of I4.0 technologies (Martinez and Pfister 2023).

3 Research objectives and methodology

"Despite the global push towards digital transformation, the construction industry in developing economies lags behind due to several challenges, including financial constraints, technological complexity, regulatory issues, organizational barriers, and market-related factors. Based on the literature review, the following research gaps are identified:"

i. *Lack of Focus on Developing Countries* Much of the existing research on Industry 4.0 (I4.0) adoption is centered on developed nations, where technological readiness, infrastructure, and regulatory frameworks are more robust (Ribeiro et al. 2024; Newman et al. 2021). However, the specific challenges faced by developing countries, such as limited funding, skills shortages, and inadequate regulatory support, have not been extensively explored. Developing countries

face additional financial barriers, such as high initial investment and lack of access to capital (Demirkesen and Tezel 2022; Kumar et al. 2021a, b). This research addresses the gap by focusing on the Indian construction industry and exploring the barriers to I4.0 implementation within a developing economy context.

ii. *Comprehensive Analysis of Barriers* Previous studies have identified various barriers to I4.0, such as technological complexity, workforce skill gaps, and regulatory challenges (Osunsanmi et al. 2018; Singh et al. 2023). However, most studies focus on individual barriers without providing a comprehensive, multi-dimensional analysis. The existing literature often needs a holistic perspective on the interconnectedness of these barriers. By categorizing barriers into financial, technological, regulatory, organizational, and market-related groups and using Fuzzy Analytical Hierarchy Process (FAHP), this study offers a more detailed, prioritized analysis of these barriers in the construction industry of developing nations (Mansour et al. 2023; Bajpai and Misra 2022).

iii. *Lack of Prioritization Framework* Despite the identification of barriers to I4.0, there needs to be a universally accepted framework for their prioritization, which limits practical decision-making for industry stakeholders (Liu et al. 2021). The lack of structured methods to evaluate the significance of different barriers leaves decision-makers without clear guidance on allocating resources efficiently. This study fills that gap by using Multi-Criteria Decision-Making (MCDM) techniques, precisely the fuzzy logic approach, to create a data-driven model for ranking barriers to I4.0 implementation (Pathak et al. 2022). This structured prioritization provides actionable insights for stakeholders in the construction industry.

iv. *Scarcity of Case Studies in the Construction industry* Although there is a growing body of literature on I4.0 adoption in sectors such as manufacturing and logistics, construction still needs to be explored (Shafei et al. 2024). The unique challenges the construction industry faces, such as project-specific customization, complex supply chains, and the need for real-time data integration, are not adequately addressed in the current research (Maskuriy et al. 2019). "By applying its analysis to a real-world case study of an Indian construction company, this research fills a significant gap in the literature. It validates the effectiveness of the proposed methodology within the construction industry."

By addressing these gaps, "this study contributes to a deeper understanding of how the construction industry in developing nations can overcome barriers to the adoption of I4.0 technologies, offering practical recommendations for industry leaders, policymakers, and researchers."

"The goal of this study is to examine the barriers experienced by the Indian construction industry in the integration of I4.0 technologies. This study has two primary goals: first, to identify the barriers to I4.0 adoption in the Indian construction industry, and second, to prioritize these barriers using the "FAHP" technique."

This study uses a two-phase hybrid strategy to explore and prioritize barriers. "Initially, barriers to I4.0 implementation in the Indian construction industry are identified via a literature analysis and expert consultation. The barriers and sub-barriers are then analyzed and prioritized using the FAHP technique." Fig. 1 illustrates the proposed/planned research design.

3.1 Fuzzy AHP (FAHP)

"The Analytic Hierarchy Process (FAHP) is a popular numerical method in the field of Multi Criteria Decision Making (MCDM)," developed by Saaty (1991). AHP is a powerful decision-making method that handles uncertainty and ambiguity in decision-making processes. "Its importance in I4.0, notably in the construction industry, stems from the complex and diverse nature of construction projects. Table 2 depicts contemporary uses of Fuzzy AHP".

Chang (1996), created an extent analysis method that applies to each criterion g_i.

$$M_{gi}^{1}, M_{gi}^{2}, \ldots\ldots\ldots.., M_{gi}^{m}, i = 1,2,3,4, \ldots\ldots.., n$$

"The $M_{gi}^{j}(j = 1,2,3,4, \ldots\ldots\ldots.., m)$ are Triangular Fuzzy Numbers (TFN), as depicted in Table 3. Triangular Fuzzy Numbers (TFN) are used to represent uncertainty in decision-making processes. A TFN is expressed as a triplet (l,m,u)(l, m, u)(l,m,u), where: l(lower bound) represents the smallest possible value, m(middle value) represents the most likely value, and u(upper bound) represents the largest possible value."

This fuzzy number is called "triangular" because its membership function is shaped like a triangle. TFNs are widely used in the Fuzzy Analytical Hierarchy Process (FAHP) to handle imprecision in expert judgments, allowing for a flexible and more realistic representation of data in scenarios where exact numbers are hard to determine. In this study, TFNs facilitate the comparison of barriers to Industry 4.0 by quantifying subjective opinions with these ranges.

Chang's analysis outlines the following steps:

Step 1 "The value of fuzzy synthetic (S_i) for the ith criterion is given as:"

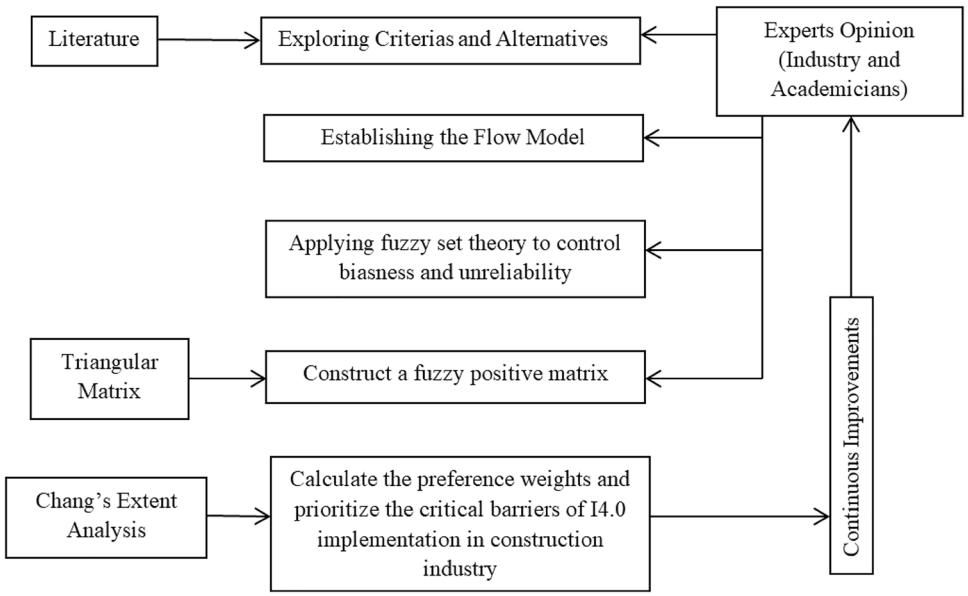


Fig. 1 Proposed research design

Table 2 Contemporary uses of Fuzzy AHP

Author(s)	Research area	Country
Shafei et al. (2024)	Rank construction 4.0 technologies	Malaysia
Mansour et al. (2023)	Analyzed I4.0 critical factors to assess readiness	Malaysia
Hall et al. (2023)	Identify and prioritize barriers of Building Information Modeling (BIM)	New Zealand
Musarat et al. (2022)	Analyzed I4.0 technologies for improving health and safety	Malaysia
Bajpai and Misra (2022)	Identify and prioritize barriers of digitalization	India
Lee and Park (2022)	Examined the competitiveness from an industrial standpoint to rejuvenate smart technologies	Korea
Pathak et al. (2022)	Identify and prioritize barriers to the development of renewable energy technologies	India
Bademosi and Issa (2021)	Identify and prioritize positive and negative factors in adoption of robotics and automation	United States
Wu et al. (2021)	Analyze barriers to implementation of BIM for industrialized buildings	China
Shaawat et al. (2018)	Analyze challenges in sustainable construction	Saudi Arabia

Table 3 Triangular fuzzy number (TFN) scale. *Source* Prakash and Barua (2015)

Lingual variables	Fuzzy triangular numbers
Equal	(1, 1, 1)
Very low	(1, 2, 3)
Low	(2, 3, 4)
Moderate	(3, 4, 5)
High	(4, 5, 6)
Very strong	(5, 6, 7)
Extreme	(7, 8, 9)

$$S_i = \sum_{j=1}^{m} M_{gi}^{j} \times \left[\sum_{i=1}^{n}\sum_{j=1}^{m} M_{gi}^{j}\right]^{-1} \tag{1}$$

$$\sum_{j=1}^{m} M_{gi}^{j} = \left(\sum_{j=1}^{m} l_{ij}, \sum_{j=1}^{m} m_{ij}, \sum_{j=1}^{m} u_{ij}\right)$$

$$\left[\sum_{i=1}^{n}\sum_{j=1}^{m} M_{gi}^{j}\right]^{-1} = \left(\frac{1}{\sum_{i=1}^{n}\sum_{j=1}^{m} u_{ij}}, \frac{1}{\sum_{i=1}^{n}\sum_{j=1}^{m} m_{ij}}, \frac{1}{\sum_{i=1}^{n}\sum_{j=1}^{m} l_{ij}}\right)$$

where 'm' signifies the most favorable value, 'u' denotes the maximum value and 'l' represents the minimum value.

Step 2 The level of probability of $S_2 = (l_2, m_2, u_2) \geq S_1 = (l_1, m_1, u_1)$ is defined as

$$V(S_2 \geq S_1) = {}_{x \geq y}{}^{sup}\left[\min(\mu_{s_1}(x), \mu_{s_1}(y))\right]$$

$$V(S_2 \geq S_1) = \begin{cases} 1, & if\ m_2 \geq m_1 \\ V(S_2 \geq S_1) = 0, & if\ l_1 \geq \mu_2 \\ \frac{l_1 - u_2}{(m_2 - u_2) - (m_1 - l_1)} = \mu_d \end{cases} \quad (2)$$

where μ_d is the maximum convergence of fuzzy numbers $\mu_{s_1} and \mu_{s_2}$. (Fig. 2).

Step 3: "The probability of a convex fuzzy number being bigger than k convex fuzzy numbers $S_i (i = 1,2, \ldots .., k)$ can be defined as:"

$$\begin{aligned} &V(S \geq S_1, S_2, \ldots \ldots \ldots .., S_k) \\ &V[(S \geq S_1), (S \geq S_2), \ldots \ldots \ldots \ldots ., (S \geq S_k)] \\ &= minV[(S \geq S_i), \\ &i = 1,2, \ldots, kassumingthatd\prime(A_i) = minV(S_i \geq S_k) \end{aligned} \quad (3)$$

$Fork = 1,2, \ldots .., n, (k \neq i)$, the weight vectors are illustrated by

$$W\prime = (d\prime(A_1), d\prime(A_2), \ldots \ldots .., d\prime(A_k))^T \quad (4)$$

Step 4: Following normalisation, the normalised weight vectors are depicted by:

$$W\prime = (W_1\prime, W_2\prime, \ldots \ldots \ldots, W_n\prime)^T \quad (5)$$

3.2 Computational complexity of the proposed methodology

The manuscript proposes the use of a Fuzzy Analytical Hierarchy Process (FAHP) for prioritizing barriers to Industry 4.0 (I4.0) implementation in the Indian construction industry. The computational complexity of FAHP arises from several factors, mainly the number of criteria and pairwise comparisons required, as well as the fuzzification and defuzzification processes inherent in the fuzzy logic framework. Specifically, FAHP involves multiple steps:

i. *Pairwise Comparisons* FAHP requires pairwise comparisons of all criteria and sub-criteria. If there are n criteria, the number of comparisons needed is $\frac{n(n-1)}{2}$. This creates a quadratic relationship in terms of computational complexity.
ii. *Fuzzification and Defuzzification* Each comparison is represented by triangular fuzzy numbers, which need to be synthesized and converted into crisp values. These operations are computationally expensive, typically requiring operations such as fuzzy multiplication, addition, and defuzzification using centroid methods. The complexity of these operations depends on the number of fuzzy numbers and their degree of precision, which adds further computational overhead.
iii. *Normalization and Weight Calculation* After defuzzification, the weights are normalized, and a final priority vector is computed. While the normalization itself is linear, synthesizing the fuzzy matrices and computing the synthetic extent values adds to the overall complexity.

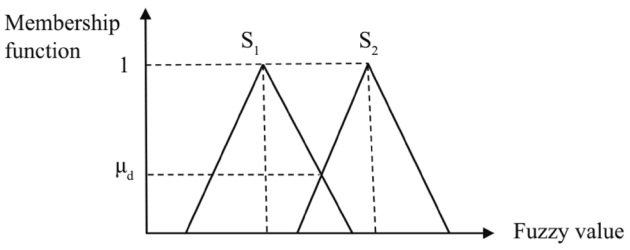


Fig. 2 Convergence of fuzzy numbers

Thus, the overall computational complexity of the FAHP method can be approximated as **O(n^2)** for pairwise comparisons, with additional overhead for fuzzification and defuzzification steps, potentially reaching **O(n^3)** depending on the precision of the fuzzy numbers used.

3.3 Comparison to existing methodologies

i. *Traditional AHP* Traditional Analytical Hierarchy Process (AHP) is widely used in multi-criteria decision-making (MCDM). The complexity of AHP primarily arises from the pairwise comparison matrix, where each criterion is compared with others, leading to a quadratic complexity **O(n^2)** (Saaty 1991). However, AHP lacks the ability to handle uncertainty and imprecision in judgments, making it less suitable for scenarios involving high ambiguity, like Industry 4.0 implementation in construction. FAHP overcomes this limitation by integrating fuzzy logic, allowing decision-makers to express preferences more flexibly using linguistic variables.
ii. "*Fuzzy TOPSIS* Fuzzy Technique for Order Preference by Similarity to Ideal Solution (Fuzzy TOPSIS) is another common method in MCDM problems." Its computational complexity is similar to FAHP, with the complexity for ranking alternatives being quadratic. However, Fuzzy TOPSIS does not inherently incorporate hierarchical structuring of criteria, which FAHP offers. Fuzzy TOPSIS is often preferred when the decision problem focuses on ranking a set of alternatives rather than dealing with hierarchical structures (Prakash and Barua 2015).

iii. "*DEMATEL and Fuzzy DEMATEL* Decision Making Trial and Evaluation Laboratory (DEMATEL) and its fuzzy variant are also used for barrier analysis in complex systems." DEMATEL captures the cause-effect relationships among criteria, adding another layer of computational complexity compared to AHP or FAHP. DEMATEL's matrix operations and handling of causal relationships typically make it more computationally intensive, often reaching **O(n^3)** for larger matrices (Wu et al. 2021). However, FAHP's structured hierarchy and direct pairwise comparison make it simpler and more intuitive for prioritization tasks.

iv. *Hybrid Methods* Other hybrid methods like **PCA-Fuzzy AHP** (Kumar et al. 2021a, b) and **SWARA-COPRAS** integrate multiple techniques to overcome limitations of single MCDM methods. While these methods can achieve more accurate results, they often come at the cost of increased computational complexity. Hybrid methods involving fuzzy logic typically show **O(n^3)** complexity, making them more computationally demanding than pure FAHP approaches.

The FAHP method proposed in the manuscript has a computational complexity of **O(n^2)** to **O(n^3)**, primarily depending on the number of criteria and the fuzzification process. Compared to traditional AHP, FAHP better handles uncertainty but incurs higher computational costs. Fuzzy TOPSIS and DEMATEL also offer alternatives, with DEMATEL being more computationally intensive due to its causal relationship analysis. The choice between these methodologies depends on the decision context, but FAHP provides a good balance between computational efficiency and the ability to model uncertainty, particularly for the barrier prioritization problem discussed in the manuscript.

4 Implementation of the proposed model

The proposed model is implemented in a Construction Company (identity withheld for confidentiality reasons) situated in northern India. The company is considered for the case study because it is a prominent player in the Indian construction industry. The company is renowned for its extensive expertise in significant construction projects using modern technologies.

4.1 Data collection

A survey was developed with insights from the literature study and disseminated to experts and decision-makers. Their expertise was used to explore common barriers in the construction industry to the acceptance of I4.0 technologies. The study considers decision-makers from both industry and academia to mitigate potential biases toward their respective organizations. "Eight professionals with over 15 years of experience in the construction industry were chosen to refine identified barriers and construct a pairwise decision matrix." "Twenty barriers were eventually classified into five major categories based on both literature research findings and expert opinions, as detailed in Table 1. A visual illustration of the Fuzzy AHP framework is shown in Fig. 3, which showcases the use of the Fuzzy AHP method to prioritize barriers to Industry 4.0 implementation." The process involves several key steps:

i. *Problem Definition* Barriers to Industry 4.0 are identified through literature reviews and expert input, categorized into financial, technological, organizational, market, and regulatory barriers.

ii. *Hierarchy Structuring* The identified barriers are arranged into a hierarchical structure, with the main objective at the top and sub-barriers underneath.

iii. *Pairwise Comparisons* Experts perform pairwise comparisons between barriers using Triangular Fuzzy Numbers (TFNs) to express uncertainty. The fuzzy scale captures the relative importance of each barrier.

iv. *Fuzzy Synthesis* Fuzzy synthetic values are computed to aggregate expert judgments, reflecting the overall importance of each barrier.

v. *Defuzzification* The fuzzy values are converted into crisp values to simplify the comparison of barriers.

vi. *Ranking and Sensitivity Analysis* Barriers are ranked based on their defuzzified scores, and sensitivity analysis ensures that the results are stable.

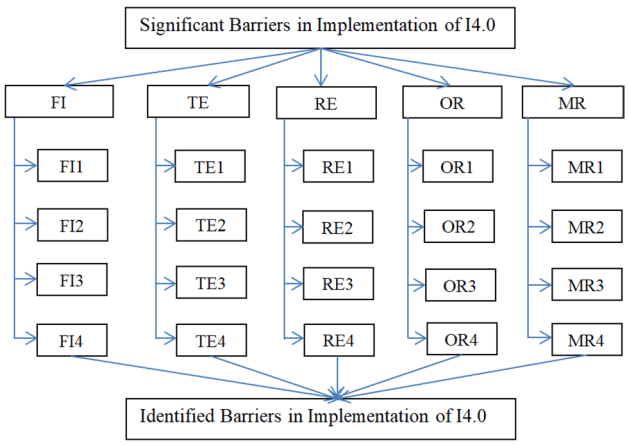


Fig. 3 Fuzzy analytic hierarchy framework for identifying barriers to I4.0 implementation

Fuzzy logic handles uncertainty and ambiguity in expert judgments, making the decision-making process more flexible and realistic. It enables the quantification of subjective opinions, which is crucial in complex scenarios like Industry 4.0 implementation, where precise data may not always be available. This approach enhances accuracy and helps decision-makers prioritize actions based on reliable, nuanced evaluations.

4.2 Computation of "fuzzy synthetic weight"

"Decision-makers conducted pairwise analysis of major criteria using a TFN scale (Table 3)." Table 4 displays the pairwise assessments of major barriers based on lingual statements and expressions.

"Equation (1), calculates fuzzy synthetic weights for five major barriers, as shown below:"

S(FI) = (7.00,11.00,15.00) x [0.020,0.031,0.039] = (0.140,0.341,0.585).

S(OR) = (4.91,6.33,8.50) x [0.020,0.031,0.039] = (0.098,0.196,0.331).

S(TE) = (6.00,8.50,11.33) x [0.020,0.031,0.039] = (0.120,0.263,0.441).

S(MR) = (2.11,2.58,3.83) x [0.020,0.031,0.039] = (0.042,0.079,0.149).

S(RE) = (4.50,6.66,9.00) x [0.020,0.031,0.039] = (0.090,0.206,0.351).

Equations (2) and (3) are used to compute the minimum possible degree:

m(FI) = min $V(S_1 \geq S_k)$ = min(1,1,1,1), whereas values for further criterias are m(OR) = 0.568, m(TE) = 0.794, m(MR) = 0.033, m(RE) = 0.609.

"The final weight vector is determined by": $W_v = (1.000, 0.568, 0.794, 0.033, 0.609)^T$

"The final weight values are calculated after normalization as:"

W = (0.655, 0.372, 0.520, 0.021, 0.399).

"Table 5 illustrates the final relative weights and ranking of the major barriers or criteria. The ranking of the main barriers is FI > TE > RE > OR > MR." Financial barriers are ranked as the most critical category in the manuscript, surpassing technological, regulatory, organizational, and market-related barriers. These barriers include high initial investment costs, lack of funding, and uncertainty in cost–benefit analysis, which are critical challenges for implementing Industry 4.0 in the construction industry.

Table 5 Relative weight and ranks of major barriers

Criteria	Relative weights	Rank
FI	0.655	1
OR	0.372	4
TE	0.520	2
MR	0.021	5
RE	0.399	3

The dominance of financial barriers implies that decision-makers prioritize addressing funding and investment issues before tackling technological upgrades or regulatory compliance. This ranking suggests that without resolving financial constraints, advancements in other areas, such as technology adoption and workforce development, are likely to be stalled, impacting the overall progress of Industry 4.0 integration in the construction industry.

"Likewise, the pairwise assessment of sub-barriers was conducted as illustrated in Tables 6, 7, 8, 9, and 10." These tables evaluate how each sub-barrier contributes to the overall category by assigning relative importance using a Triangular Fuzzy Number (TFN) scale. Each sub-barrier is compared with others based on expert opinions, assessing which barrier poses a more significant challenge to Industry 4.0 adoption.

"Ultimately, the global weights of individual sub-barriers were computed by multiplying the relative weights of the major barriers by those of the sub-barriers." Table 11 shows the final ranking of I4.0 barriers based on the global weights assigned to each barrier.

5 Results and discussions

The fuzzy AHP approach provides a rational answer for comparing two barriers, assisting decision-makers in prioritizing them. This article uses AHP approach to investigate the barriers to the application of I4.0 in the Indian construction industry. "It encourages organizations to understand the barriers of implementing the I4.0 in their construction processes. It illustrates the ranking of barriers

Table 4 Pairwise assessment of major barriers

	FI	OR	TE	MR	RE
FI	(1.00,1.00,1.00)	(1.00,2.00,3.00)	(1.00,2.00,3.00)	(2.00,3.00,4.00)	(2.00,3.00,4.00)
OR	(0.33,0.50,1.00)	(1.00,1.00,1.00)	(0.33,0.50,1.00)	(3.00,4.00,5.00)	(0.25,0.33,0.50)
TE	(1.00,0.50,0.33)	(1.00,2.00,3.00)	(1.00,1.00,1.00)	(1.00,2.00,3.00)	(2.00,3.00,4.00)
MR	(0.25,0.33,0.50)	(0.20,0.25,0.33)	(0.33,0.50,1.00)	(1.00,1.00,1.00)	(0.33,0.50,1.00)
RE	(0.25,0.33,0.50)	(2.00,3.00,4.00)	(0.25,0.33,0.50)	(1.00,2.00,3.00)	(1.00,1.00,1.00)

Table 6 Pairwise assessment of financial sub-barrier (FI)

	FI1	FI2	FI3	FI4
FI1	(1.00,1.00,1.00)	(2.00,3.00,4.00)	(0.33,0.50,1.00)	(3.00,4.00,5.00)
FI2	(0.25,0.33,0.50)	(1.00,1.00,1.00)	(0.33,0.50,1.00)	(1.00,2.00,3.00)
FI3	(1.00,2.00,3.00)	(1.00,2.00,3.00)	(1.00,1.00,1.00)	(0.25,0.33,0.50)
FI4	(0.2,0.25,0.33)	(0.33,0.50,1.00)	(2.00,3.00,4.00)	(1.00,1.00,1.00)

Table 7 Pairwise assessment of organizational sub-barrier (OR)

	OR1	OR2	OR3	OR4
OR1	(1.00,1.00,1.00)	(3.00,4.00,5.00)	(0.33,0.50,1.00)	(2.00,3.00,4.00)
OR2	(0.20,0.25,0.33)	(1.00,1.00,1.00)	(2.00,3.00,4.00)	(0.25,0.33,0.50)
OR3	(1.00,2.00,3.00)	(0.25,0.33,0.50)	(1.00,1.00,1.00)	(1.00,2.00,3.00)
OR4	(0.25,0.33,0.50)	(2.00,3.00,4.00)	(0.33,0.50,1.00)	(1.00,1.00,1.00)

Table 8 Pairwise assessment of technological sub-barrier (TE)

	TE1	TE2	TE3	TE4
TE1	(1.00,1.00,1.00)	(2.00,3.00,4.00)	(1.00,2.00,3.00)	(0.20,0.25,0.33)
TE2	(0.25,0.33,0.50)	(1.00,1.00,1.00)	(0.33,0.50,1.00)	(3.00,4.00,5.00)
TE3	(0.33,0.50,1.00)	(1.00,2.00,3.00)	(1.00,1.00,1.00)	(2.00,3.00,4.00)
TE4	(3.00,4.00,5.00)	(0.20,0.25,0.33)	(0.25,0.33,0.50)	(1.00,1.00,1.00)

Table 9 Pairwise assessment of market sub-barrier (MR)

	MR1	MR2	MR3	MR4
MR1	(1.00,1.00,1.00)	(0.25,0.33,0.50)	(0.20,0.25,0.33)	(3.00,4.00,5.00)
MR2	(2.00,3.00,4.00)	(1.00,1.00,1.00)	(2.00,3.00,4.00)	(0.33,0.50,1.00)
MR3	(3.00,4.00,5.00)	(0.25,0.33,0.50)	(1.00,1.00,1.00)	(1.00,2.00,3.00)
MR4	(0.20,0.25,0.33)	(1.00,2.00,3.00)	(0.33,0.50,1.00)	(1.00,1.00,1.00)

Table 10 Pairwise assessment of regulatory sub-barrier (RE)

	RE1	RE2	RE3	RE4
RE1	(1.00,1.00,1.00)	(0.20,0.25,0.33)	(3.00,4.00,5.00)	(0.25,0.33,0.50)
RE2	(3.00,4.00,5.00)	(1.00,1.00,1.00)	(0.33,0.50,1.00)	(0.20,0.25,0.33)
RE3	(0.20,0.25,0.33)	(1.00,2.00,3.00)	(1.00,1.00,1.00)	(2.00,3.00,4.00)
RE4	(2.00,3.00,4.00)	(3.00,4.00,5.00)	(0.25,0.33,0.50)	(1.00,1.00,1.00)

based on their prioritized weightage values, with financial, technological, regulatory, organizational, and market-related barriers grouped in descending order, as shown in Table 5." Financial barriers are the most significant barriers to the implementation of I4.0 in the construction industry, and sub-barriers are prioritized as FI1 > FI3 > FI4 > FI2. Initial Costs incurred in purchasing modern machinery, improving infrastructure, and obtaining digital tools required for automation and digitization act as significant determinants (Demirkesen and Tezel (2022). Limited budget allocations and lack of funding such as loans, venture capital, or investment are major constraints. This limited access to external funding sources exacerbates the issue of obtaining the resources required for I4.0 implementation (Watts et al. 2023). Moreover, the complexity of conducting comprehensive cost–benefit analyses due to data availability and quality poses additional barriers to I4.0 implementation (Newman et al. 2021; Taher 2021). Decision-makers confront difficulties in effectively forecasting the return financial benefits (ROI) of technology investments due to factors such as changing market dynamics, fast-expanding technology trends, and project-specific variables (Singh et al. 2023). Construction projects comprise of sensitive information such as project blueprints, financial data, and stakeholders' personal information. Implementing I4.0 in construction processes exposes them to cyber security

Table 11 Final ranks and global weights of sub- barriers

Criteria	Relative weights	Barriers	Relative weights	Global Weights	Rank
FI	0.655	FI1	0.746	0.489	1
		FI2	0.276	0.181	11
		FI3	0.471	0.309	2
		FI4	0.379	0.248	8
OR	0.372	OR1	0.720	0.268	5
		OR2	0.334	0.124	16
		OR3	0.464	0.173	12
		OR4	0.391	0.145	15
TE	0.520	TE1	0.516	0.268	4
		TE2	0.486	0.253	7
		TE3	0.532	0.277	3
		TE4	0.461	0.240	9
MR	0.021	MR1	0.444	0.009	19
		MR2	0.607	0.013	17
		MR3	0.595	0.012	18
		MR4	0.281	0.006	20
RE	0.399	RE1	0.395	0.158	14
		RE2	0.425	0.170	13
		RE3	0.488	0.195	10
		RE4	0.651	0.260	6

dangers such as data breaches, ransomware attacks, and malicious hacking (Humayun et al. 2021). Construction projects are frequently unique and highly customized, necessitating expertise solutions to fulfill specific needs. "Building Information Modeling (BIM), Internet of Things (IoT), augmented reality (AR), robotics, and sophisticated analytics are all examples of I4.0 technologies. Aripin et al. (2019) argued that integrating and managing these complex technologies within the ecology of construction projects necessitates extensive knowledge and resources."

Moreover, due to poor standards and protocols, companies are hesitant to invest in I4.0 technologies due to concerns about compatibility and implementation obstacles (Singh et al. 2023). The construction industry employs a varied workforce with varying degrees of literacy and abilities, but a lack of digital skills may cause a struggle to operate and troubleshoot I4.0 technologies, leading to productivity losses and frustration. This necessitates specific training programs that cater demands of the construction industry (Siriwardhana and Moehler 2023; Ebekozien et al. 2023). Therefore, ranking of technological sub-criteria is TE3 > TE1 > TE2 > TE4. "The relative significance of the sub-criteria within the category of regulatory barriers is RE4 > RE3 > RE2 > RE1, which illustrates that Companies may be hesitant to incur the financial risk, especially if tax rebates are not available to mitigate potential losses (Mustaffa et al. 2022)." While automation can improve productivity and safety, it also raises ethical concerns regarding labor displacement and its effect on lives in the construction industry. Resistance from labor unions and advocacy groups concerned about job loss may hinder the implementation of I4.0 technologies (Boya et al. 2022). Addressing these ethical and social challenges necessitates consultation with local communities and stakeholders to ensure that technological adoption is consistent with their values and goals. Moreover, without clarity on liability issues, companies may be hesitant to employ I4.0 technologies for fear of legal ramifications in case of an incident (Yap et al. 2022). Companies may perceive these risks as barriers due to the absence of robust risk management strategies and regulatory guidance. Incorporating privacy regulations may need robust encryption, access controls, and data anonymization mechanisms, which might add complexity and cost to the implementation of I4.0 technology (Almatari et al. 2023). The ranking of organizational barrier stands as OR1 > OR3 > OR4 > OR2, implying that without strong leadership and vision at the organizational level, there may be resistance to change from employees who are comfortable with traditional methods (Olsson et al. 2021; Govindan and Arampatzis 2023). Risk-averse leaders may be hesitant to invest in I4.0 efforts unless there is demonstrable proof of success or established case studies in the industry (Orzes et al. 2018). Moreover, the lack of performance measures poses challenges for decision-makers to justify the ROI of I4.0 investments to stakeholders such as investors, clients, and regulatory agencies, thus limiting future funding and support for technology adoption efforts (Singh et al. 2023; Demirkesen and Tezel 2022). Finally, the market-related

barriers ranking is MR2 > MR3 > MR1 > MR4, which outlines that limited awareness of I4.0 potential benefits (Demirkesen and Tezel 2022) and fear of single vendor lock-in (Balkhy et al. 2021) are the most significant barriers in the implementation of I4.0 in the construction industry.

6 Analysis of sensitivity

According to the findings of Chang et al. (2007), even little changes in relative weights can result in a different ranking of obstacles in the final assessment. "The sensitivity analysis is carried out to confirm the stability of the ranking of barriers by observing how it changes with variations in the relative weights of criteria for implementing I4.0" (Singh et al. 2023; Kumar et al. 2021a, b; Singh and Misra 2021). The sensitivity analysis confirms the dominant impact of financial barriers on the prioritization of other factors. The value of the financial barrier is incrementally raised from 0.1 to 0.9, with increments of 0.1, and the resulting alterations in the ranking of other criteria are observed, illustrated in Table 12. Technological barriers show the most significant variations in their weight values. "The ranking of specific sub-barriers fluctuates with the main barriers." The variation in global weight of financial barriers to 0.1 results in data management and security (TE3) at the top position and market volatility (MR4) at the lowest position. While approaching closer to the normalized value of 0.665, high initial investment (FI1) emerges as the most significant barrier. The sensitivity analysis results are illustrated in Table 13, and Fig. 4 depicts variances in the criteria ranking.

Table 12 Impact of incremental increase (0.1 to 0.9) in FI barrier values on other criterias

Criteria	Normalized weight	Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9
FI	0.655	0.100	0.200	0.300	0.400	0.500	0.600	0.700	0.800	0.900
OR	0.372	0.618	0.577	0.536	0.495	0.454	0.413	0.330	0.288	0.246
TE	0.520	0.862	0.805	0.748	0.691	0.634	0.577	0.462	0.404	0.346
MR	0.021	0.033	0.031	0.029	0.027	0.025	0.023	0.018	0.015	0.012
RE	0.399	0.663	0.619	0.575	0.531	0.487	0.443	0.399	0.399	0.399

Table 13 Barriers rank after sensitivity analysis

	0.1	0.2	0.3	0.4	0.5	0.6	0.665 (Normalized)	0.7	0.8	0.9
FI1	13	13	11	7	1	1	1	1	1	1
FI2	16	16	16	16	16	14	11	10	4	4
FI3	14	14	14	13	9	6	2	2	2	2
FI4	15	15	15	15	13	9	8	3	3	3
OR1	4	3	3	3	4	4	5	6	7	7
OR2	12	12	13	14	15	16	16	16	16	16
OR3	9	8	8	9	10	11	12	12	12	12
OR4	11	11	12	12	14	15	15	15	15	15
TE1	3	2	2	2	3	3	4	5	6	6
TE2	5	5	5	5	6	7	7	8	9	9
TE3	1	1	1	1	2	2	3	4	5	5
TE4	6	6	6	6	7	8	9	9	10	10
MR1	19	19	19	19	19	19	19	19	19	19
MR2	17	17	17	17	17	17	17	17	17	17
MR3	18	18	18	18	18	18	18	18	18	18
MR4	20	20	20	20	20	20	20	20	20	20
RE1	10	10	10	11	12	13	14	14	14	14
RE2	8	9	9	10	11	12	13	13	13	13
RE3	7	7	7	8	8	10	10	11	11	11
RE4	2	4	4	4	5	5	6	7	8	8

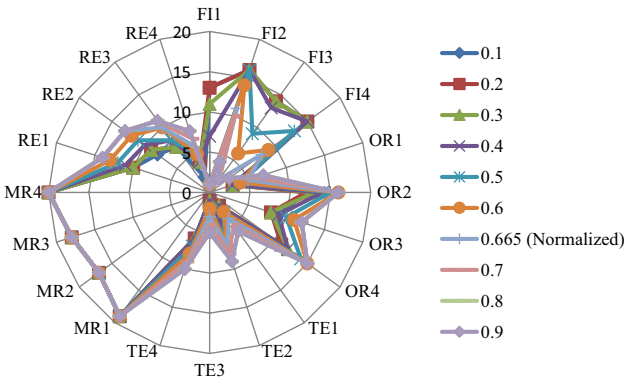


Fig. 4 Ranking variations after sensitivity analysis

7 Conclusion

Due to increasing concerns about resource scarcity and mounting pressure to transition to sustainable business models, organizations have commenced integrating I4.0 practices into their construction projects. "Nevertheless, in developing countries like India, various barriers hinder the efficient implementation of I4.0 in the construction industry." In the case of practical implementation, decision-makers face difficulty in making sound judgments, making it complex to overcome them all at once. "This study explores and prioritizes the barriers using a robust MCDM method for implementing I4.0 in the Indian construction industry. It encompasses 24 barriers categorized into five major groups, identified through a literature review and expert consultation. The fuzzy AHP methodology was utilized to determine the influence ranking of each barrier." An empirical case from an Indian construction company validates the proposed approach. "The findings of this study demonstrate that high initial investment, lack of funding, data security, poor tax rebates, and lack of leadership are the most notable barriers to I4.0 implementation in the construction industry. The outcomes of this study would aid decision-makers and governmental bodies in strategizing tangible regulations that incentivize organizations to adopt I4.0. Finally, sensitivity analysis measures volatility in barriers rank when altering the criteria weights."

8 Managerial implications

The empirical investigation conducted in the Indian construction industry presents a thorough compilation of barriers hindering I4.0 adoption. This investigation of barriers provides valuable insights for construction industry managers to develop strategic plans and integrate processes for implementing I4.0 technologies. This ensures that resources such as budget, time, and personnel are directed toward addressing the most critical barriers first, thereby maximizing the chances of successful implementation. Prioritization enables decision-makers to anticipate resistance to change and develop change management strategies to mitigate potential challenges, such as employee training and cultural shifts. "Findings of this article illustrate the importance of government initiatives in providing economic incentives and tax benefits to encourage organizations to invest in technology and innovation." Managers can use these insights to ensure regulatory compliance and allocate resources efficiently. "Overall, identifying and assessing barriers to I4.0 in the construction industry enables managers to make informed decisions, allocate resources effectively, and traverse the complexities of technology implementation to create innovation and competitive advantage." The study's conclusions not only assist the case company in overcoming considerable barriers but also support its progress towards sustainable development, thereby contributing to the national economy.

9 Practical implications

"The research conducted in this study on Industry 4.0 (I4.0) implementation within the construction industry, particularly in developing nations like India, provides critical insights for industry professionals, policymakers, and decision-makers. Below are the detailed practical implications based on the findings:"

i. *Financial Planning* High initial investments and funding challenges are major barriers to adopting Industry 4.0 in the construction industry. Companies should focus on long-term financial strategies, exploring public–private partnerships or government subsidies to overcome these financial constraints.
ii. *Workforce Development* A significant barrier is a lack of skilled labor for handling advanced technologies like BIM, IoT, and robotics. Implementing targeted training programs and upskilling initiatives is crucial for companies to harness the full potential of Industry 4.0 technologies.
iii. *Data Security and Management* Integrating digital technologies raises concerns about data management and cyber security. Construction companies must invest in secure data handling processes and adhere to regulatory compliance frameworks, such as GDPR, to mitigate risks associated with data breaches.
iv. *Standardization and Protocols* The paper emphasizes standardizing protocols for better system interoperability. Construction firms should collaborate with industry bodies to adopt global standards that

ensure smoother integration of digital tools and processes.

v. *Government and Regulatory Support* Regulatory barriers, including poor tax incentives, hinder Industry 4.0 implementation. Policymakers should provide tax benefits, rebates, and more transparent legal frameworks to encourage firms to adopt advanced technologies without facing financial or legal uncertainties.

These insights can help construction firms and policymakers strategize effectively for a smoother transition to Industry 4.0, enhancing productivity, efficiency, and competitiveness.

10 Validity and future scope

The study emphasized using the Fuzzy Analytical Hierarchy Process (FAHP), a well-established Multi-Criteria Decision-Making (MCDM) technique, which ensures the rigorous prioritization of barriers to Industry 4.0 implementation in the construction industry. The barriers and their prioritization were validated through expert input, ensuring that the study reflects practical and expert-backed insights. We have reiterated that the sensitivity analysis conducted verifies the stability and robustness of the results, confirming the validity of the prioritization framework.

Future research can apply the proposed model to other developing countries or sectors beyond construction to examine if similar barriers exist and if the methodology remains effective across different contexts. We will suggest studying how barriers to Industry 4.0 implementation evolve over time, especially with the rapid advancement of technologies and changing market dynamics. Incorporating other decision-making frameworks, such as Fuzzy DEMATEL, or hybrid approaches combining FAHP with other methods like TOPSIS, could provide more nuanced results. Future research could focus on barriers associated with specific technologies under Industry 4.0 (e.g., BIM, IoT, Robotics) to provide more tailored insights. Further longitudinal studies could track the effectiveness of overcoming these barriers in real-world projects, providing feedback loops for refining the model.

11 Limitations

"This study has limited scope of exploring and prioritizing barriers to I4.0 implementation in India's construction industry. The construction industry is inherently complex, with numerous interrelated factors influencing the implementation of I4.0 technologies. AHP may struggle to capture the full complexity of these relationships, which can be further analyzed through the Fuzzy DEMATEL approach." "Moreover, the suggested theoretical model can be validated by Confirmatory Factor Analysis (CFA) or Structural Equation Modeling (SEM)." The findings of this study can also undergo comparison using various MCDM techniques, such as SWARA and COPRAS, within a fuzzy environment.

Funding The authors did not receive any funding for the submitted work.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest.

Human and animal participants This article does not contain any involvement of Human Participants and/or Animals.

Informed consent Informed consent has been obtained from all individual participants involved in the study."

Appendix A List of abbreviations

S. No.	Abbreviation	Full form	Definition
1	AHP	Analytical Hierarchy Process	A multi-criteria decision-making method that uses pairwise comparisons to rank alternatives or criteria
2	FAHP	Fuzzy Analytical Hierarchy Process	An extension of AHP that integrates fuzzy logic to handle uncertainty and ambiguity in decision-making
3	I4.0	Industry 4.0	The fourth industrial revolution, focused on digital technologies like IoT, robotics, and AI in manufacturing

S. No.	Abbreviation	Full form	Definition
4	MCDM	Multi-Criteria Decision Making	A set of methods for evaluating and prioritizing alternatives based on multiple criteria
5	TFN	Triangular Fuzzy Number	A fuzzy number represented by three values (lower, middle, and upper), used to model uncertainty in data
6	BIM	Building Information Modeling	A digital representation of the physical and functional characteristics of a construction project
7	IoT	Internet of Things	A network of interconnected devices that communicate and exchange data to optimize operations in real-time
8	CFA	Confirmatory Factor Analysis	A statistical technique used to test whether a set of variables aligns with an expected factor structure
9	SEM	Structural Equation Modeling	A multivariate statistical analysis technique used to analyze structural relationships between measured variables
10	SWARA	Step-wise Weight Assessment Ratio Analysis	A decision-making method used for ranking alternatives by assigning weights step-by-step
11	COPRAS	Complex Proportional Assessment	A decision-making method that ranks alternatives by comparing their performance in relation to criteria
12	GDPR	General Data Protection Regulation	A European regulation designed to protect the privacy and data of individuals within the European Union

References

Al Balkhy W, Sweis R, Lafhaj Z (2021) Barriers to adopting lean construction in the construction industry—The case of Jordan. Buildings 11(6):222. https://doi.org/10.3390/buildings11060222

Almatari HAQ, Chan M, Masrom MAN (2023) Factors inhibiting the adoption of industrial revolution 4.0 in Malaysian construction industry. Smart Sustain Built Environ. https://doi.org/10.1108/SASBE-03-2023-0041

Aripin IDM, Zawawi EMA, Ismail Z (2019) Factors influencing the implementation of technologies behind industry 4.0 in the Malaysian construction industry. MATEC Web Conf 266:01006. https://doi.org/10.1051/matecconf/201926601006

Bademosi F, Issa RR (2021) Factors influencing adoption and integration of construction robotics and automation technology in the US. J Constr Eng Manag 147(8):04021075. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002094

Bajpai A, Misra SC (2022) Barriers to implementing digitalization in the Indian construction industry. Int J Qual Reliab Manag 39(10):2438–2464. https://doi.org/10.1108/IJQRM-10-2021-0367

Boya A, Akinradewo O, Aigbavboa C, Ebekozien A, Ramabodu M (2022) Bottlenecks to the implementation of automation and robotics in the construction industry. In: International conference on computing in civil and building engineering, Springer, Switzerland, pp 139–149. https://doi.org/10.1007/978-3-031-11373-8_12

Chang DY (1996) Applications of the extent analysis method on fuzzy AHP. Eur J Oper Res 95(3):649–655. https://doi.org/10.1016/0377-2217(95)00300-2

Chang CW, Wu CR, Lin CT, Chen HC (2007) An application of AHP and sensitivity analysis for selecting the best slicing machine. Comput Ind Eng 52(2):296–307. https://doi.org/10.1016/j.cie.2006.12.008

Demirkesen S, Tezel A (2022) Investigating major challenges for industry 4.0 adoption among construction companies. Eng Constr Archit Manag 29(3):1470–1503. https://doi.org/10.1108/ECAM-10-2020-0821

Ebekozien A, Aigbavboa CO, Adekunle SA, Aliu J, Thwala WD (2023) Training needs of built environment professionals: the role of fourth industrial revolution. Eng Constr Archit Manag. https://doi.org/10.1108/ECAM-12-2021-1185

Govindan K, Arampatzis G (2023) A framework to measure readiness and barriers for the implementation of Industry 4.0: a case approach. Electron Comm Res Appl 59:101249. https://doi.org/10.1016/j.elerap.2022.101249

Hall AT, Durdyev S, Koc K, Ekmekcioglu O, Tupenaite L (2023) Multi-criteria analysis of barriers to building information modeling (BIM) adoption for SMEs in New Zealand construction industry. Eng Constr Archit Manag 30(9):3798–3816. https://doi.org/10.1108/ECAM-06-2022-0606

Hanuar NNA, Kasim N, Zainal R, Abd Rahim MHI, Jusoh ZM (2023) Implementation of the fourth industrial revolution (IR 4.0) for controlling workers' safety at construction sites. Res Manag Technol Bus 4(1):1017–1033. https://doi.org/10.1108/RMOTB-04-2023-0024

Hong J, Shen GQ, Li Z, Zhang B, Zhang W (2018) Barriers to promoting prefabricated construction in China: a cost–benefit analysis. J Clean Prod 172:649–660. https://doi.org/10.1016/j.jclepro.2017.10.147

Humayun M (2021) Industry 4.0 and cyber security issues and challenges. Turkish J Comput Math Edu (TURCOMAT) 12(10):2957–2971

Javaid M, Haleem A, Singh RP, Suman R (2021) Substantial capabilities of robotics in enhancing industry 4.0 implementation. Cognit Robot 1:58–75. https://doi.org/10.1016/j.cogr.2021.07.001

Kumar S, Suhaib M, Asjad M (2021a) Narrowing the barriers to Industry 4.0 practices through PCA-Fuzzy AHP-K means. J Adv Manag Res 18(2):200–226. https://doi.org/10.1108/JAMR-09-2020-0186

Kumar V, Vrat P, Shankar R (2021b) Prioritization of strategies to overcome the barriers in Industry 4.0: a hybrid MCDM approach. Opsearch. https://doi.org/10.1007/s12597-020-00487-7

Lau SEN, Aminudin E, Zakaria R, Saar CC, Roslan AF, Abd Hamid Z, Ahamad AH (2021) Talent as a spearhead of Construction 40 transformation: analysis of their challenges. IOP Conf Ser Mater Sci Eng 1200(1):012025. https://doi.org/10.1088/1757-899X/1200/1/012025

Lee B, Park SK (2022) A study on the competitiveness for the diffusion of smart technology in the construction industry in the era of the 4th industrial revolution. Sustainability 14(14):8348. https://doi.org/10.3390/su14148348

Liu Z, Lu Y, Shen M, Peh LC (2021) Transition from building information modeling (BIM) to integrated digital delivery (IDD) in sustainable building management: a knowledge discovery approach-based review. J Clean Prod 291:125223. https://doi.org/10.1016/j.jclepro.2021.125223

Mansour H, Aminudin E, Mansour T (2023) Implementing industry 4.0 in the construction industry-strategic readiness perspective. Int J Constr Manag 23(9):1457–1470. https://doi.org/10.1080/15623599.2021.1981220

Martinez E, Pfister L (2023) Benefits and limitations of using low-code development to support digitalization in the construction industry. Autom Constr 152:104909. https://doi.org/10.1016/j.autcon.2023.104909

Maskuriy R, Selamat A, Maresova P, Krejcar O, David OO (2019) Industry 4.0 for the construction industry: review of management perspective. Economies 7(3):68. https://doi.org/10.3390/economies7030068

Musarat MA, Alaloul WS, Irfan M, Sreenivasan P, Rabbani MBA (2022) Health and safety improvement through industrial revolution 4.0: Malaysian construction industry case. Sustainability 15(1):201. https://doi.org/10.3390/su15010201

Mustaffa NK, Mat Isa CM, Ekundayo D, Joseph VRA (2022) Barriers and strategies for improving carbon emissions management approaches in Malaysian construction. Constr Econ Build 22(3):99–123. https://doi.org/10.5130/AJCEB.v22i3.7878

Nagy O, Papp I, Szabó RZ (2021) Construction 4.0 organisational level challenges and solutions. Sustainability 13(21):12321. https://doi.org/10.3390/su132112321

Newman C, Edwards D, Martek I, Lai J, Thwala WD, Rillie I (2021) Industry 4.0 deployment in the construction industry: a bibliometric literature review and UK-based case study. Smart Sustain Built Environ 10(4):557–580. https://doi.org/10.1108/SASBE-09-2020-0139

Oesterreich TD, Teuteberg F (2016) Understanding the implications of digitisation and automation in the context of Industry 4.0: a triangulation approach and elements of a research agenda for the construction industry. Comput Ind 83:121–139. https://doi.org/10.1016/j.compind.2016.09.006

Okonkwo C, Okpala I, Awolusi I, Nnaji C (2023) Overcoming barriers to smart safety management system implementation in the construction industry. Results Eng 20:101503. https://doi.org/10.1016/j.rineng.2023.101503

Olsson NO, Arica E, Woods R, Madrid JA (2021) Industry 4.0 in a project context: Introducing 3D printing in construction projects. Project Leadership Soc 2:100033. https://doi.org/10.1016/j.plas.2021.100033

Omrany H, Al-Obaidi KM, Husain A, Ghaffarianhoseini A (2023) Digital twins in the construction industry: a comprehensive review of current implementations, enabling technologies, and future directions. Sustainability 15(14):10908. https://doi.org/10.3390/su151410908

Orzes G, Rauch E, Bednar S, Poklemba R (2018) Industry 4.0 implementation barriers in small and medium-sized enterprises: A focus group study. In: 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), IEEE, pp 1348–1352. https://doi.org/10.1109/IEEM.2018.8607862

Osunsanmi TO, Aigbavboa C, Oke A (2018) Construction 4.0: the future of the construction industry in South Africa. Int J Civ Environ Eng 12(3):206–212

Pathak SK, Sharma V, Chougule SS, Goel V (2022) Prioritization of barriers to the development of renewable energy technologies in India using integrated modified Delphi and AHP method. Sustain Energy Technol Assess 50:101818. https://doi.org/10.1016/j.seta.2021.101818

Prakash C, Barua MK (2015) Integration of AHP-TOPSIS method for prioritizing the solutions of reverse logistics adoption to overcome its barriers under fuzzy environment. J Manuf Syst 37:599–615. https://doi.org/10.1016/j.jmsy.2014.12.018

Ribeiro DB, Coutinho ADR, Satyro WC, Campos FCD, Lima CRC, Contador JC, Gonçalves RF (2024) The DAWN readiness model to assess the level of use of Industry 4.0 technologies in the construction industry in Brazil. Constr Innov 24(2):515–536. https://doi.org/10.1108/CI-06-2022-0140

Saaty TL (1991) Some mathematical concepts of the analytic hierarchy process. Behaviormetrika 18(29):1–9. https://doi.org/10.2333/bhmk.18.29_1

Sajjad M, Hu A, Waqar A, Falqi II, Alsulamy SH, Bageis AS, Alshehri AM (2023) Evaluation of the success of industry 4.0 digitalization practices for sustainable construction management: Chinese construction industry. Buildings 13(7):1668. https://doi.org/10.3390/buildings13071668

Shaawat ME, Jamil R, Al-Enezi MM (2018) Analysis of challenges in sustainable construction industry by using analytic hierarchy process: a case study of Jubail Industrial City, Saudi Arabia. Int J Sustain Real Estate Constr Econ 1(2):109–122

Shafei H, Rahman RA, Lee YS, Che Ibrahim CKI (2024) Implications of Construction 4.0 technologies to enhancing well-being: a fuzzy TOPSIS evaluation. Constr Innov. https://doi.org/10.1108/CI-05-2023-0128

Singh A, Kumar V, Verma P, Kandasamy J (2023) Identification and severity assessment of challenges in the adoption of industry 4.0 in Indian construction industry. Asia Pacific Manag Rev 28(3):299–315. https://doi.org/10.1016/j.apmrv.2021.10.003

Singh A, Misra SC (2021) Identifying challenges in the adoption of industry 4.0 in the Indian construction industry. In: Progress in Advanced Computing and Intelligent Engineering: Proceedings of ICACIE 2019, vol 1. Springer, Singapore, pp 380–398. https://doi.org/10.1007/978-981-15-1081-0_30

Siriwardhana S, Moehler R (2023) Mastering the skills of Construction 4.0: a review of the literature using science mapping. Smart Sustain Built Environ. https://doi.org/10.1108/SASBE-09-2022-0149

Spaan M, Abraham YS (2023) Barriers to and enablers of affordable housing construction: insights from construction industry professionals. Eng Proc 53(1):36. https://doi.org/10.3390/engproc2023053036

Taher G (2021) Industrial revolution 4.0 in the construction industry: challenges and opportunities. Manag Stud Econ Syst 6(3/4):109–127

Tayal A, Agrawal S, Yadav R (2024) Implementation of industry 4.0 in construction industry: a review. Int J Syst Assur Eng Manag 15:1–20. https://doi.org/10.1007/s13198-024-02432-6

Tayal A, Agrawal S, Yadav R (2023) ARank-FSC: assessment and ranking of Industry 4.0 enablers using fuzzy SWARA and fuzzy COPRAS in Indian construction. Intelligent Decision Technologies. Advance online publication. https://doi.org/10.3233/IDT-230005

Umeokafor N, Teizer J (2023) Barriers to digitalisation of construction health, safety, and well-being regulations enforcement and compliance checking. In: CIB W099 & W123 Annual International Conference: Digital Transformation of Health and Safety in Construction, pp 441–450.

Watts G, McDermott P, Kolo S (2023) Framing the barriers to construction industry transformation. Built Environ Project Asset Manag 13(1):185–199. https://doi.org/10.1108/BEPAM-07-2022-0100

Wu P, Jin R, Xu Y, Lin F, Dong Y, Pan Z (2021) The analysis of barriers to BIM implementation for industrialized building construction: a China study. J Civ Eng Manag 27(1):1–13. https://doi.org/10.3846/jcem.2021.13985

Yap JBH, Lam CGY, Skitmore M, Talebian N (2022) Barriers to the adoption of new safety technologies in construction: a developing country context. J Civ Eng Manag 28(2):120–133. https://doi.org/10.3846/jcem.2022.16521

You Z, Feng L (2020) Integration of industry 4.0 related technologies in construction industry: a framework of cyber-physical system. IEEE Access 8:122908–122922. https://doi.org/10.1109/ACCESS.2020.3006953

Zulu SL, Saad AM, Ajayi SO, Dulaimi M, Unuigbe M (2023) Digital leadership enactment in the construction industry: barriers undermining effective transformation. Eng Constr Archit Manag. https://doi.org/10.1108/ECAM-02-2022-0188

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Check for updates

ADVANCED THEORY AND SIMULATIONS
www.advtheorysimul.com

Parametric Analysis of Organic LED and Photo-Diode Based Sensor for Heart Rate Detection

Sugandha Yadav, Poornima Mittal, and Shubham Negi*

Conventional Light emitting diode and photo-diode integration is the best suitable combination for point-of-care diagnostic in the field of sensor applications. However, these approaches based on inorganic devices are not scalable to incorporate multiple bio-marker. As a solution, this work proposes the methodology for bio-medical application using organic LED (OLED) and photo-diode to make a highly efficient, ultra-flexible, and low-cost sensor. In this paper, a green charge generation based OLED and organic photo diode (OPD) are utilized for heart rate measurements. The methodology is basically based on the fluorescence emission where green light having wavelength of 497 nm falls on the blood cells. Furthermore, a different intensity wavelength is sensed by the photodiode. The proposed OLED exhibits current and luminescence parameters of 0.92 A and 8 363.6 cd m^{-2}, correspondingly. Moreover, the presented OPD demonstrates dark currents and photocurrent of 10.2 and 174.5 nA, subsequently. This proposed technique can be also investigated in terms of other biomedical applications such as oxygen sensor, cancer detection, corona detection etc.

1. Introduction

Ultra-thin and flexible sensors are highly promising for health-care applications.[1–7] To produce flexible wearable sensors, two devices are vital: 1) organic LED (OLED) and 2) organic photo diode (OPD). OLED is generally used for light emission which falls on the sample and OPD receives the light which is emitted by the sample. Various researchers have utilized these two devices for so many health-care based sensor applications to make flexible devices. Titov et al.[8] utilized a combination of OLED and OPD which is suitable for low-cost compact optical detection for the sensing applications. Furthermore, Katchman et al.[9] demonstrated quantitative detection of IpG antibodies using human serum with the help of light emitting diode and photo diode.

Lim et al.[10] presented an oxygen sensor device using the same OLED-OPD combination. Herein, simultaneous monitoring of $tcpO_2$ is incorporated for the detection in any part of the human body. Addition to this, Lochner et al.[11] reported a low-cost flexible pulse oximeter using organic devices, LED and photo diode. The proposed device is able to measure oxygenation and pulse rate with the errors of 2% and 1%, correspondingly. Furthermore, Khan et al.[12] demonstrated a reflectance oximeter sensor device using a light emitter, a detector and an optical barrier between emitter and detector. Using this method, the researchers designed three sensor geometries which are rectangular geometry, bracket geometry, and circular geometry. Additionally, Lee et al.[13] proposed a compact flexible OLED-OPD sensing device for covid-19 detection. The device successfully detects SARS-CoV-2 RNA within 20 min using fluorescence detection.

Our work investigates few biomedical applications based on the light detection methodology using OLED-OPD integration. The in-depth discussion about the methodology is provided in the below sections. The complete paper is categorized in seven different sections. As discussed already above, Section I is about the introduction which includes various sensing applications using OLED and OPD. Afterward, Section II showcases the light detection methodology using these devices. Moreover, Section III illustrates the novel OLED structure for heart rate detection. Thereafter, in Section IV, OPD is investigated in order to measure the heart rate in human. Additionally, Section V depicts the methodology for the heart rate detection. Next, Section VI incorporates the comparison of few researchers' work for heart rate detection. Eventually, Section VII incorporates the necessary conclusions.

2. Light Detection Methodology Using OLED-OPD Integration

The block diagram in **Figure 1** depicts methodology for light-based detection technique using OLED and photo-diode. Here, OLED is used as light emitting device whereas, organic photo-diode is utilized as light detector as per their basic characteristics. The light from LED falls on sample which further excites the sample based on the nano particles present inside the sample and fluorescence emission occurs.

Eventually, photodiode provides current values at the specific wavelengths as per the detection technique. This idea of excitation and absorption in terms of wavelength is very promising for

S. Yadav, P. Mittal
Department of Electronics and Communication
Delhi Technological University
New Delhi 110042, India
E-mail: poornimamittal@dtu.ac.in
S. Negi
Department of Electronics and Communication Engineering
Chandigarh University
Mohali, Punjab 140413, India

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/adts.202500037

DOI: 10.1002/adts.202500037

 © 2025 Wiley-VCH GmbH

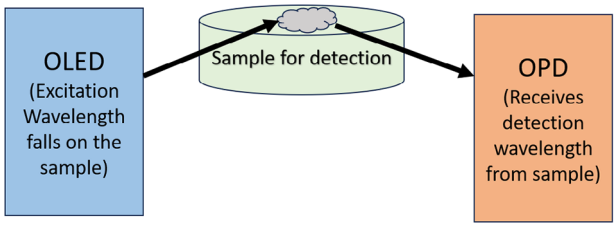


Figure 1. Light based-detection methodology using OLED and OPD.

the light-based detection of various diseases such as oxygen sensing, cancer detection, stroke detection, covid-19 detection,[14–20] etc. Some of the literature survey related to these sensors are discussed in the above section.

3. Proposed Novel OLED Using CGL for Bio-Medical Applications

Herein, a novel structure of green OLED (D_1) is presented. The reference device of this structure is already validated in our previous research work.[21] **Figure 2b** outlines the structure of the

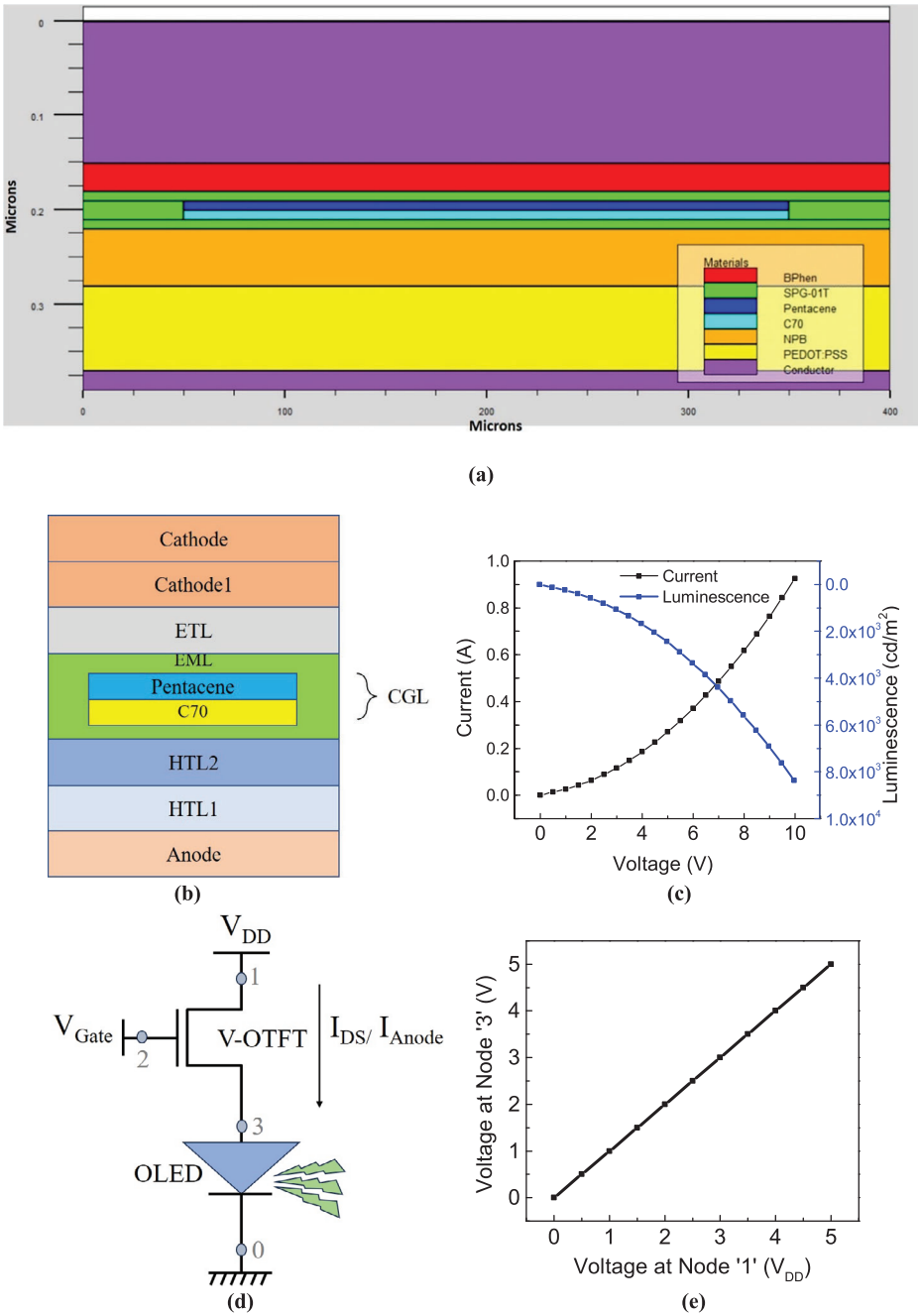


Figure 2. a) Simulated Structure and b) Basic structure of CGL based Novel Green OLED D_1, and c) Current and Luminescence curves of D_1, d) Variation of anode voltage at node 3, and e) Current and luminescence curve of OLED.

© 2025 Wiley-VCH GmbH

25130390, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adts.202500037 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1. Materials of each layer of OLED along with thickness.

Layer	Material and dimensions for proposed device (D_1)	
	Material	Thickness [nm]
Anode	ITO	20
HTL1	PEDOT:PSS	90
HTL2	NPB	60
CGL	C70	10
	Pentacene	10
EML	Polymer green SPG-01T	40
ETL	BPhen	30
Cathode	LiF	1
Cathode 1	Al	150

proposed OLED device. The device contains eight different layers starting from cathode, thereafter electron transport layer (ETL), emissive layer [EML], charge generation layer, hole transport [HTL], and anode. **Table 1** depicts the materials and thickness for all the mentioned layer. The results are obtained using Silvaco Atlas simulation tool wherein, a model, i.e., Langevin's recombination rate is applied in the 2-D device simulator for the study of recombination rate. Moreover, to understand the carrier's transportation in a device structure, Poole-Frenkel mobility model is utilized which provides the results under specified boundary conditions. Figure 2a illustrates the simulated structure of the proposed device.

It is evident that every light has its own specific wavelength. Herein, the calculation of wavelength for the utilized green OLED is performed. As shown in Table 1, the material for EML SPG-01T contains the energy band gap of 2.5 eV.[22]

We know that energy gap (E_g) = $h\nu$where, $\nu = \frac{c}{\lambda}$ hence, Eg = $\frac{hc}{\lambda}$ (in Joule), and

$$\lambda = \frac{hc}{\mathrm{Eg}} \ (in\ meters) \tag{1}$$

In Equation (1), the parameters h, λ, and c represent Planck's constant, wavelength, and speed of light, subsequently.

After incorporating all the values in Equation (1),

$$\lambda = \frac{6.626 \times 10^{-34} \times 3 \times 10^{8}}{2.5 \times 1.6 \times 10^{-19}} \tag{2}$$

$$\lambda = 497\mathrm{nm} \tag{3}$$

It is clear from the final value of λ that the proposed OLED exhibits green color wavelength as the green light ranges from 495 to 570 nm.

To refine the improvement of the reference device on the basis of luminescence and current density, a charge generation layer is incorporated. With the help of CGL, the lost part of light during the recombination can be restored which eventually enhances the device performance in terms of efficiency and EL emission.[23] These layers are made of two different materials (n type & p type) and act same as electrodes. The charge carriers are generated based on the external applied biasing and inserted to the EML where the recombination happens and light is produced.[24] The reason behind opting a suitable CGL layer is completely based on the work function parameter. The work function of charge generation layer should be closely matched with the work functions of the nearby layer. In this research, The CGL is placed in between Bphen and NPB, which contains the HOMO as 6.0 and 5.4 eV, correspondingly. These work functions are closely matched with CGL layer which contains Pentacene (5 eV) and C70 (6.1 eV) materials as hole injector and electron injector, respectively. Thereby, easy facilitation of charge carriers occurs.

The 20 nm CGL is placed inside the EML (see Figure 2b) in the proposed device. It is the combination of two materials which are Pentacene (10 nm) and C70 (10 nm). Inclusion of CGL in OLED contributed significant improvement in charge injection. Thereby, the device exhibits higher current density and refined luminous properties.[25–28] The reason behind the significant enhancement is that in D_1 top layer (Pentacene) of CGL performs as hole injector element and bottom layer (C70) conducts like electron injector. As a result, holes in Pentacene layer will start moving toward cathode and will arrive to EML. In the same manner, electrons in C70 layer will be attracted to anode side and reach to EML. Herein, CGL in D_1 is enclosed inside the EML. Thereby, generated excess electrons and holes are directly moving to the EML. This automatically provides the direct increment in the recombination rate of D_1.

A circuit is designed (see Figure 2d) for driving the above-mentioned OLED using a voltage-controlled device, i.e., vertical OTFT.[29] In vertical TFTs, the channel length is very much reduced (in nano-meters) because of vertical channel formation.[30–32] Moreover, organic TFTs are preferable over inorganic transistor as the organic devices contain excellent mechanical flexibility, low-cost large area fabrication.[33–35] Also, these devices offer the possibility to adapt to low temperature processes for fabrication such as solution printing, dip coating, etc.[36] Hence, V-OTFT is a good choice for utilization to drive any circuit.

The driver V-OTFT is connected in series with the OLED device which balances the amount of current inside the OLED. Thereby the operation can be controlled. It is clear from the Figure 2d that the current is same from V_{DD} to anode terminal because of the series connection of both the devices. Hence, the current across anode is now completely dependent on V_{DD} and gate voltage which eventually control the luminescence of OLED.

The V-OTFT provides the current when it is conducted in the linear or saturation region for which the equations are given below,

$$V_{DS} < V_{GS} - V_{th}; \quad Linear\ Region \tag{4}$$

$$V_{DS} \geq V_{GS} - V_{th}; \quad Saturation\ Region \tag{5}$$

where, "V_{th}" is the threshold voltage, "V_{GS}" and "V_{DS}" stand for gate to source voltage and drain to source voltages, subsequently. Now to drive the OLED, 5 V voltage is required at node 3 (see Figure 2d) to attain maximum current density and luminescence. The circuit is operated in linear region to achieve the control on the OLED performance parameters by varying the V_{DS}. Here, the gate voltage is constant at 5 V and the drain voltage VDD is vary-

© 2025 Wiley-VCH GmbH

25130390, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adts.202500037 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 2. Materials and thickness for each layer of D_2.

Layer	Material and dimensions for Proposed Device (D_2)	
	Material	Thickness [nm]
Anode	ITO	20
HTL1	PEDOT:PSS	90
CGL	C70	10
	Pentacene	10
Acceptor	C60	40
Donor	CuPc	20
ETL	BPhen	12
Cathode	Al	150
Cathode 1	LiF	1

ing from 0 to 5 V. Figure 2c illustrates the results of the proposed driver circuit.

The Figure 2e illustrates a variation in anode voltage at node 3. It is proven here that the proposed circuitry is capable of generating a potential difference of 5 V at anode for the proper operation of OLED. The current and luminescence curves are obtained as 0.92 A and 8 363.6 cd m^{-2}, respectively, at 5 V. It can be concluded that anode voltage of the OLED is controlled by varying the V_{DD} as illustrated in Figure 2d which further regulates the luminescence parameter of the light emitting diode.

4. Novel OPD Using CGL for Bio-Medical Applications

First, a multilayered OPD is validated in our previous work[37] with the simulation tool Silvaco Atlas to set the boundary conditions for the proposed work in terms of materials, thickness, and other important parameters. Thereafter, in this work, a novel structure of photo-diode (D_2) is proposed which contains CGL layer inside the active layer. This CGL layer is incorporated to design a highly efficient OPD which consists of Pentacene and C70 as hole injectors and electron injectors, correspondingly. Addition to this, the electrodes terminals anode and cathode consist of indium Tin Oxide (ITO) and Aluminum (Al), respectively. Furthermore, ETL and hole transport layer (HTL) consist of the materials BPhen and PEDOT:PSS (poly(3,4-ethylenedioxy thiophene):polystyrene sulfonate), correspondingly. Moreover, the active layer which is the combination of acceptor and donor is composed of C60 and CuPc, respectively. **Table 2** demonstrates all the incorporated materials along with their thickness for each layer. **Figure 3**a depicts the simulated structure of the novel photo diode.

The in-depth explanation of the working of the CGL based proposed device D_2 is provided now. The CGL is the mixture of two different layers; one is C70 and the other is Pentacene which are opted to generate free electrons and holes, correspondingly. As illustrated in Figure 3b, CGL layers are placed inside the active layer. The thickness of the incorporated C70 and Pentacene is 50 and 20 nm, respectively. After providing external voltage to D_2 in such a pattern where anode is attached to the negative terminal and cathode is attached to the positive terminal of the battery, a barrier potential is generated. Hence, a depletion region is developed. The charge carriers inside the depletion region are bounded with this barrier potential and are not able to move further. Thereby, no current flow is attainable. At the moment, when light encounters the acceptor-donor layer, the immobile electrons-holes pairs will procure some energy and become mobile. Furthermore, large number of electrons from C70 layer will be attracted toward anode. In the same way, Pentacene generates large number of free holes which further reach to the cathode. These mobile charge carriers will be added up with the active layer's charge carriers. As a result, significant large number of dark currents and photocurrent are obtained as 10.2 and 174.5 nA, respectively as shown in Figure 3c. Hence, it is made clear that charge generation layer played vital role for refining the performance of the proposed D_2 structure. Figure 3d depicts the current values at different wavelengths.

5. Heart Rate Detection

Photoplethysmography (PPG) is very popular medical technology to measure heart rate and blood oxygen in real time.[38,39] This device was most widely used device during the 2019 pandemic to measure the health conditions. The working principle follows the mentioned (In Section II) light detection methodology which contains a LED and a photo-diode. A target light beam falls from the LED on the blood tissue, and photo-diode senses it and converts it into an electrical signal. Various researchers have devoted their efforts in designing a highly efficient, motion tolerant, and low noise PPG sensor. Kim et al.[40] proposed an "All in one sensor" by integrating OLED an OPD which can detect heart rate, blood pressure, and heart rate variability. Moreover, Lochner et al.[11] presented all organic optoelectronic sensor to measure the human pulse rate using green and red OLEDs and single flexible OPD. The device exhibits only 1% error in the pulse rate measurement.

Pandey et al.[41] also designed an OLED-OPD flexible sensor in which heart rate is measured with 96% accuracy. Addition to this, the ideas for flexible and wearable devices for pulse oximetry sensor are discussed by Dcosta et al.[42] In this research, the geometry, design, processing technique, encapsulation, etc. are also provided. Furthermore, Jhuma et al.[43] presented a review which is focused on the growth of PPG sensors in terms of their flexibility. Moreover, Simone et al.[44] proposed an organic reflectance PPG array using 16 × 16 OPD pixel. The pixel device exhibits the sensitivity up to 950 nm and low dark current as 10^{-6} mA cm^{-2}. Additionally, Bilgaiyan et al.[45] investigate OPD performance for refining the sensing ability of PPG sensor.

In our work, we have presented a methodology for the detection of heart rate using flexible novel green OLED (D_1) and OPD (D_2) as outlined in **Figure 4**. The utilization of green OLEDs in PPGs is quite prevalent. This light contains high skin absorption coefficient. So, it provides better contrast signal.[47] **Table 3** outlines the few researchers' work for the comparison with the proposed work in terms of OLED color and its wavelength. It can be clearly seen that all the excitation wavelengths are from range of green OLED. The detection methodology is basically based on the emission of wavelength from OLED and detection of wavelengths at OPD. A target light beam of 497 nm emitted from OLED falls on human blood tissues. Some of this light is absorbed by the blood and remaining light is detected by the photo-diode, cor-

© 2025 Wiley-VCH GmbH

25130390, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adts.202500037 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

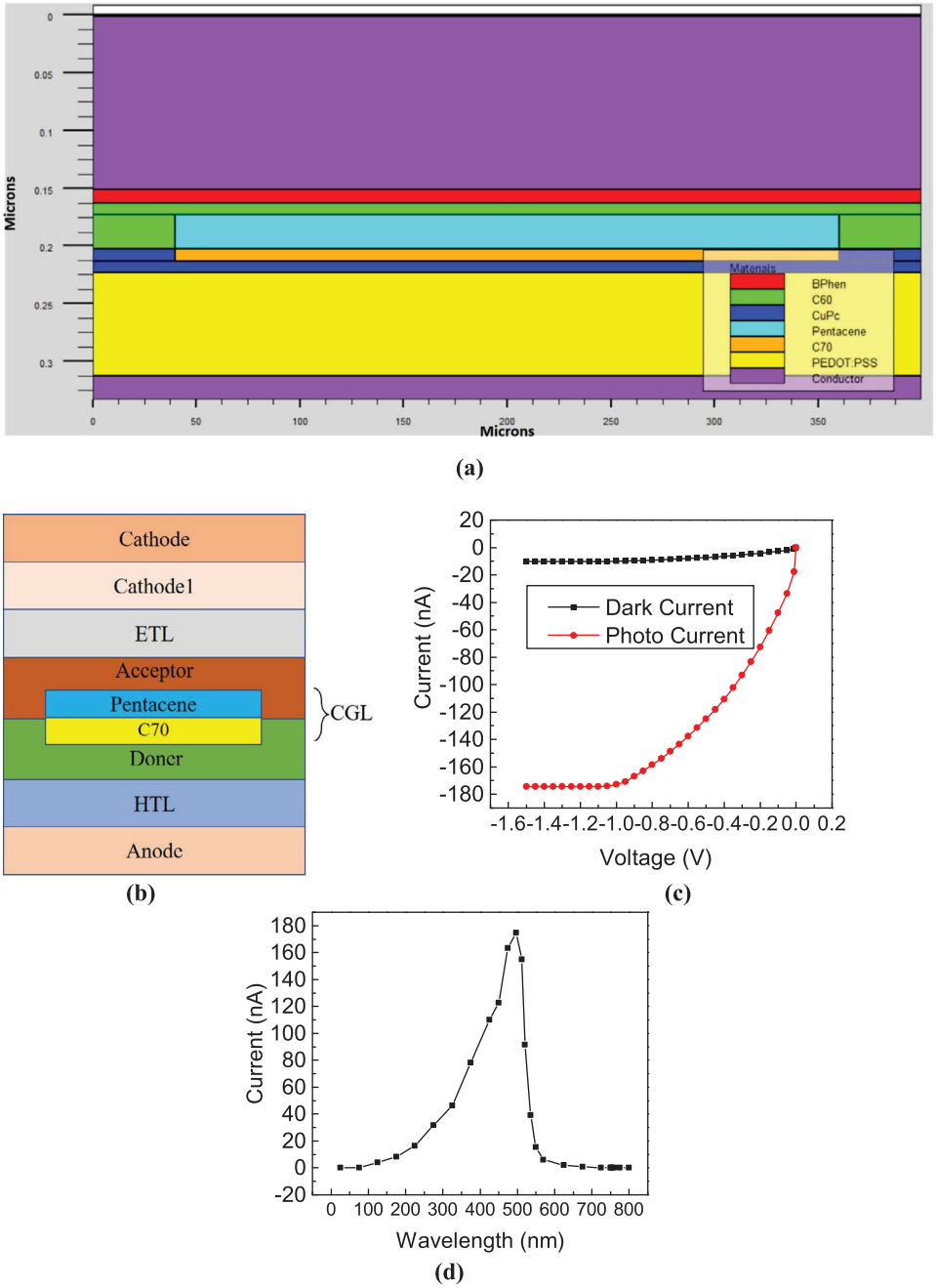


Figure 3. a) Simulated OPD, b) Proposed Design of OPD, c) Dark and Photo Currents and d) Current Values at various Wavelengths.

respondingly a photo current is produced. If the volume of the arterial blood is higher in case of systole (during the heart's contraction phase), it will absorb large amount of light. Hence, low intensity wavelength will be reflected toward the photo-diode. Thereby, less current will be obtained by the photo-diode. On the contrary, in case of diastole (during the heart's relaxation phase) when blood volume is less, blood tissues will absorb less amount of light. So higher intensity wavelength will be detected by the photo-diode which will eventually produce higher output current. Herein, this methodology concludes that less photocurrent results in higher heart-rate and vice-versa as outlined in **Table 4**.

The above discussed methodology calculates the current outputs at different light intensities having the same wavelength of 497 nm. The variation of light intensities is because of the change in blood volume. Herein, it is investigated that the proposed photo diode is capable of producing the different current values at different light intensities by varying the input voltage from 0–5 V using V-OTFT.

Here, ten luminescence values (L_1–L_{10}) at different voltages are taken from the output of the proposed D_1 device as outlined in **Table 5**. As drawn in **Figure 5a**, when the light beams (B_1–B_{10}) having different luminous values (L_1–L_{10}) with the same target

© 2025 Wiley-VCH GmbH

25130390, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adts.202500037 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

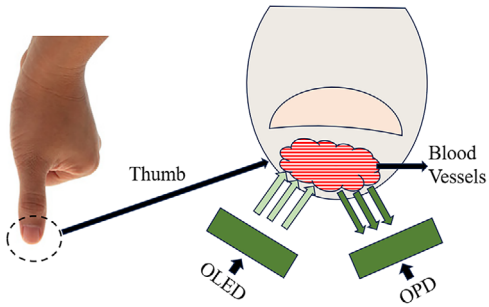


Figure 4. Heart Rate Detection Methodology.

Table 3. Comparison of existing work in terms of wavelength.

Ref.	OLED Color	Wavelength [nm]	Sample	Detection
Prop.	Green	497	Blood cells	Heart Rate
[46]	Green	527	Blood cells	Heart Rate
[11]	Green	532	Blood cells	Heart Rate
[40]	Green	530	Blood cells	Heart Rate
[41]	Green	530	Blood cells	Heart Rate

wavelength of 497 nm are sensed by the photo diode, it will produce different values of currents as indicated in **Table 6**. Because, all the light beams contain different intensity and luminescence. Figure 5b shows the light intensities at various light beams. This proposed method can be compared with the real time situations where volume of blood in the blood vessels is varied all the time. Accordingly, OPD produces output currents and the device concludes the heart rate.

Table 4. Technique for Heart Rate Detection.

Excitation Wavelength	Detection Wavelength	Blood Volume	Output Current	Heart Rate
497	497	Higher	Low	High
		Lower	High	Low

Table 5. Luminescence at different voltages.

Voltage [V]	Luminescence [cd m^{-2}]
0.5	L_1 = 124.9
1	L_2 = 239.8
1.5	L_3 = 388.9
2	L_4 = 574.5
2.5	L_5 = 795.6
3	L_6 = 1 053.1
3.5	L_7 = 1 346.4
4	L_8 = 1 675.5
4.5	L_9 = 2 040.4
5	L_{10} = 2 440.6

Table 6. Current variations at different light beams for detection.

Light beams fall on OPD with different luminous values at different voltages (emission wavelength = 497 nm)	Currents (nA) obtained at detector corresponding to B_1–B_{10}
B_1	17.7
B_2	33.5
B_3	47.7
B_4	72.5
B_5	83.3
B_6	110.5
B_7	125.1
B_8	137.7
B_9	143.5
B_{10}	153.9

Table 7. Different bio-medical applications based on OLED-OPD sensors.

Ref.		Input/emission wavelength [nm]	Output/ Detection wavelength [nm]	Application
[11]	532	640	Pulse oximeter sensor	
[9]	520	605	Cancer detection	
[48]	330	440	Ovarian cancer detection	
[16]	380	497	Oxygen sensor	
[12]	720	420	Optoelectronic sensor	
[49]	490	625	Oxygen sensor	

6. Other Bio-Medical Applications Based on OLED-OPD Sensor

This section draws some other applications based on the sensing devices using OLED-OPD integration as highlighted in **Table 7**. Various researchers have utilized these flexible combinations using different emission and detection wavelengths for bio-medical diseases. Lochner et al.[11] presented a pulse oximeter sensor by using OLED and photo diode. Herein, the input wavelength is considered as 532 nm which represents the green color light, and detection wavelength is opted as 640 nm which is responsible for the red color light. Moreover, Katchman et al.[9] proposed point-of-care detection of circulating cancer biomarkers. In this research, emission wavelength is chosen as 520 nm (green color), and detection wavelength is chosen as 604 nm (orange color).

Negi et al.[48] worked on an ovarian cancer detection using 330 nm emission wavelength and 440 nm detection wavelength. Furthermore, Lim et al.[16] discussed the methodology for wearable oxygen sensor using OLED-OPD integration. The emission and detection wavelength are opted here as 380 and 497 nm. Additionally, Khan et al.[12] proposed an organic multi-channel optoelectronic sensor. This work is suitable for various health monitoring applications. Herein, red color light is utilized for the emission with the wavelength of 720 nm, and violet color wavelength is used for the detection with 420 nm wavelength.

© 2025 Wiley-VCH GmbH

25130390, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adts.202500037 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

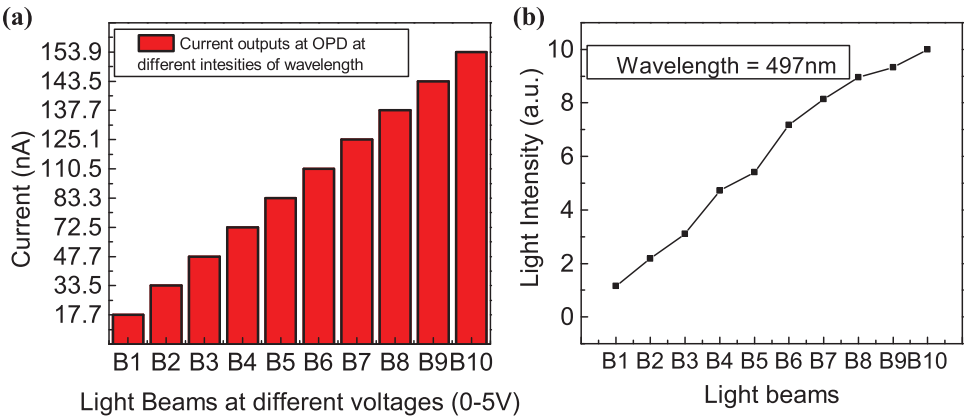


Figure 5. a) Current values at different light beams, and b) Light intensity at different beams.

7. Conclusion

This work introduces a technique for heart rate detection using CGL based ultra-flexible OLED and OPD. A green OLED is opted as a light source which contains the wavelength of 497 nm. To drive the novel OLED, a vertical thin film transistor is utilized which is also flexible and includes less channel length for smooth and efficient performance. Afterward, this paper proposed an organic photo-diode which produces a peak photocurrent and dark current of 10.2 and 174.5 nA, respectively. Herein, the current density and luminescence by the OLED are obtained as 0.92 A and 8 363.6 cd m^{-2}, correspondingly. The additional layer, i.e., CGL used in the both devices makes them highly efficient. The methodology depicts that during the diastole, higher current will be produced by the OPD as blood cell absorbs less green light. On the contrary, the OPD provides lesser current in the case of systole due to less absorption of the green light by presented blood cells in the blood vessels. The discussed method may be further explored in the other sensing bio-medical applications such as cancer detection, PPG sensor, oxygen sensor, hemoglobin sensor etc.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Sugandha Yadav dealt with conceptualization, methodology, data curation, writing–original draft, and writing–review & editing. Poornima Mittal performed conceptualization, methodology, supervision, data curation, writing–review & editing, and validation. Shubham Negi carried out conceptualization, supervision, writing–review & editing, visualization, investigation, and formal analysis.

Ethics Approval

The manuscript is not submitted to anywhere else for the publication. The submitted work is an original work and have not been published elsewhere in any form or language. The authors do not have any conflict of interest. The research does not involve any human participant and/or animal thus, informed consent is not required.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

biomedical sensor, heart rate measurements, OLED, OPD, V-OTFT

Received: January 7, 2025
Revised: March 26, 2025
Published online:

[1] J. C. Y. Kenry, C. T. Lim, *Microsyst. Nanoeng.* **2016**, *2*, 16043.
[2] T. Mayr, T. Abel, E. Kraker, S. Köstler, A. Haase, C. Konrad, M. Tscherner, B. Lamprecht, *Procedia Eng.* **2010**, *5*, 1005.
[3] R. Liu, T. Xiao, W. Cui, J. Shinar, R. Shinar, *Anal. Chim. Acta.* **2013**, *778*, 70.
[4] D. Gudeika, O. Bezvikonnyi, D. Volyniuk, E. Skuodis, P. H. Lee, C. H. Chen, J. V. Grazulevicius, *J. Mater. Chem. C* **2020**, *8*, 9632.
[5] T. Yokota, P. Zalar, M. Kaltenbrunner, H. Jinno, N. Matsuhisa, H. Kitanosako, Y. Tachibana, W. Yukita, M. Koizumi, T. Someya, *Sci. Adv.* **2016**, *2*, 1501856.
[6] D. Threm, J. L. Gugat, A. Pradana, M. Radler, J. Mikat, M. Gerken, *IEEE Photonics Technol. Lett.* **2012**, *24*, 912.
[7] W. L. Jia, M. J. Moran, Y. Y. Yuan, Z. H. Lu, S. Wang, *J. Mater. Chem.* **2005**, *15*, 3326.
[8] I. Titov, M. Köpke, N. C. Schneidewind, J. Buhl, Y. Murat, M. Gerken, *IEEE Sens. J.* **2020**, *20*, 7540.
[9] A. B. Katchman, J. T. Smith, U. Obahiagbon, S. Kesiraju, Y. K. Lee, B. O'Brien, K. Kaftanoglu, J. B. Christen, K. S. Anderson, *Sci. Rep.* **2016**, *6*, 29057.
[10] C. J. Lim, S. Lee, J. H. Kim, H. J. Kil, Y. C. Kim, J. W. Park, *ACS Appl. Mater. Interfaces* **2018**, *10*, 41026.
[11] C. M. Lochner, Y. Khan, A. Pierre, A. C. Arias, *Nat. Commun.* **2014**, *5*, 5745.
[12] Y. Khan, D. Han, J. Ting, M. Ahmed, R. Nagisetty, A. C. Arias, *IEEE Access* **2019**, *7*, 128114.
[13] S. Yadav, P. Mittal, S. Negi, *J. Comput. Electron.* **2023**, *22*, 1696.
[14] C. Murawski, M. C. Gather, *Adv. Opt. Mater.* **2021**, *9*, 2100269.
[15] R. Liu, T. Xiao, W. Cui, J. Shinar, R. Shinar, *Anal. Chim. Acta.* **2013**, *778*, 70.
[16] C. J. Lim, S. Lee, J. H. Kim, H. J. Kil, Y. C. Kim, J. W. Park, *ACS Appl. Mater. Interfaces* **2018**, *10*, 41026.

© 2025 Wiley-VCH GmbH

25130390, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adts.202500037 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

[17] S. Yadav, P. Mittal, S. Negi, *IEEE Sens. J.* **2024**, *24*, 40678.
[18] S. S. Dhanabalan, S. Sriram, S. Walia, S. R. Avaninathan, M. F. Carrasco, M. Bhaskaran, *Adv. Photonics Res.* **2021**, *2*, 2000076.
[19] I. Lee, N. Park, H. Lee, C. Hwang, J. H. Kim, S. Park, *Appl. Sci.* **2021**, *11*, 2313.
[20] M. Ha, S. Lim, H. Ko, *J. Mater. Chem. B.* **2018**, *6*, 4043.
[21] S. Yadav, P. Mittal, S. Negi, *J. Comput. Electron.* **2023**, *22*, 1696.
[22] K. Papadopoulos, D. Tselekidou, A. Zachariadis, A. Laskarakis, S. Logothetidis, M. Gioti, *Nanomaterials* **2024**, *14*, 1608.
[23] A. Jetly, R. Mehra, *Opt. Mater.* **2019**, *88*, 304.
[24] S. Yadav, P. Mittal, S. Negi, *ECS J. Solid State Sci. Technol.* **2024**, *13*, 026001.
[25] D. Zhao, H. Liu, Y. Miao, H. Wang, B. Zhao, Y. Hao, F. Zhu, B. Xu, *Org. Electron.* **2016**, *32*, 1.
[26] H. Sun, Q. Guo, D. Yang, Y. Chen, J. Chen, D. Ma, *ACS Photonics* **2015**, *2*, 271.
[27] S. Hamwi, J. Meyer, M. Kröger, T. Winkler, M. Witte, T. Riedl, A. Kahn, W. Kowalsky, *Adv. Funct. Mater.* **2010**, *20*, 1762.
[28] Y. Liu, X. Wu, Z. Xiao, J. Gao, J. Zhang, H. Rui, X. Lin, N. Zhang, Y. Hua, S. Yin, *Appl. Surf. Sci.* **2017**, *413*, 302.
[29] S. Yadav, P. Mittal, S. Negi, *Phys. Scr.* **2024**, *99*, 025940.
[30] Q. Chen, Y. Yan, X. Wu, S. Lan, D. Hu, Y. Fang, D. Lv, J. Zhong, H. Chen, T. Guo, *ACS Appl. Mater. Interfaces* **2019**, *11*, 35888.
[31] N. A. Roslan, S. M. Abdullah, M. Z. M. Halizan, T. M. Bawazeer, N. Alsenany, M. S. Alsoufi, W. H. A. Majid, A. Supangat, *J. Electron. Mater.* **2018**, *47*, 2184.
[32] Y. Do Kim, J. W. Park, I. N. Kang, S. Y. Oh, *Ultramicroscopy* **2008**, *108*, 1237.
[33] P. Mittal, *J. Soc. Inf. Disp.* **2021**, *29*, 642.
[34] P. Mittal, Y. S. Negi, R. K. Singh, *J. Comput. Electron.* **2015**, *14*, 360.
[35] M. C. Um, J. Kwak, J. P. Hong, J. Kang, S. H. Lee, C. Lee, J. I. Hong, *J. Mater. Chem.* **2008**, *18*, 4698.
[36] B. Kumar, B. K. Kaushik, Y. S. Negi, *IET Circuits, Devices & Systems* **2014**, *8*, 131.
[37] S. Yadav, P. Mittal, S. Negi, *ECS J. Solid State Sci. Technol.* **2024**, *13*, 026001.
[38] Y. Lee, J. W. Chung, G. H. Lee, H. Kang, J. Y. Kim, C. Bae, S. Kim, *Sci. Adv.* **2021**, *7*, abg9180.
[39] D. Tordera, B. Peeters, H. B. Akkerman, A. J. van Breemen, J. Maas, S. Shanmugam, G. H. Gelinck, *Adv. Mater. Technol.* **2019**, *4*, 1900651.
[40] C. Kim, K. S. Bae, G. Kim, D. Y. Lee, G. Moon, D. Yang, C. Lee, *Commun. Eng.* **2024**, *3*, 92.
[41] R. K. Pandey, P. C. P. Chao, *IEEE Trans. Biomed. Circuits Syst.* **2021**, *16*, 36.
[42] J. V. Dcosta, D. Ochoa, S. Sanaur, *Adv. Sci.* **2023**, *10*, 2302752.
[43] F. A. Jhuma, R. Hattori, *Proc. of Int. Exchange and Innovation Conf. on Engineering & Sciences (IEICES)*, **2022**, Vol *8*, pp. 246-252.
[44] G. Simone, D. Tordera, E. Delvitto, B. Peeters, A. J. van Breemen, S. C. Meskers, G. H. Gelinck, *Adv. Opt. Mater.* **2020**, *8*, 1901989.
[45] A. Bilgaiyan, F. Elsamnah, H. Ishidai, C. H. Shim, M. A. B. Misran, C. Adachi, R. Hattori, *ACS Appl. Electron. Mater.* **2020**, *2*, 1280.
[46] C. Murawski, M. C. Gather, *Adv. Opt. Mater.* **2021**, *9*, 2100269.
[47] A. N. Bashkatov, E. A. Genina, V. I. Kochubey, V. V. Tuchin, *J. Phys. D: Appl. Phys.* **2005**, *38*, 2543.
[48] S. Negi, P. Mitta, B. Kumar, A. Bhatia, In 2018 5th IEEE Uttar Pradesh Section Int. Conf. on Electrical, Electronics and Computer Engineering (UPCON), IEEE, Gorakhpur, India **2018**, pp. 1–4.
[49] N. Schneidewind, M. Köpke, J. Buhl, Y. Murat, M. Gerken, IEEE Int. Conf. on Flexible and Printable Sensors and Systems (FLEPS), IEEE, **2019**, pp. 1–3.

© 2025 Wiley-VCH GmbH

25130390, 0, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/adts.202500037 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Materials Research Proceedings 49 (2025) 186-196 https://doi.org/10.21741/9781644903438-19

Performance study of a retrofitted electric two-wheeler using GT-suite

Alpha Omar JALLOW[1,a*], Alphaus NKANDU[2,b], Towett SAMUEL[1,c], Ashish Rajeshwar KULKARNI[1,d], Amit PAL[2,e]

[1]Department of Electrical Engineering, Delhi Technological University, Delhi, India

[2]Department of Mechanical Engineering, Delhi Technological University, Delhi, India

[a]jallowalphaomar100@gmail.com, [b]alphausnkandu2@gmail.com, [c]sktowett@gmail.com, [d]bashishkulkarni@dtu.ac.in, [e]amitpal@dce.ac.in

Keywords: Electric Vehicle (EV) Retrofit, Internal Combustion Engine (ICE), Brushless DC Motor, Lithium-ion Battery, GT Suite Simulation, Electric Two-Wheeler (e-TW)

Abstract. This paper presents the performance study of a retrofitted Bajaj DTS Si motorcycle, converted into an electric two-wheeler (e-TW). Initially, this two-wheeler had internal combustion engine (ICE). This paper presents two distinctive simulation models of retrofitted e-TWs, using GT Suite platform. The first model explores the feasibility of upgrading from the ICE vehicle to an e-TW, with a 1.5 kW Hub motor and 48V/28Ah battery. This model projects speed of 52 km/h and a distance over 29km. The second model explores the feasibility of upgrading from the first model to a new e-TW, with a 5 kW BLDC motor and 72V/78Ah battery. This model projects speed of 80.44 km/h and a range over 130.25 km on a single charge of the battery. This paper confers the details of designing Simulink models using the GT Suite platform. It includes theoretical calculations, simulation methodology and analysis of results, highlighting the software's utility in electric vehicle design and optimization.

Introduction

In major Indian cities, 66% of pollution is contributed by internal combustion engine (ICE) driven two-wheelers. One hundred forty-eight million two-wheelers were sold in India during the year 2011-2020. These two-wheelers include motorcycles, scooters, and mopeds. Twenty-one million two-wheelers were sold during 2019 itself. However, the radical increase in air pollution obliges the transition from ICE based vehicles to electric two-wheelers (e-TWs) [1].

In response to the pressing need to combat air pollution and reduce carbon emissions, there is a growing imperative to transition away from internal combustion engine (ICE) vehicles towards electric alternatives. However, the high cost of new EVs often presents a significant barrier to widespread adoption. Retrofitting existing ICE vehicles with electric drivetrains offers a pragmatic and cost-effective solution to accelerate the transition towards clean transportation.

Furthermore, discarding vehicles running on ICE results in misuse of resources and faster environmental degradation. Retrofitting, therefore, is the best option for replacement of already existing ICE vehicles.

Through meticulous research and engineering innovation, this paper aims to redefine the electric motorbike landscape. This offers consumers an electric alternative that not only meets but surpasses their expectations, but also paves the way for a future where EVs set the standard for performance and environmental responsibility.

The Indian driving cycle (IDC) was developed and formulated in 1985 after thorough road testing and research conducted by the Automotive Research Association of India (ARAI) [2].

This paper explores the feasibility and benefits of retrofitting the body of an ICE two-wheeler with an electric motor bike, by comparing the ICE model with the retrofitted 1.5 kW hub motor

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 license. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. Published under license by Materials Research Forum LLC.

and the retrofitted 5 kW BLDC motor [3]. This paper utilized the IDC to replicate the typical operation of the retrofitted bike on Indian terrain.

Description of research methodology

This work focuses on enhancement of speed and range of a conventional ICE vehicle by its retro-fitment to an e-TW. Considering the basic retrofitted BAJAJ DTS SI motorcycle, having speed of 52 km/h and distance range of 29 km, the research aims to surpass the capabilities of ICE vehicles. This paper strives to match and eventually exceed the performance standards of traditional vehicles while prioritizing efficiency and sustainability [4].

Equipment of Retrofitted EV.

The basic block diagram representing a retrofitted EV is shown in Figure 1. The retrofitted EV consists of following main equipment: (1) Throttle, (2) Battery Controller, (3) Motor Controller, (4) Motor, (5) Wheels.

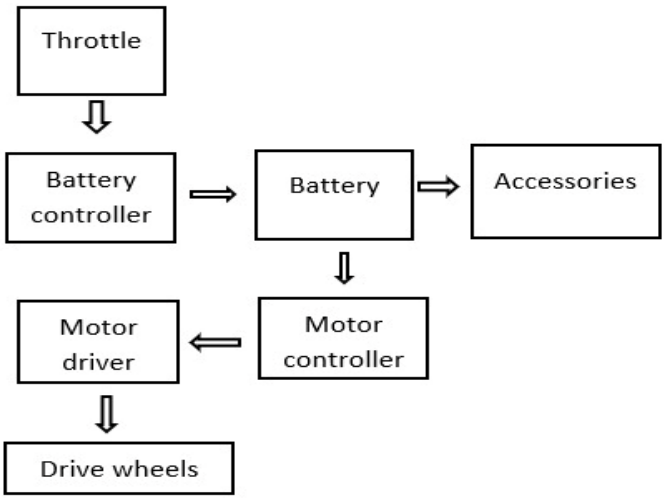


Figure 1. Schematic diagram of a two-wheeler

In Figure 1, the rider's twist of the throttle translates into a control signal. This signal is sent to the battery controller, which manages the flow of electricity from the battery. Based on the throttle input, the battery controller regulates voltage and current output to the motor controller.

The motor controller then translates this signal into precise power delivery for the motor driver. Finally, the motor driver controls the high-power current that spins the motor, propelling the motorbike forward through the wheels.

Contribution of the Work.

This research work contributes following significant aspects to the EV field:

1. Usage of vehicle dynamics is incorporated in order to perform the computation of motor power and battery capacity.
2. Mathematical modelling of battery pack is presented in accordance with the motor capacity and range of the vehicle.
3. The GT Suite v2023 is utilized for designing the simulation models in order to verify feasibility of analytical models.
4. For validating the obtained results from analytical models and simulation models, the assimilation of IDC is performed for highlighting the significance of the work in real time scenario with respect to Indian road conditions.

Analytical Modelling.

In order to compute motor power and battery capacity using vehicle dynamics, four forces acting on the vehicle are needed to be considered. These forces are: (1) Rolling resistance force (F_{rr}), (2) Aerodynamic drag force (F_{ad}), (3) Gradient force (F_g) and (4) Acceleration force (F_a) [5]. Various specifications of the vehicle used in the analytical modelling is presented in Table 1.

Table 1. Vehicle Specifications for analytical modelling

Parameter	**Value**
Total Mass, m	$220\ kg$
Gear Ratio, G_r	6
Transmission efficiency, η	0.97
Rolling coefficient, F_r	0.013
Drag coefficient, C_d	0.9
Vehicle Frontal area, A_f	$0.75\ m^2$
Mass density of air, ρ	$1.225\ kg/m^3$
Wheel radius, r	$0.216\ m$
Velocity, v	$78\ km/h$
Distance Covered, s	$100\ km$
Gradient Angle, θ	0°

The rolling resistance force is computed using Equation (1). The velocity, V, in SI units, is $21.67\ m/s$. The aerodynamic drag force is given by Equation (2). By considering zero initial velocity (u) and computing acceleration, a, force due to acceleration, F_a, is calculated using Equation (3)-(4). The Gradient force is also calculated by using Equation (5).

$$F_{rr} = m \times g \times F_r \times \cos\theta \quad (1)$$

$$F_{ad} = 0.5 \times \rho \times A_f \times C_d \times V \quad (2)$$

$$a = \frac{V^2 - u^2}{2 \times s} \quad (3)$$

$$F_a = m \times a \quad (4)$$

$$F_g = m \times g \times \sin\theta \quad (5)$$

The calculation of total traction force is performed using Equation (6). The wheel torque, T_w, is calculated using Equation (7). Equation (8) gives the wheel speed, N_w. Thus, motor torque (T_m), motor speed (N_m) and motor power (P_m) are computed using Equation (9), (10), (11) respectively.

$$F = F_{rr} + F_{ad} + F_a + F_g \quad (6)$$

$$T_w = F \times r \quad (7)$$

$$N_w = \frac{V}{2 \times \pi \times \frac{r}{60}} \quad (8)$$

$$T_m = \frac{T_w}{G_r \times \eta} \quad (9)$$

$$N_m = G_r \times N_w \times \eta \quad (10)$$

$$P_m = \frac{2 \times \pi \times N_m \times T_m}{60} \quad (11)$$

For battery pack design calculations, assuming the controller efficiency of 0.95, total energy of battery, E_{b_t}, is computed using Equation (12)-(13).

$$t = s/v \quad (12)$$

$$E_{b_t} = \frac{P_m \times t}{controller\ efficiency} \quad (13)$$

The available energy from battery is thus calculated using Equation (14). {The Depth of Discharge (DOD) value is assumed to be 0.9}. In this work, the battery pack of voltage rating 72 V is used, with nominal voltage of each cell being 3.65 V. So, the number of cells in series are found to be 20. Hence, battery capacity, C_b, is calculated using Equation (15). Assuming cell capacity of 13 Ah, the number of cells in parallel are found to be 6. Thus, all the computation results are presented in Table 2.

$$E_b = E_{b_t} \times DOD \quad (14)$$

$$C_b = \frac{E_b}{72} \quad (15)$$

Table 2. Analytical modelling computation results

Parameter	Value
F_{rr}	28.06 N
F_{ad}	194.15N
a	0.00235 m/s^2
F_a	0.5165 N
F_g	0 N
F	222.72 N
T_w	48.11 Nm
N_w	958.02 rpm
T_m	8.27 Nm
N_m	5575.70 rpm
P_m	4.83kW
t	1.28 h
E_{b_t}	6.18 kWh
E_b	5.56kWh
C_b	77.26 Ah

Design of proposed simulation models

The proposed simulation model of retrofitted e-TW is designed in GT Suite and the model is represented in Figure 2. This simulation model consists of all major components interconnected together to form an EV bike model. A brief discussion of all the necessary components of the designed model is presented in this section.

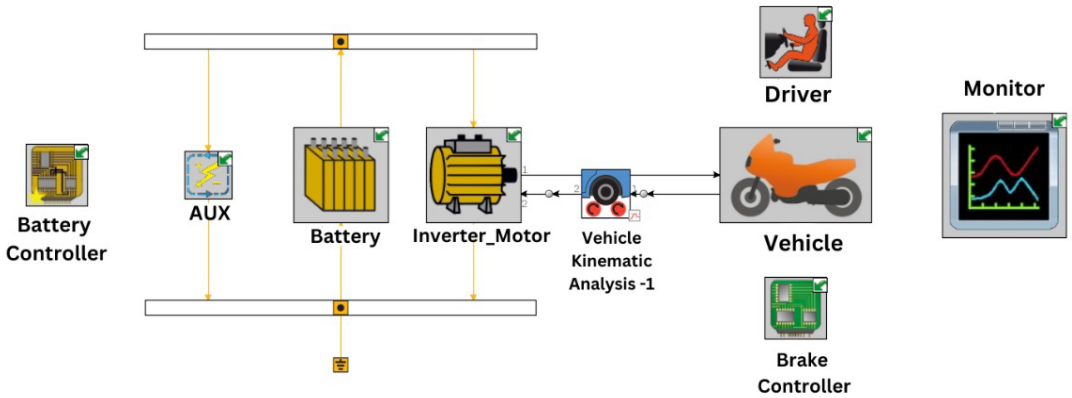


Figure 2: Designed Simulation model using GT-Suite

Battery and Battery Controller.

This model has used a $72V/78Ah$ Nickel Manganese Cobalt (NMC) Lithium-ion battery. They offer a good balance of energy density, power density, and thermal stability, making them popular choices for EV applications [6-7].

The control system in an EV bike manages various aspects of battery operation, including charging, discharging, thermal management, and state-of-charge (SoC) estimation. It also includes Battery Management System (BMS) components responsible for monitoring individual cell voltages, temperatures, and currents to ensure safe and efficient battery operation [8-9]. The BMS composes of specification like minimum discharge voltage limit of 27.5V, maximum charging voltage limit of 40.15V, and maximum C rate during discharge of 3.

Motor and Motor Controller.

The electric motor is the primary propulsion system of an EV bike, responsible for converting electrical energy from the battery into mechanical energy to drive the wheels. A 5 kW Brushless DC (BLDC) motor is used in this project due to its high efficiency, compact size, and low maintenance requirements [10-11].

The motor inverter, also known as the motor controller or power electronic unit, serves as the interface between the battery pack and the electric motor. It controls the speed, torque, and direction of the motor by regulating the frequency, amplitude, and phase of the electrical signals supplied to the motor windings.

Vehicle and Driver.

The GT-Suite modelling offers aerodynamic simulation capabilities to simulate airflow around the vehicle body. Thus, by optimizing aerodynamic shapes and reducing drag improves the energy efficiency of the vehicle. To achieve this, it allows to specify mass of the vehicle and driver, frontal area of the bike, gear ratio, tire radius, and drag coefficient values. The specifications mentioned in Table 1 are utilized in simulation modelling using GT Suite.

The modelling of driver is carried out using driver block. This block has inputs such as throttle, braking, steering, and manoeuvring through predefined control algorithms or user-defined functions. The designed model is tested using IDC in order to represent different driving conditions throughout the test duration.

Results and discussion

This section presents various tests performed to assess the performance of the designed model and their corresponding results are discussed. The GT-Suite platform has the facility of dashboard monitors, as shown in Figure 3.

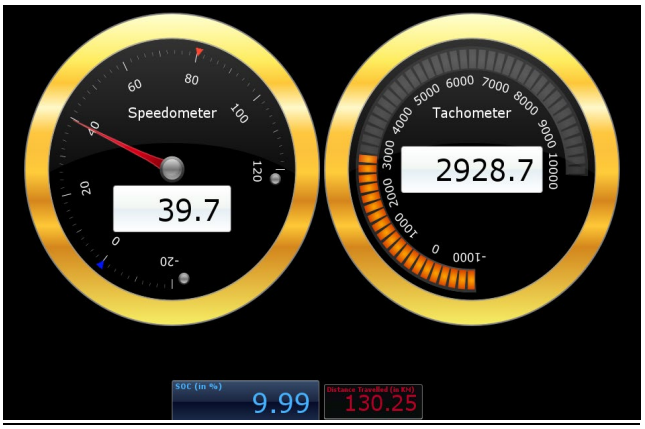


Figure 3. GT Suite dashboard monitor

The dashboard is pivotal in providing essential information to the driver and enriching the overall driving experience. In the designed simulation model, a user-friendly and informative dashboard interface is crafted. It showcases the state of charge (SoC) level, tachometer and speedometer readings, drive cycle, and distance travelled by the e-TW. The different conditions under which the designed model is tested are presented in four different case studies.

Case 1: Drive Cycle Test

The drive cycle test in GT-Suite serves as a valuable tool for engineers and researchers to analyse, optimize, and validate vehicle performance in a wide range of driving conditions. Thus, ultimately leading to the development of more efficient, cleaner, and safer vehicles.

The Standard Indian Drive Cycle is represented in Figure 4. This graph represents the standard driving pattern of an e-TW on Indian terrain.

The vehicle speed plot for drive cycle test is presented in Figure 5. The simulation results revealed valuable insights into the performance of the bike under the IDC conditions. It is observed that the electric bike demonstrated efficient energy utilization and satisfactory range coverage under the stop-and-go traffic conditions characteristic of the IDC. Additionally, the analysis of motor efficiency and battery SoC variations provided useful information for optimizing the electric bike's powertrain configuration and energy management strategies. In this test, it has been observed that the maximum speed attained is 42 km/h just as the standard IDC condition and the simulation repeats every 110 s. Comparison of Figure 4 and Figure 5 shows that the retrofitted bike followed the standard IDC during this test.

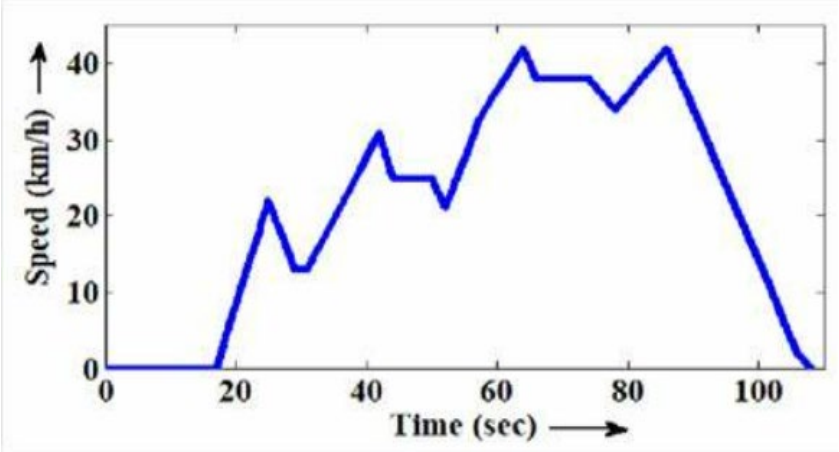


Figure 4. Standard Indian Drive Cycle

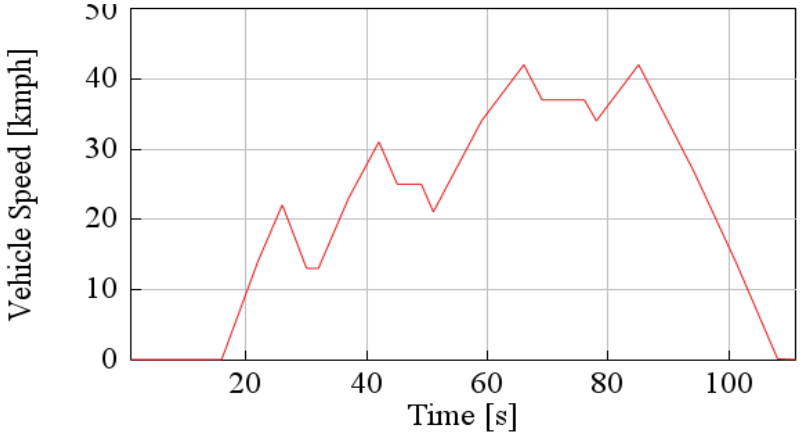


Figure 5. India drives cycle (IDC) actual response representation

Case 2: Maximum Speed Test

This test is conducted to assess the vehicle's performance capabilities and limitations under high-speed operating conditions. It also allows for the determination of the vehicle's maximum speed. The vehicle speed plot for maximum speed test is presented in Figure 6. The simulation results from the maximum speed test revealed the maximum achievable speed of the bike is $80.44\ kmph$. Additionally, the analysis of battery current ($88.17\ A$) and power consumption ($5.72\ kW$) provided insights into the thermal constraints and potential overheating risks associated with prolonged high-speed operation.

The motor speed plot, obtained using Tachometer, is presented in Figure 7. Further, the simulation results from the motor maximum speed case revealed valuable insights into the brake torque ($6586\ Nm$), efficiency (97%), and motor losses ($0.42\ kW$). The upper limit of the motor is found to be $5927\ rpm$. In addition, a plot depicting the maximum motor torque was acquired. The maximum wheel torque plot, presented in Figure 8, illustrates the results of the maximum wheel torque test, revealing a peak torque of 55.79 Nm achievable by the wheel. Obtaining the wheel torque facilitates the calculation of motor torque using the gear ratio and motor efficiency.

Torque reaches its maximum when speed is zero, as the vehicle is stationary and requires significant torque to overcome rolling resistance. Initially, tractive effort demands substantial torque. However, as speed increases, torque decreases to 45 Nm, as most power is directed towards reaching maximum speed, thereby diminishing torque.

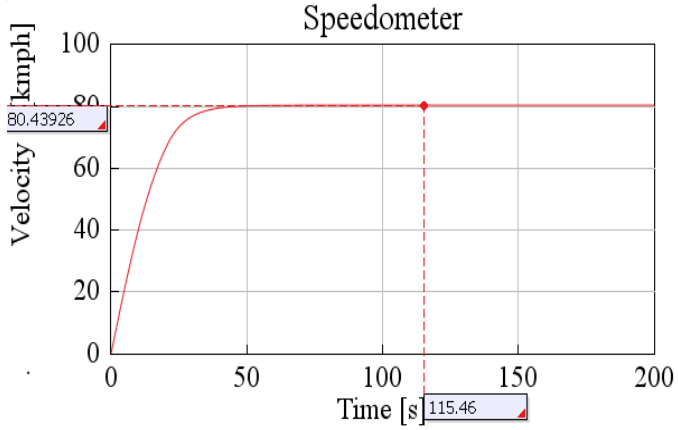


Figure 6. Speedometer result indicating peak speed attained by vehicle

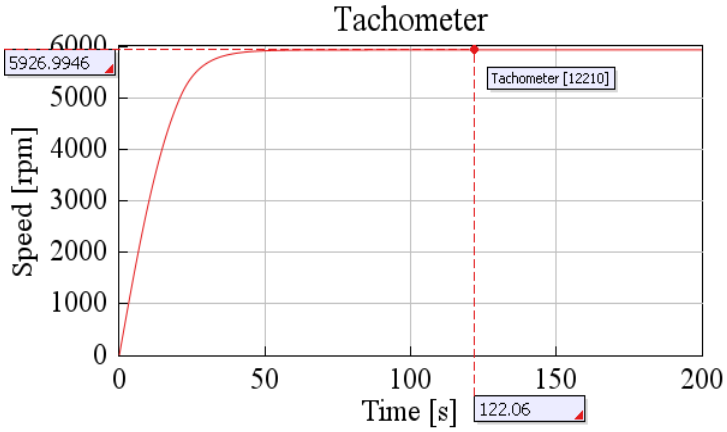


Figure 7. Tachometer result indicating peak speed attained by motor

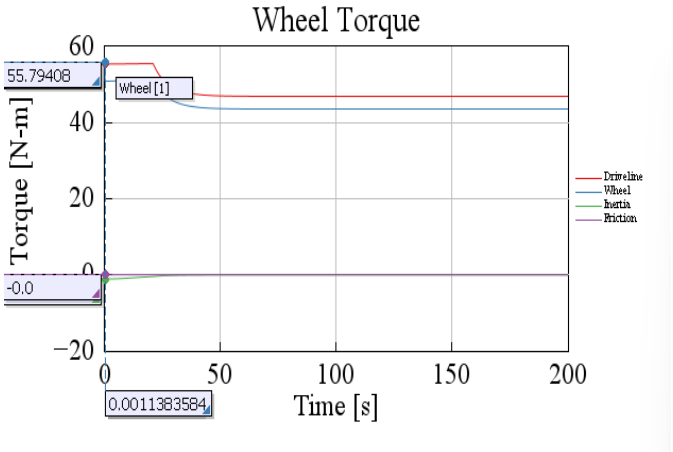


Figure 8. Analysis results indicating torque attained by the wheel

Case 3: Acceleration Test

The acceleration test serves to optimize powertrain performance, evaluate vehicle acceleration, and validate simulation models for the development of EV.

The acceleration plot is presented in Figure 9. The simulation results from the acceleration test revealed the acceleration behaviour of the bike ($8.56\ m/s^2$). The vehicle's acceleration profile, time to reach specific speed ($10\ s\ at\ 40.4\ km/h$), and wheel torque output characteristics ($54.1\ Nm$) during acceleration manoeuvres are observed. Additionally, the analysis of energy consumption provided insights into the powertrain efficiency (97 %) and energy utilization ($3.36\ kW$) during acceleration.

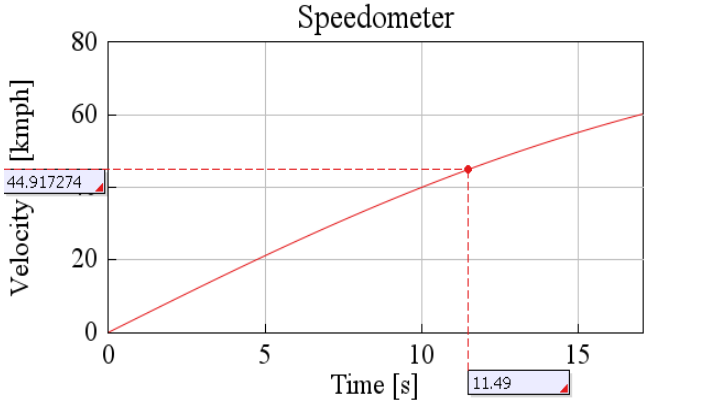


Figure 9. Speedometer result indicating the vehicle acceleration

Case 4: Vehicle's Range Test

The vehicle range test conducted plays a crucial role in evaluating, optimizing, and validating the range of EVs highlighting its overall performance. It also allows to acquire the knowledge of the total distance covered on a single charge of battery.

The displacement plot with respect to time is presented in Figure 10. The simulation results from the vehicle's range test provided valuable insights into the range capabilities and energy consumption patterns of the bike. The vehicle's range is found to be $130.25\ km$, with observed maximum current consumed ($32.5A$), battery terminal power ($2.05\ kW$), battery SOC ($100\% - 10\%$), and maximum driving speed ($42.1\ kmph$).

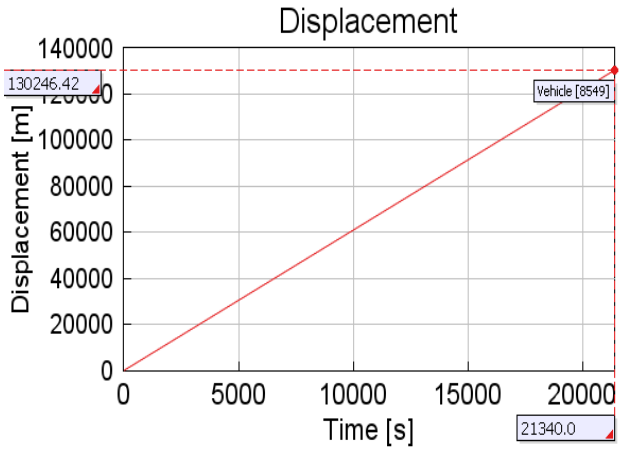


Figure 10. Analysis results depicting vehicle's range

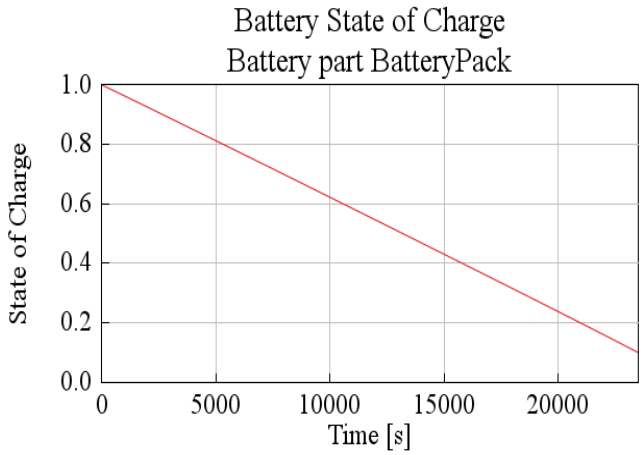


Figure 11. Battery State of Charge (SoC) during vehicle's range test

The battery's SoC behaviour is shown in Figure 11 during the IDC Vehicle Range test. During this test, battery discharges from 100 % to 10 %, indicating a depth of discharge of 90 %. A decline in battery SoC is observed during different driving phases, such as acceleration, cruising, and idling.

Comparative analysis of vehicle performance

This section is dedicated to the discussion regarding the comparative analysis results, obtained from analytical modelling and designed simulation modelling. The compilation of results is presented in Table 3.

Table 3: Comparative analysis results ($\boldsymbol{p.u.}$ means per unit)

Characteristic of e-TW	Analytical modelling	Simulation modelling	Error
Wheel Torque	48.11Nm	55.79 Nm	$-0.14\ p.u.$
Motor Torque	8.27Nm	9.59 Nm	$-0.16\ p.u.$
Maximum Speed	78km/h	80.44 km/h	$-0.03\ p.u.$
Maximum Range	100 km	130.25 km	$-0.30\ p.u.$
Wheel Speed	958.02 rpm	982.78 rpm	$-0.02\ p.u.$
Motor Speed	5575.7 rpm	5927.00 rpm	$-0.06\ p.u.$

It has been observed that the simulation model yields slightly higher values compared to the analytical model. Considering the small error (< -0.3) across all comparative analysis results, it can be demonstrated that the GT Suite platform is highly accurate.

Materials Research Proceedings 49 (2025) 186-196 https://doi.org/10.21741/9781644903438-19

Further, comparison of performance of ICE two-wheeler is performed with two types of retrofitted vehicles. The comparative results are presented in Table 4. It depicts the various specifications of the ICE vehicle, specifications of retrofitted model with 1.5 kW hub motor, and specifications of retrofitted model with 5 kW BLDC motor (both designed by GT suite).

Table 4. Comparison of ICE vehicle with two retrofitted e-TW

Attribute	**ICE vehicle**	**Retrofitted model (1. 5 kW Hub motor)**	**Retrofitted model (5 kW BLDC motor)**
Battery Capacity	$12V, 5Ah$	$48V, 28Ah$	$72V, 76Ah$
Maximum Velocity	85 km/h	52 km/h	$80.44km/h$
Range	60 km/l	29 km	130.25 km
Maximum Torque	7.85 Nm	22.5 Nm	9.59 Nm
Maximum Speed	5000 rpm	636.6 rpm	5927.00 rpm
Running Cost	High	Low	Low
Emissions	High	None	None
Rank	**3**	**2**	**1**

In Table 4, the ranks are awarded to the vehicles in accordance to their performances. It is observed that retrofitted model with 5 kW BLDC motor, has the maximum vehicle range with one charge of the battery and it is also capable to attain the maximum speed of 80.44 $kmph$, which is in agreement with the ICE vehicle. Also, this retrofitted model has increased torque capabilities. Hence, rank 1 is awarded to this retrofitted model. Similarly, the comparison shows that the retrofitted model with low power rating of 1.5 kW Hub motor secures Rank 2, since it also saves emissions and provide 22.5 Nm of maximum torque. However, it does not attain commendable displacement range and maximum speed. Therefore, it lacks in the vehicle's overall performance, as compared to retrofitted model with 5 kW BLDC motor. Further, ICE vehicle secures rank 3 because it offers high emissions, which are less environmentally friendly. Thus, retrofitting of ICE vehicles to EVs with suitable components are proved to be the better solution for sustainable, more favourable and less cost-intensive transport systems.

Conclusion

This study effectively evaluates motor power, battery capacity, and vehicle performance metrics using analytical method as well as simulation modelling. The analysis and simulation methodology highlight the potential of GT Suite in EV design and optimization. The analysis results showcase a significant improvement in speed and range with the upgraded motor and battery configuration. Ultimately, this study contributes following valuable insights into the practicality and benefits of transitioning from ICE vehicles to electric alternatives, offering a promising pathway for sustainable transportation solutions:

- The analytical modelling is successfully carried out using practical values of vehicle dynamics. This helps in recognizing the understanding of motor power and battery capacity computation for knowledge enhancement in real time scenario of the road condition.
- Battery pack calculation is successfully carried out for understanding of the series or parallel combination of cells by implementation of mathematical modelling. These values of combination of battery cells (series and/or parallel) are utilized in simulation models.
- The simulation model of retrofitted e-TW is designed having 5 kW BLDC motor using highly accurate GT Suite platform. This platform helped in analysing four significant case studies, namely, drive cycle test (using IDC), maximum speed test, vehicle range test and acceleration test.

- Using comparative analysis results, it is found that the simulated model shows negligible errors in motor speed, wheel speed, maximum speed and maximum range. However, there are removable errors in wheel torque and motor torque, if few parameters are altered during analytical modelling.

Acknowledgment
This work has been supported by Centre of Excellence for Electric Vehicles and Related Technologies, Electrical Engineering Department, Delhi Technological University, India.

References
[1] Kumar, N. A., Navaneeth, M., & Joseph, A. S. (2021). Retrofitting of Conventional Two-wheelers to Electric Two-Wheelers. Asia-Pacific Power and Energy Engineering Conference, APPEEC, 2021-November. https://doi.org/10.1109/APPEEC50844.2021.9687799
[2] S, A. and Manjunatha, R. (2022) 'Design and modelling of Ultra-Capacitor based hybrid energy storage system in electric two-wheeler for Indian driving cycle', 2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE) [Preprint]. https://doi.org/10.1109/icdcece53908.2022.9793082.
[3] Friedl, H. et al. (2018) 'Range Extender Technology for Electric vehicles', 2018 5th International Conference on Electric Vehicular Technology (ICEVT) [Preprint]. https://doi.org/10.1109/icevt.2018.8628391.
[4] Song, Y. et al. (2019) 'Improved SOC estimation of lithium-ion batteries with novel SOC-OCV curve estimation method using equivalent circuit model', 2019 4th International Conference on Smart and Sustainable Technologies (SpliTech) [Preprint]. https://doi.org/10.23919/splitech.2019.8783149.
[5] Muruganandhan, D., Muthunagai, R., Balaji, D., Prasanna Ganesh, K., & Vishnu, S. (2023). Simulation Model Development for Conversion of IC Engine Vehicle to Electric Vehicle. 2023 International Conference on System, Computation, Automation and Networking, ICSCAN 2023. https://doi.org/10.1109/ICSCAN58655.2023.10395002
[6] Arifwardana, J.F. et al. (2022) 'Experimental of heat generation on electric motorcycle NMC and LiFePO4 Battery Pack', 2022 7th International Conference on Electric Vehicular Technology (ICEVT) [Preprint]. https://doi.org/10.1109/icevt55516.2022.9924981.
[7] Corno, M. et al. (2015) 'Electrochemical model-based state of charge estimation for Li-Ion Cells', IEEE Transactions on Control Systems Technology, 23(1), pp. 117–127. https://doi.org/10.1109/tcst.2014.2314333.
[8] Song, Y. et al. (2019) 'Improved SOC estimation of lithium-ion batteries with novel SOC-OCV curve estimation method using equivalent circuit model', 2019 4th International Conference on Smart and Sustainable Technologies (SpliTech) [Preprint]. https://doi.org/10.23919/splitech.2019.8783149.
[9] H. He, R. Xiong, X. Zhang, F. Sun, J. Fan, "State-of-Charge Estimation of the Lithium-Ion Battery Using an Adaptive Extended Kalman Filter Based on an Improved Thevenin Model," IEEE Transactions on Vehicular Technology, vol. 60(4), pp. 1461-1469, 2011.
[10] Sudhir Gupte , A.D.Patel Institute of Technology , Vidyanagar published a technical paper on "experimental analysis and feasibility study of 1400 cc diesel engine car converted into hybrid electric vehicle by using BLDC hub motors" in 4th International Conference on Advances in Energy Research 2013(ICAER2013).
[11] Kulkarni, A. R., Singh, S., Kant, K., & Bharti, H. (2020). Experimental Electric Retrofitting of an Ice Vehicle with Simulation and Cost Analysis. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 9(9S), 2278-3075.

Microchemical Journal 213 (2025) 113681



Contents lists available at ScienceDirect

Microchemical Journal

journal homepage: www.elsevier.com/locate/microc



Photocatalytic degradation of crystal violet dye using phytofabricated zinc oxide nanoparticles from anthocyanin extract

Sarla Yadav, Pooja Singh, Gunjan Varshney, Saurabh, Poonam Singh, Raminder Kaur *

Department of Applied Chemistry, Delhi Technological University (Formerly Delhi College of Engineering), Delhi 110042, India

ARTICLE INFO

Keywords:
Phyto-fabrication
Zinc oxide nanoparticles
Anthocyanin extract
Photocatalysis
Crystal violet dye
Water treatment

ABSTRACT

This study investigates a coprecipitation method for the production of zinc oxide nanoparticles (ZnO NPs) using anthocyanin (ATH) extract derived from red cabbage (RC) that acts as a reducing, capping and stabilising agent. The efficacy of these NPs in the photocatalytic degradation of crystal violet (CV) dye is assessed thoroughly considering all the relevant influencing parameters such as pH levels, catalytic dosage, and dye concentration, among others. The structural, morphological, surface charge and presence of different functional groups of the synthesized NPs were then analysed using XRD, SEM-EDX, BET, and FTIR. The average particle size of the synthesized ZnO NPs and ZnATH-(O) NPs was 31.12 nm and 27.91 nm, respectively. The BET analysis determined that the surface area of ZnO NPs was 15.685 m^2/ g, and that of ZnATH-(O) NPs was 25.187 m^2/ g. In contrast to ZnO NPs, ZnATH-(O) NPs exhibit efficient photocatalytic degradation of CV dye. This is due to their improved stability, reduced size, increased surface area and optimized structure. The influence of various parameters like initial dye concentration, catalyst dosage, pH and ionic strength was also assessed under 90 min visible light irradiation. It was observed that ZnATH-(O) NPs show maximum photocatalytic degradation of CV dye in 90 min of visible light irradiation with a degradation efficiency of 96.12 %. This green synthesis method mitigates the environmental consequences associated with conventional chemical synthesis and utilizes the plentiful and renewable resources present in RC, thereby offering a sustainable alternative to a deteriorating environment.

1. Introduction

The rapid industrialization and urbanization of recent decades have led to a significant increase in the release of synthetic dyes into aquatic environments [1–3]. Among these, crystal violet (CV) dye, which is frequently used in printing, textiles, and other sectors, presents a significant risk because of its high toxicity, persistence, and possible carcinogenicity [4–6]. Conventional techniques used to treat water contaminated with dye, such as coagulation [7], flocculation and adsorption frequently lack effectiveness, cost-efficiency, and environmental sustainability [8]. Therefore, the quest for inventive and environmentally friendly methods to reduce the detrimental impacts of these contaminants has become a crucial area of attention in environmental research.

Photocatalysis, capable of harnessing light to activate semiconductor materials that can degrade organic pollutants, emerges as a promising solution [9,10]. Zinc oxide nanoparticles (ZnO NPs) have gathered considerable interest in this field because of their distinctive optical and electrical characteristics, such as a broad energy gap (3.37 eV), a high exciton binding energy (60 meV), and a robust photocatalytic activity [5,11]. These properties enable ZnO NPs to generate reactive oxygen species (ROS) such as hydroxyl radicals ($OH^{\cdot}$) and superoxide (O_2^{-}) anions under UV irradiation, which can effectively degrade complex organic pollutants, including dye molecules, into less degraded entities like CO_3^{2-}, NO_3^{2-} and H_2O [12]. Moreover, ZnO NPs are biocompatible, easy to synthesise, and non-toxic, making them an ecologically benign choice for water treatment and other applications [13,14].

Conventional methods adopted to synthesize ZnO NPs frequently employ toxic compounds like sodium hydroxide, zinc nitrate, and other organic solvents. Additionally, these approaches need energy-intensive procedures such as sol–gel, hydrothermal, and chemical vapour deposition (CVD) techniques [15]. These technologies not only give rise to environmental and safety concerns but also provide obstacles in achieving constant NPs size and morphology, which are crucial for achieving optimal photocatalytic performance. As an outcome of these challenges, the field of green synthesis has gained prominence,

* Corresponding author.
E-mail address: raminderkaur@dce.ac.in (R. Kaur).

https://doi.org/10.1016/j.microc.2025.113681
Received 5 February 2025; Received in revised form 29 March 2025; Accepted 16 April 2025
Available online 17 April 2025
0026-265X/© 2025 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

emphasising the use of natural, renewable resources and environmentally benign processes [16]. Among various natural sources, plant extracts rich in bioactive compounds, such as anthocyanins (ATH), have demonstrated substantial potential in the biosynthesis of NPs [5]. ATH found abundantly in red cabbage, perform dual functions as reducing and capping agents, enhancing the NPs' stability and functionality through their antioxidant properties [17]. The mechanism involves the reduction of zinc ions (Zn^{2+}) to zinc oxide (ZnO) facilitated by the phytochemicals present in the extract, leading to the formation of NPs with controlled size and shape.

Red cabbage (*Brassica oleracea* var. *capitata f. rubra*) (RC) is highly beneficial for this purpose because it contains a significant amount of ATH, is inexpensive and is easily accessible [18]. ATH extract is derived from RC, which is rich in micronutrients and some phytochemicals, such as Vitamin C and K, β-carotene, minerals, fibres, total polyphenols and glucosinolates. ATHs are the glycosylated forms of anthocyanidins (aglycones) [19]. The key functional groups present in ATH extract include hydroxyl (–OH) and methoxy ($–OCH_3$) groups attached to the phenolic rings, which significantly contribute to their antioxidant properties. Additionally, ATH contains glycosidic bonds where sugars like glucose or galactose are linked to the aglycones, enhancing their solubility and stability in aqueous environments. The presence of carbonyl (C=O) groups in the core structure further contributes to the electron structure [20]. ATH extracted from RC possesses potent reducing characteristics and the capacity to chelate metal ions, rendering them exceptionally efficient in the development of NPs. In addition, the utilization of RC extract aligns well with the principles of green chemistry, as it eliminates the requirement for dangerous chemicals and diminishes the total ecological impact of the synthesis process. ATH extracted from RC is highly versatile and effective, providing them with the possible potential for synthesizing ZnO NPs with improved photocatalytic characteristics.

This research presents a groundbreaking approach to synthesize ZnO NPs using ATH extract derived from RC for photocatalytic degradation of CV dye. While previous studies predominantly utilized thermal routes for synthesising ZnO NPs from ATH, this study introduces a novel method: coprecipitation using NaOH as the co-precipitating agent. Furthermore, the study aims to compare the performance of these bio-derived ZnO NPs (ZnATH-(O) NPs) with those synthesized via traditional methods [21]. The methodology involves optimizing the synthesis parameters such as pH, catalytic dosage, and concentration of the dye solution to achieve desirable catalytic properties for maximum degradation. The photocatalytic activity was assessed by monitoring the degradation of CV dye under visible light irradiation, measuring the change in absorbance over reaction time to determine the degradation rate constants. The significance of this work lies is two-fold: it advances the fundamental understanding of green nanotechnology and proposes practical solutions for environmental remediation. Through comparative analysis, it highlights the advantages of these ZnATH-(O) NPs, paving the way for future innovations in the fields of photocatalysis and environmental science.

2. Materials and methods

2.1. Materials

Red cabbage (*Brassica oleracea* var. *capitata*) (RC) was procured from the local market in Delhi, India, for ATH extraction. Hydrochloric acid (HCl), Anhydrous Ethanol (EtOH) and sodium hydroxide (NaOH) were acquired from Thermofischer chemicals. Zinc acetate dihydrate [$Zn(CH_3COO)_2{\cdot}2H_2O$] was obtained from S.D Fine Chemical Limited. All the experiments were conducted using distilled water.

2.2. Synthesis

2.2.1. Extraction of anthocyanin (ATH) from RC

ATH was extracted from RC (*Brassica oleracea* var. *capitata*) by employing a similar methodology as reported in the literature [22–24], using pure EtOH as a solvent with modifications as shown in Fig. 1. The procedure involved cutting and washing of fresh RC using distilled water. Then, excess moisture was removed using a muslin cloth. 100 g of RC were pulverised and mixed with 150 ml of pure EtOH solvent in an Erlenmeyer flask for 36 h at 4 °C to extract phytochemicals. Subsequently, the mixture was filtered, and the resulting filtrate was centrifugated at 4000 rpm for 10 min for further purification. Afterwards, the pH of the mixture was first maintained at 2, using HCl. Then, the mixture was kept in the dark at 4 °C for 36 h. Afterwards, the pH of the solution was adjusted to 7 by adding NaOH, and it was further kept at a temperature of 4 °C for future use.

2.2.2. Synthesis of zinc oxide nanoparticles (ZnO NPs) without ATH extract

A cost-effective co-precipitation method was employed to synthesize ZnO NPs from the precursor $Zn(CH_3COO)_2{\cdot}2H_2O$ and NaOH, as shown in Fig. 2. 1 g of $Zn(CH_3COO)_2{\cdot}2H_2O$ was added to a beaker containing 50 ml of distilled water during the synthesis process, and it was constantly agitated for 15 min at room temperature. Additionally, the pH of the mixture was kept at 10 by gradually adding 1 M NaOH solution, and the resulting mixture was then stirred at 450 rpm for 4 h in an ambient condition. Finally, the white-coloured precipitate formed was filtered and rinsed with 1:1 EtOH solvent and distilled water. Then, the product was dried at 60 °C in an oven for 24 h. Afterwards, the dried product was crushed into a fine powder using a mortar and pestle and then calcined for two hours at 450 °C in a muffle furnace.

2.2.3. Phyto-fabrication of ZnO NPs with ATH extract (ZnATH-(O) NPs)

The phyto-fabrication of ZnO NPs was carried out using ATH extract, depicted in Fig. 3. Firstly, 1 g of $Zn(CH_3COO)_2{\cdot}2H_2O$ was dissolved in 50 ml distilled water and stirred for 15 min at room temperature. Afterwards, 5 ml of extracted ATH solution was added and stirred for the next 30 min. At this point, the colour of the solution turned orange. Additionally, a 1 M NaOH solution was added to the solution in a dropwise manner to maintain its pH at 10, and the resulting solution was then agitated further for 4 h. The resulting white precipitate was then filtered and rinsed with distilled water and EtOH solvent in a 1:1 ratio. The isolated product underwent a 24 h drying process at 60 °C, followed by a 2 h calcination in a muffle furnace at 450 °C. The probable mechanism for the synthesis of ZnO NPs using ATH extract is shown in Fig. 4.

2.3. Characterization techniques

A powder X-ray diffractometer (XRD) using CuKα radiation (Rigaku Ultima IV model) was employed for the crystallographic analysis of the synthesized NPs. The XRD patterns were recorded within a 2θ range of 10 °C–80 °C, with a step size of 0.02° and a scanning rate of 1.0 s/step, at 298 K. A Perkin Elmer BX-II Fourier-transform infrared spectrometer equipped with a KBr disk was utilized to investigate the functional groups present in the prepared NPs within the range of 500–4000 cm^{-1}. The produced NPs were further analyzed using SEM-EDX analysis with Zeiss EVO-18 instrument. The nitrogen adsorption–desorption approach was employed to achieve the average dispersion of pore sizes and specific surface area (S_{BET}) of NPs at 77 K, which was examined by the Quanta chromium Nova-1000 equipment. The Zeta potential of the synthesized NPs was determined by the Malvern Zeta-Sizer Nano series meter (Version 7.13). The photocatalytic degradation of CV dye was assessed using a WinCary 300 UV–vis spectrophotometer at the wavelength of maximum absorption (λ_{max}) of 580 nm.

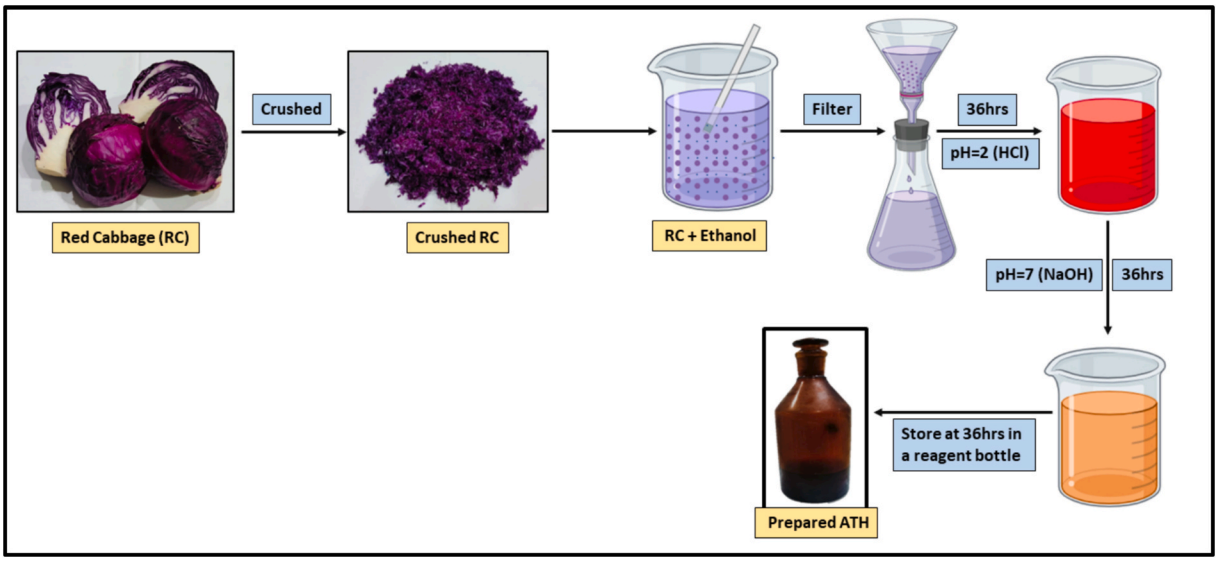


Fig. 1. Extraction of ATH from RC (*Brassica oleracea var. capitata*).

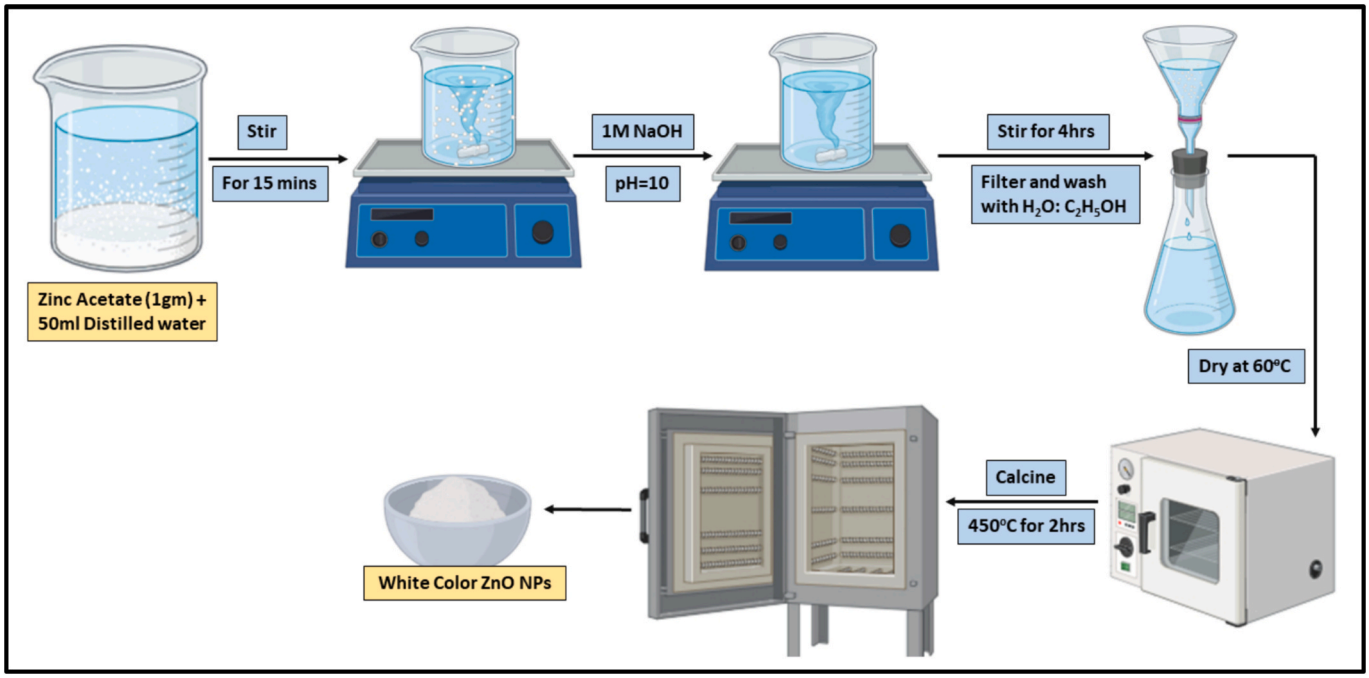


Fig. 2. Schematic representation of the synthesis of ZnO NPs.

2.4. Evaluation of photocatalytic activity

The degrading capacity of NPs towards azo dyes was evaluated for CV dye. Firstly, a predetermined quantity of CV dye was used to generate a stock solution with a 100 mg/L concentration. Standard solutions of varying concentrations (10–50 mg/L) were prepared by diluting 100 mg/L stock solution of CV dye. A cylindrical batch reactor with an annular shape was employed to carry out the photocatalysis of CV dye. The reactor was illuminated with a 125 W high-pressure xenon bulb. The synthesized NPs were used to photodegrade CV dye solution in the presence of a visible light source to assess its photocatalytic activity. Before irradiation, a 30 min adsorption of CV dye was performed using NPs synthesized via both methodologies (ZnO NPs and ZnATH-(O) NPs), was performed to establish adsorption–desorption equilibrium between the pollutant (CV dye) and photocatalyst. Subsequently, the CV dye solution was subjected to a light source exposure in the presence of ZnATH-(O) NPs while being continuously stirred for 90 min at 450 rpm. At a fixed time interval of 10 min, 3 mL samples were taken and then centrifuged for 5 min at 4000 rpm. The degradation percentage was determined by applying Eq. (1).

$$\boldsymbol{Percentagedegradation}(\%) = \left(\frac{\boldsymbol{C_i} - \boldsymbol{C}}{\boldsymbol{C_i}}\right) \times 100 \quad (1)$$

The variable $\boldsymbol{C}$ represents the CV dye solution concentration in the reaction mixture containing photocatalyst after time t. On the other hand, $\boldsymbol{C_i}$ represents the initial concentration of CV without any catalyst. After analysis, the photocatalytic degradation of CV dye was found to follow pseudo-first-order kinetics, as shown in Eq. (2).

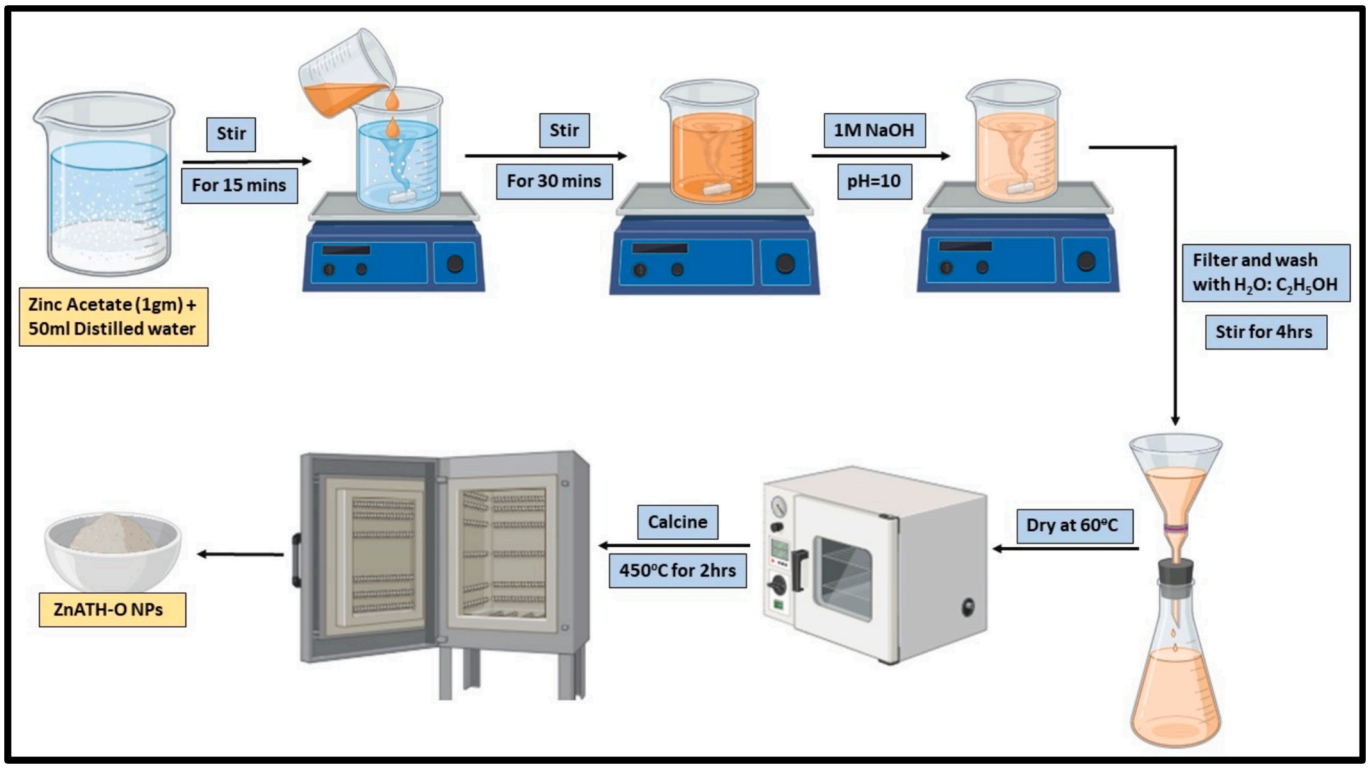


Fig. 3. Systematic methodology for the preparation of ZnATH-(O) NPs.

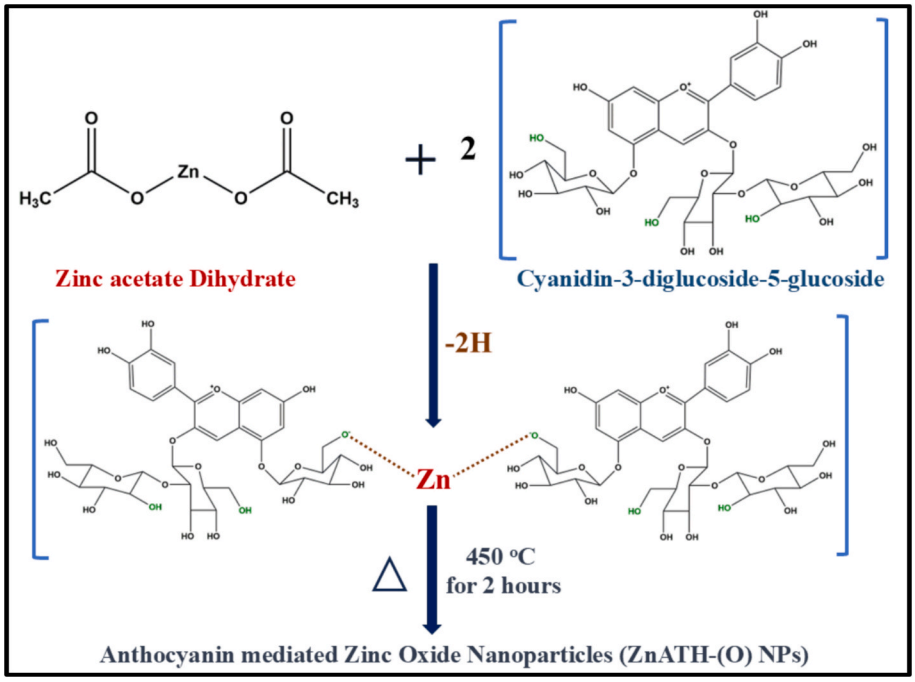


Fig. 4. Probable mechanism of green synthesis of ZnO NPs using ATH leaf extract. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$$ln\left(\frac{C}{C_i}\right) = Kt \quad (2)$$

Here, C represents the final concentration of CV dye at time t, C_i denotes the initial dye concentration at t = 0, and K represents the rate constant.

The Beer-Lambert law (Eq. (3)) was employed to create a calibration plot that correlates absorbance with concentration to ascertain the unidentified concentration of the dye.

$$A = \varepsilon Cl \quad (3)$$

The reaction kinetics of the process were quantified in terms of photocatalytic degradation as mentioned in Eq. (4).

$$\ln\left(\frac{C}{C_i}\right) = ln\left(\frac{A_2}{A_1}\right) = kt \quad (4)$$

where A_1 and A_2 denotes the initial and final absorbance of CV dye. The impact of various influencing variables such as reaction time, dye concentration, dose of catalyst, and pH of dye solution on catalytic degradation of CV dye using ZnATH-(O) NPs were also measured. In addition, the influence of the ionic strength of NaCl on the degradation efficiency, degradation of CV dye in real water samples and the recyclability of ZnATH-(O) NPs as photocatalyst were also examined.

3. Result and discussion

3.1. Characterization of ZnO NPs and ZnATH-(O) NPs

The comparative XRD patterns of ZnO NPs and ZnATH-(O) NPs are shown in Fig. 5 a. The ZnATH-(O) NPs were found to imitate the hexagonal wurtzite ZnO structure described in the literature [25,26]. The ZnO NPs exhibit distinct and strong diffraction peaks at 2θ = 31.72°, 34.36°, and 36.2°, which correspond to the (1 0 0), (0 0 2), and (1 0 1) crystallographic planes. The d-spacing values for these planes are 2.81 Å, 2.60 Å, and 2.47 Å, respectively. As, these values are consistent with those reported in the literature and match the standard JCPDS file (JCPDS No. 01-080-0074) for ZnO.

The average size of the crystallites was determined using Debye Scherrer's equation:

$$D = \frac{k^*\lambda}{\beta^* cos\theta} \quad (5)$$

where D represents the diameter of the crystallite, k is the Scherrer constant, λ represents the wavelength of the incident X-ray (0.15418 nm), β refers to the full width at half maximum (FWHM) of peaks, and θ represents the diffraction angle. The mean crystallite size of ZnO NPs was determined to be 31.12 nm, using Debye Scherrer's equation.

The XRD pattern of ZnATH-(O) NPs has an identical structure with that of ZnO NPs and shows a minor shift in peak positions, which can be attributed to the incorporation of ATH extract into the Zn-(O) matrix, confirming the dual function of ATH extract as reducing and capping agent. The mean crystallite size of ZnATH-(O) NPs was determined to be 27.91 nm, which corroborates the successful reduction in size of NPs synthesized using ATH extract as compared to that of ZnO NPs synthesized via the conventional co-precipitation method without ATH extract [27]. However, despite this incorporation, the overall crystal planes (1 0 0), (0 0 2), and (1 0 1) remain dominant, indicating that the hexagonal wurtzite structure of ZnO is still retained, as these values were compared with the JCPDS file (JCPDS No. 01–076-0704).

The synthesized NPs were further analysed using FTIR spectroscopy, which was recorded in the range of 500–4000 cm^{-1}, as shown in Fig. 5 b. The ZnO NPs display a wide and strong IR band at 3488 cm^{-1}, corresponding to the stretching vibrations of O–H bonds caused by adsorbed water molecules on the surface [5]. In addition, the FTIR spectra of ZnO NPs display a distinctive peak at 648 cm^{-1}, which signifies the stretching vibration of metal–oxygen bonds (M−O) in ZnO NPs.

However, in the case of ZnATH-(O) NPs, additional peaks were observed at 1445 cm^{-1}, 1144 cm^{-1}, and 1053 cm^{-1}. These peaks are attributed to the stretching of C=O bonds, the presence of –COO groups, and the bending vibrations of C–H bonds in the organic moieties found in ATH extract, respectively. The characteristic peaks at 2987 cm^{-1} and 2899 cm^{-1} are due to C–H and CH_2 stretching vibration. Additionally, a broad and strong band appeared at 3467 cm^{-1}, corresponding to the stretching vibration of O–H bonds [28]. Furthermore, the decrease in the wavenumber of the M−O linkage band from 648 cm^{-1} to 604 cm^{-1} is due to the addition of ATH extract, which acts as a reducing, capping and stabilising agent, as reported in the literature [29,30]. The incorporation of ATH extract onto the ZnO matrix causes the shift in M−O vibration bands to lower wavelength, thereby decreasing the bond strength and further leading to the formation of new bonds of ZnO with the phytochemicals present in ATH extract.

The SEM analysis was carried out to examine the surface morphology and to calculate the average particle size of ZnO NPs and ZnATH-(O) NPs using Image J software, as shown in Fig. 6 a and c. The SEM image of ZnO NPs reveals a morphology resembling plates, and the particles' average size was approximately 29.54 nm [5]. Meanwhile, ZnATH-(O) NPs display a less agglomerated uneven shape, with an average particle size of approximately 21.43 nm [31]. The shift from a plate-like morphology to an irregular shape could be attributed to the introduction of ATH extract, which functions as a capping and reducing agent. This inference is supported by the observed decrease in the size of NPs, as mentioned in the XRD analysis (Table 1). EDX analysis was conducted to analyse the elemental composition of ZnO NPs and ZnATH-(O) NPs, as depicted in Figs. 6b and d. The EDX analysis confirmed that all the elements in the lattice are uniformly distributed.

The BET approach was utilised to estimate the specific surface area (S_{BET}), pore volume, and average pore size of the produced NPs (Table 2). Fig. 7 a shows the nitrogen adsorption–desorption isotherm of the produced ZnATH-(O) NPs. The S_{BET} values for ZnO NPs and ZnATH-(O) NPs were determined to be 15.685 m^2/g and 25.187 m^2/g, respectively. The pore volume of the synthesized ZnO NPs and ZnATH-(O) NPs was determined using the BJH (Barret-Joyner-Halenda) method. The pore volume was found to be 0.159 cc/g for ZnO NPs and 0.224 cc/g for

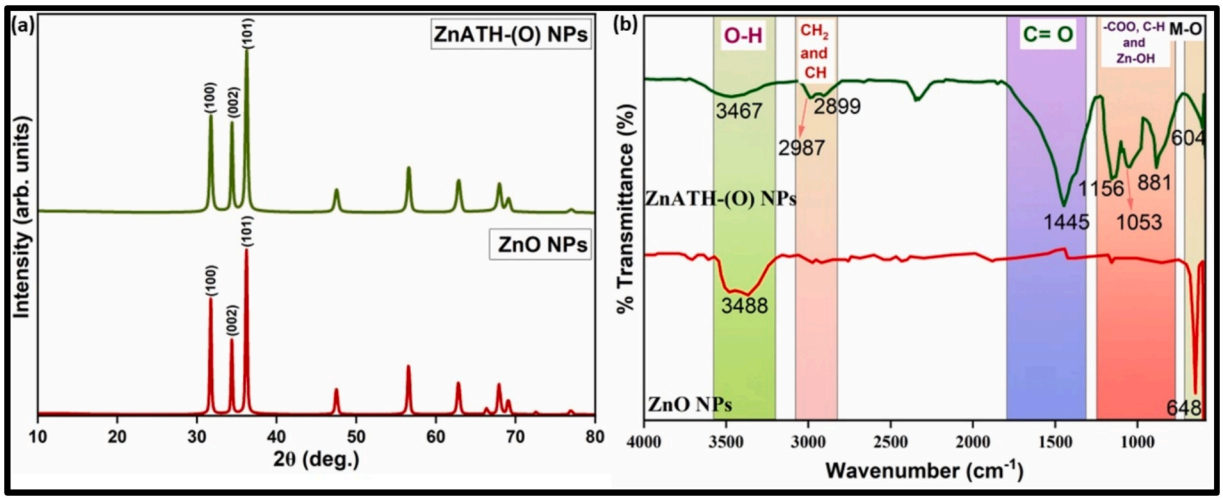


Fig. 5. Comparative (a) XRD pattern and (b) FTIR spectrum of ZnO NPs and ZnATH-(O) NPs.

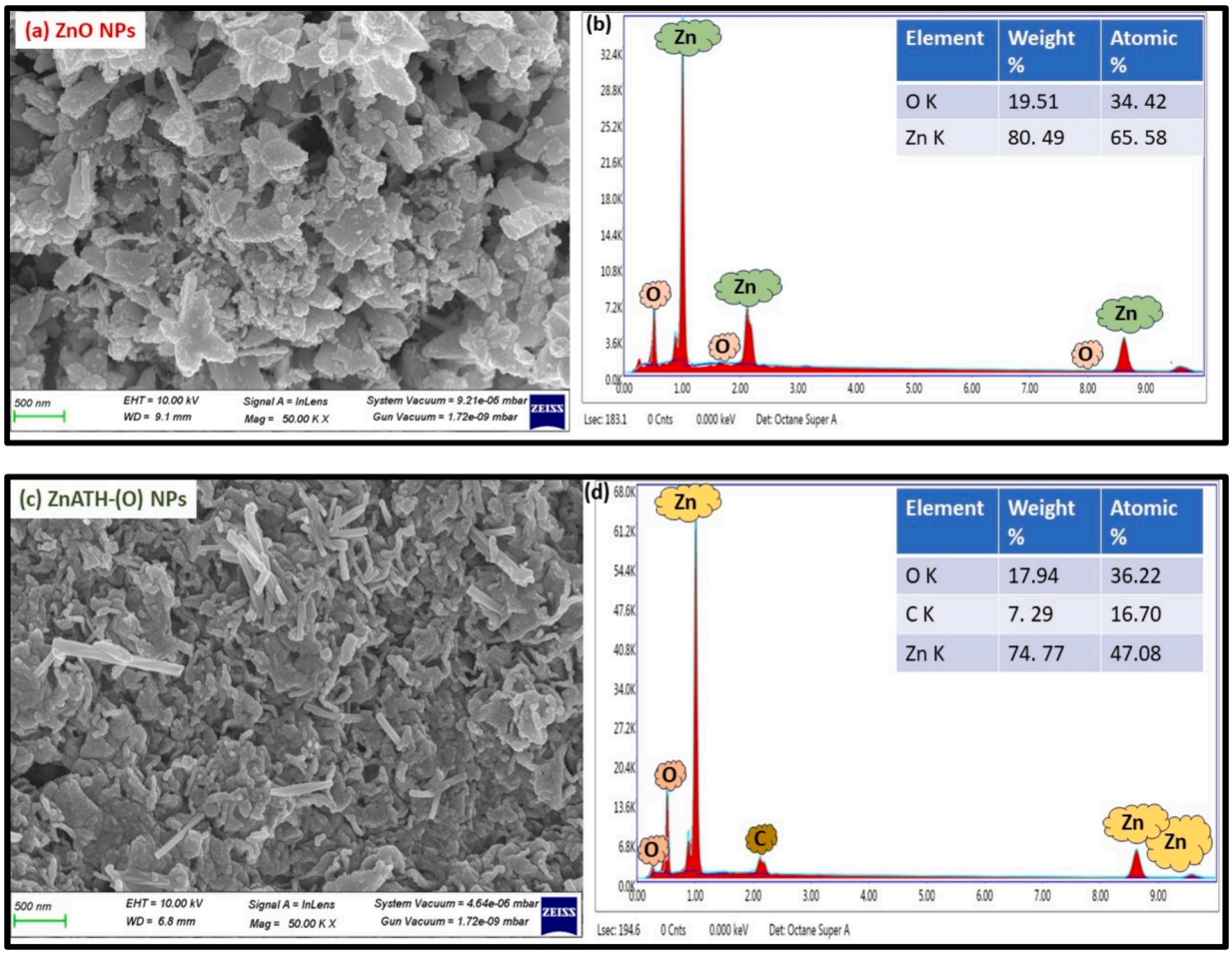


Fig. 6. SEM-EDX of (a),(b) ZnO NPs and (c),(d) ZnATH-(O) NPs.

Table 1
The average crystallite and particle size of the synthesized ZnO and ZnATH-(O) NPs calculated from XRD, SEM and BET analysis.

Nanoparticles (NPs)	XRD	SEM	BET
ZnO NPs	31.12 nm	29.54 nm	29.82 nm
ZnATH-(O) NPs	27.91 nm	21.43 nm	20.15 nm

Table 2
The Specific surface area (S_{BET}), Pore Volume, and Pore Diameter of the synthesized ZnO and ZnATH-(O) NPs calculated from BET analysis.

Nanoparticles (NPs)	Specific Surface Area (S_{BET})	Pore volume	Pore Diameter
ZnO NPs	15.685 m^2/g	0.159 cc/g	2.982 nm/g
ZnATH-(O) NPs	25.187 m^2/g	0.224 cc/g	2.015 nm/g

ZnATH-(O) NPs, as shown in Fig. 7 b. The mean pore diameter was determined to be 2.982 nm/g for ZnO NPs and 2.015 nm/g for ZnATH-(O) NPs. The data above validates that the ZnATH-(O) NPs exhibit a greater surface area and pore volume than the ZnO NPs. The IUPAC classification suggests that Type IV isotherm with an H3 hysteresis loop is followed by ZnATH- (O) NPs. The pores have a diameter of more than 2 nm, indicating that the produced NPs are non-rigid aggregates of plate-like to irregular morphology and are macroporous in nature [32,33]. The results obtained from BET analysis are consistent with SEM data, which demonstrates plate-like to irregular morphology and reduced pore size after successful incorporation of ATH extract over the ZnO matrix. Consequently, the presence of macro-porosity and the expansion of the surface area contribute to the favourable photocatalytic degradation properties of ZnATH-(O) NPs.

Zeta potential analysis was used to measure the surface charge of the produced NPs. In the present work, the surface charge was found to be −40.9 mV, suggesting that the surface of ZnATH-(O) NPs has a higher negative charge than that of ZnO NPs, having −69.07 mV. The higher negative zeta potential indicates the presence of negatively charged ions or functional groups on the surface of ZnATH-(O) NPs. These charged entities generate an electrical repulsion between the NPs, which hinders their tendency to agglomerate together or form larger groups. A higher negative zeta potential value corresponds to increased repulsion and enhanced stability of the ZnATH-(O) NPs in the solution, confirming the role of ATH as a stabilising agent.

3.2. Photocatalytic degradation of CV dye

The photocatalytic experiments were conducted to assess the efficacy of synthesized ZnO NPs and ZnATH-(O) NPs as catalysts in degrading CV dye molecules from wastewater. The photocatalytic degradation investigations were conducted using 100 mL aqueous solutions of CV, with a concentration of 20 mg/ L and 0.1 g of synthesised ZnO NPs and ZnATH-(O) NPs under visible light irradiation for 90 min. Hence, it was found that ZnO NPs as a catalyst show 81 % photodegradation of CV dye, while on the other hand, ZnATH-(O) NPs show 96 % photocatalytic degradation of CV dye, as shown in Fig. 8 a and 8b. Therefore, further analysis of other catalytic degradation parameters was carried out using ZnATH-(O) NPs as a highly efficient photocatalyst. The comparative plot of degradation efficiency of ZnO NPs and ZnATH-(O) NPs is shown in

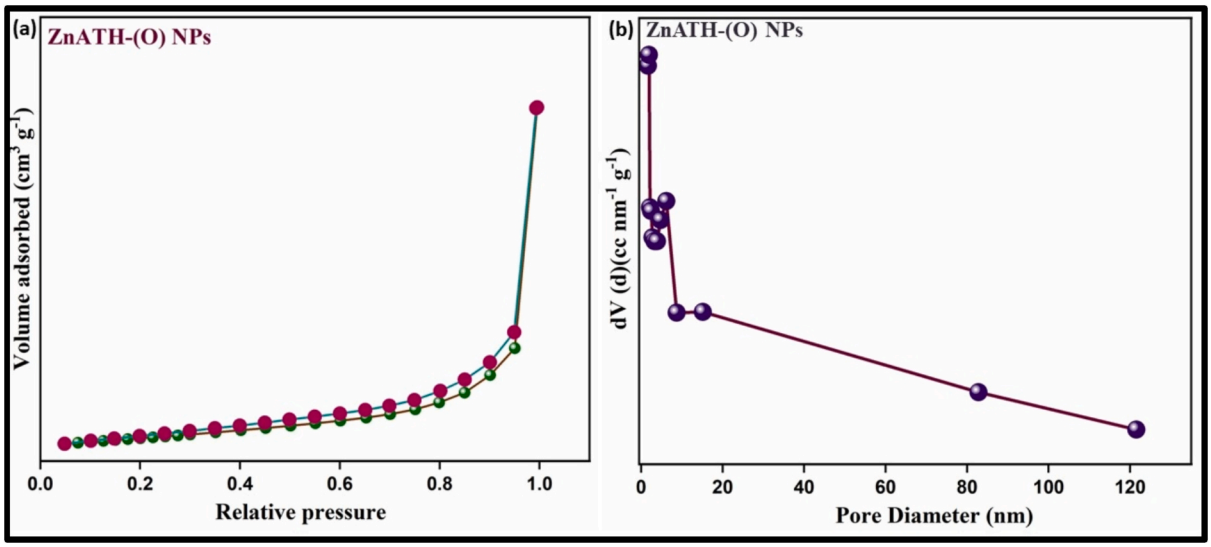


Fig. 7. (a) N_2 adsorption–desorption isotherm and (b) pore size distribution plot of ZnATH-(O) NPs.

Fig. 8 c.

3.2.1. Influence of reaction time

The study investigated the relationship between reaction time and the percentage of CV degraded over a period of 90 min using a 100 mL aqueous solution of CV with a concentration of 20 mg/L and a catalyst (ZnATH-(O) NPs) amount of 0.1 g. Fig. 8 c shows the plot of degradation efficiency versus reaction time, confirming that the degradation efficiency of CV dye increases with an increase in reaction time. The degradation efficacy increased by 62.84 % and 96.12 % in the first 30 and 90 min, respectively. The ZnATH-(O) NPs reach equilibrium after 90 min, as no substantial increase is observed after 96.12 %. The increase in degradation efficacy is attributed to the electrostatic attraction and intermolecular hydrogen bonding between CV molecules and the available active sites on the ZnATH-(O) NPs surface [5]. However, as time increases further, the quantity of available sites reduces, resulting in a slight decrease in the degradation efficiency of ZnATH-(O) NPs after 60 min. It was found that the photocatalytic degradation of CV dye molecules follows pseudo-first-order kinetics. The R^2 value was calculated to be 0.991, which signifies that the experimental results suited well with pseudo-first- order kinetic model. The rate constant was calculated to be $-0.03778\ min^{-1}$. The plot of $ln\left(\frac{C}{C_i}\right) vst$ is shown in Fig. 8 d.

3.2.2. Influence of dye concentration

The initial concentration of CV dye was varied from 10 to 50 mg/L with an increment of 10 mg/L (Fig. 9 a). The efficiency of photocatalytic degradation of CV declined progressively from 97.59 %, 97.31 %, 94.17 %, 91.48 %, and 88.58 % at concentrations of 10–50 mg/L, respectively. Hence, the efficacy of ZnATH-(O) NPs in photocatalytic degradation of CV dye was higher at lower dye concentrations and remained satisfactory up to 20 mg/L. At lower concentrations, there is more degradation of CV molecules on the surface of the ZnATH- (O) NPs than at higher concentrations. However, the amount of OH and O_2^- radicals may not be enough to degrade larger concentrations of CV dye [34]. Additional intermediates may be produced, leading to competition with the OH generated by photo-induced holes. Therefore, based on the above parameters, the highest amount of CV dye that can be degraded through a photodegradation mechanism using 0.1 g of ZnATH-(O) NPs in 90 min is determined to be 20 mg/L.

3.2.3. Influence of catalyst dosage

The study further focused on evaluating the effectiveness of photodegradation of CV dye using varying amounts of catalyst. The studies were performed using 100 ml of 20 mg/L standard solutions of CV dye with stirring at 450 rpm and irradiation using a 125 W xenon lamp. The catalyst doses were varied from 0.05 g to 0.15 g. Fig. 9 b clearly demonstrates a significant improvement in the photodegradation efficiency of CV dye from 76.74 % to 98.13 % using ZnATH-(O) NPs as photocatalyst. As the amount of catalyst increases, the active sites on the catalyst surface also increase, leading to an increase in holes and hydroxyl radical generation [35].

3.2.4. Influence of initial pH

pH is an important influencing factor in photocatalytic studies since it significantly degrades different dyes and organic contaminants. Moreover, it substantially impacts the implementation of effective wastewater treatment methods. It was observed that a photodegradation efficiency of <90 % was achieved when the pH value was 3, as indicated by data in Fig. 10 a. However, when the pH value of the dye solution was increased from 3 to 7, the photodegradation efficiency increased from 90.52 % to 98.80 %, respectively. This is due to the electrostatic interaction between the negatively charged NPs and the positively charged CV (pKa = 5.31) molecules. The results indicate that there is an increase in the concentration of hydroxyl (OH) radicals under neutral pH conditions, which leads to an increase in the photodegradation of CV dye by ZnATH-(O) NPs. However, the degradation efficiency drops from 98.80 % to 86.56 % as the pH further rises from 7 to 9. This can be attributed to the deprotonation of surface hydroxyl groups resulting in weaker electrostatic interaction between the catalyst surface and CV molecules, decreasing the effectiveness of the negative charge on the NPs' surface [36].

3.2.4.1. Point of zero charge (pH_{pzc}). The point of zero charge (pH_{pzc}) was investigated to determine the surface charge of the ZnATH-(O) NPs using the salt addition process. In the current study, 20 mL of 0.1 M NaCl solutions with initial pH (pH_i) values adjusted to 3, 5, 7, and 9 were mixed with 0.02 g of catalyst. The solutions were then allowed to settle at room temperature for 24 h. After agitating the suspensions, the final pH (pH_f) value was recorded, and ΔpH was calculated using the Eq. (6).

$$\Delta pH = pH_f - pH_i \tag{6}$$

The plot of ΔpH *vs* pH_i is displayed in Fig. 10 b. ZnATH-(O) NPs were found to have a pH_{pzc} value of 7.40, implying that at pH < 7.40, their

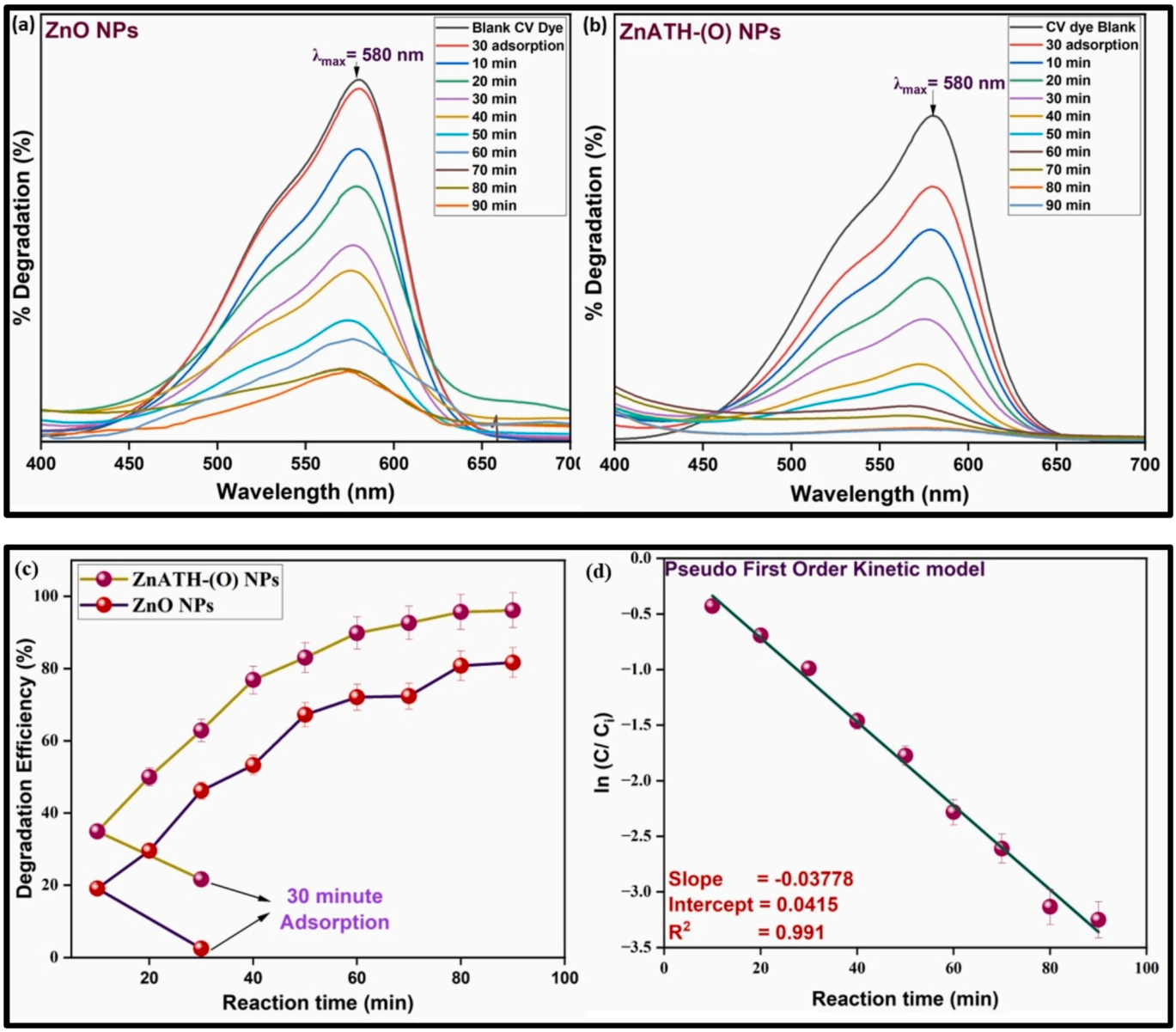


Fig. 8. (a) Degradation efficiency of ZnO NPs, (b) ZnATH-(O) NPs, (c) comparative plot of reaction time on the degradation efficiency of ZnO NPs and ZnATH-(O) NPs, and (d) Pseudo-first-order kinetic model.

surface is negatively charged. Consequently, this promotes electrostatic interaction between the cationic CV molecules and eases their degradation onto the ZnATH-(O) NPs surface. On the other hand, at pH values higher than 7.40, the positively charged catalyst surface electrostatically repels the CV cations, thereby decreasing the degradation efficiency [37]. This result is consistent with the data obtained by studying the effect of pH on the photocatalytic degradation of CV dye. It was observed that the degradation efficiency decreases on increasing the pH of the solution from 7 to 9.

3.2.5. *Influence of ionic strength*

Photocatalytic degradation of CV dye has shown that the nature and concentration of the electrolyte ionic species (such as NaCl and KCl) significantly impact the quantity of dye degradation. Inorganic salts and ionic species are commonly employed as levelling and dye-promoting agents in printing and dyeing procedures. So, it is important to examine the impact of ionic strength on the photodegradation of CV dye molecules, as shown in Fig. 11 a. It was found that the increase in ionic strength promoted the degradation of CV dye due to a combination of Na^+ with photoelectrons generated from the surface during irradiation [38].

This leads to the enhanced accessibility of H^+ for further reaction, as mentioned in Eq. (7). As the concentration of Na^+ increases from 0.00 to 0.005 M, the formation of $OH^{\cdot}$ radicals will be faster, resulting in enhanced efficacy of ZnATH-(O) NPs in degrading CV from 97.6 % to 99.8 %.

$$ZnATH-(O)NPs + h\nu \rightarrow ZnATH-(O)(h^+ + e^-) \quad (7)$$

$$Na^+ + e^- \rightarrow Na \quad (8)$$

$$h^+ + H_2O \rightarrow OH^{\bullet} + H^+ \quad (9)$$

$$h^+ + OH^- \rightarrow OH^{\bullet} \quad (10)$$

3.2.6. *Photocatalytic degradation of CV dye in real water samples*

To assess the applicability of the synthesized phytofabricated ZnO NPs, we have examined their ability to degrade CV dye from tap water,

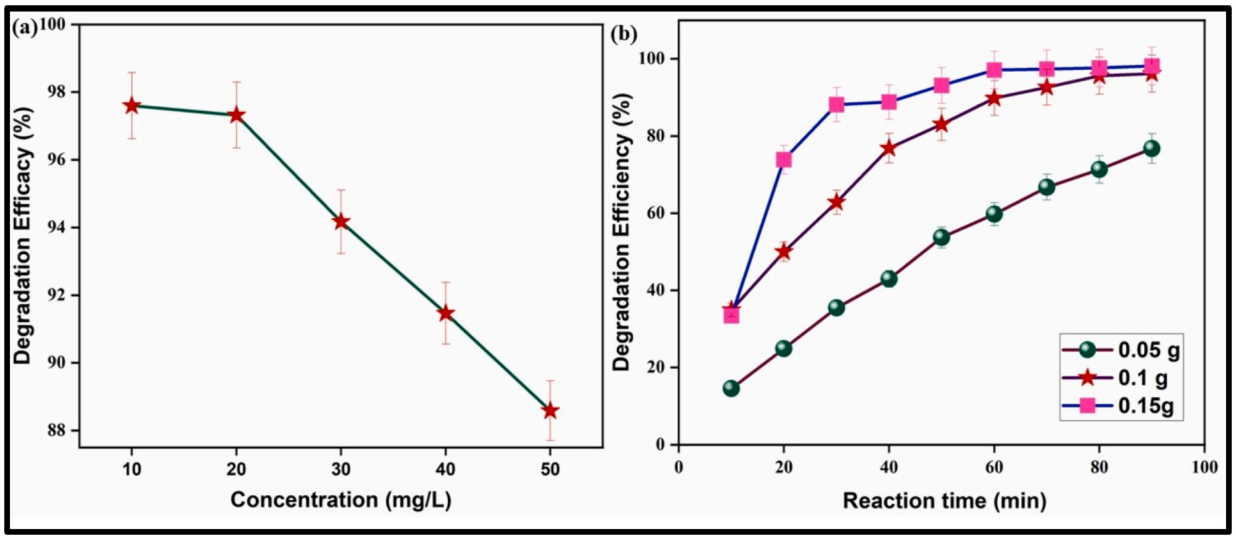


Fig. 9. (a) Impact of initial dye concentration on the degradation efficiency, and (b) impact of catalyst dosage on the degradation efficiency.

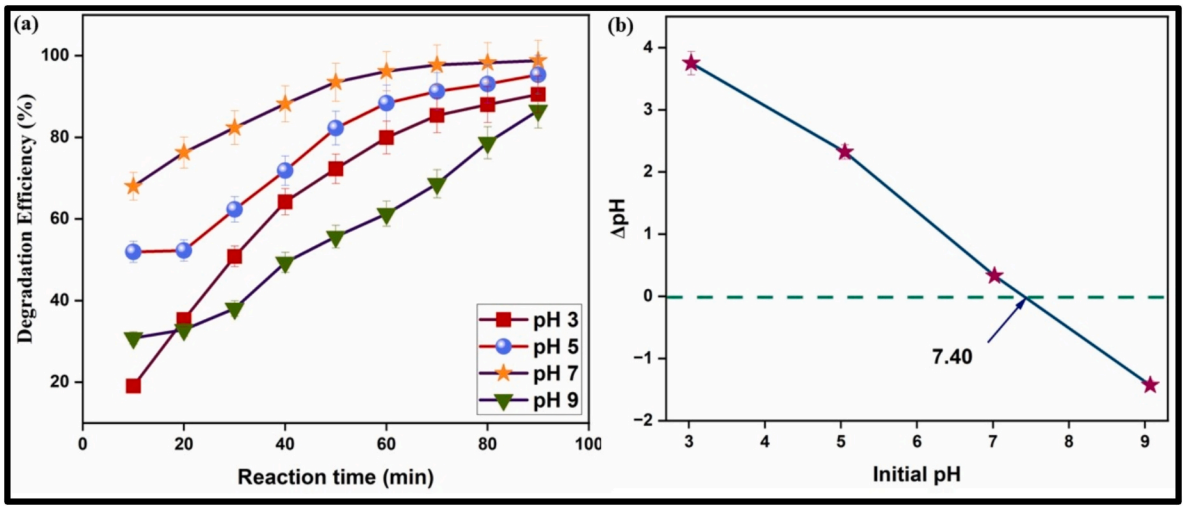


Fig. 10. (a) Influence of pH on the degradation efficiency with an increase in reaction time, and (b) pH_{pzc} plot for ZnATH-(O) NPs as a catalyst.

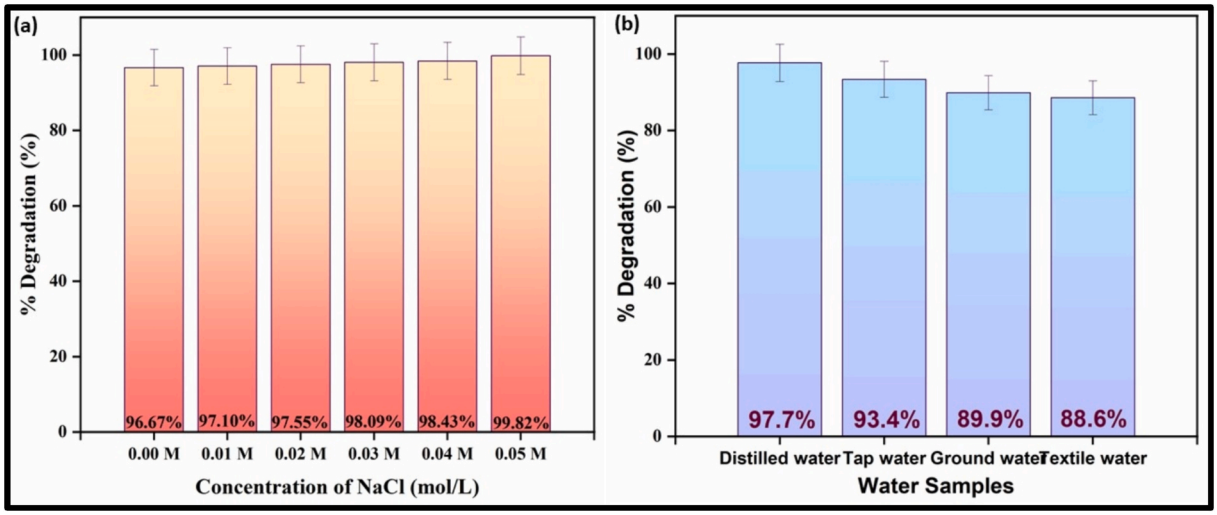


Fig. 11. (a) Influence of Ionic strength of NaCl on the degradation efficiency of ZnATH-(O) NPs, and (b) Degradation percentage of CV dye in different water samples.

groundwater, and textile water. The results obtained from analysing real water samples were subsequently compared to those of distilled water (Fig. 11b). The textile wastewater collected from the Bawana Industrial Area in New Delhi, India, was initially filtered to eliminate solid aggregates before being analysed. Due to the presence of minimal amounts of CV dye in the original samples, they were supplemented with a specific concentration of the catalyst. The percentage degradation of CV dye follows the following order: distilled water (97.7 %), tap water (93.4 %), groundwater (89.9 %), and textile water (88.6 %). Therefore, the synthesized ZnATH-(O) NPs can be used for environmental remediation in actual wastewater. These findings provide strong evidence for the current study than those previously published in the literature.

3.2.7. *Different photocatalysts for photocatalytic degradation of CV dyes*

The study also examined the most suitable parameters and degradation efficiency of recently discovered catalysts for eliminating CV, comparing them with ZnATH-(O) NPs (Table 3). The results demonstrated that the catalyst utilized in the current investigation exhibited exceptional efficacy in degrading CV dye within a 90 min of visible light irradiation, surpassing the performance of other catalysts. Furthermore, compared to other catalysts, its larger surface area is accompanied by a negatively charged metal oxide surface. The NPs surface greatly enhances its ability to adsorb cationic CV dye.

3.2.8. *Recyclability of catalyst*

A study was conducted to assess the recyclability of the catalyst, ensuring its suitability for real-time applications. After the photocatalytic degradation process, the catalyst was collected and rinsed with a water–ethanol mixture to regenerate it. Subsequently, it was dried in an oven to assess its ability to be reused. The CV dye exhibited a degradation efficiency of over 94.12 % in the initial three cycles but declined to 89.79 % after the fourth cycle (Fig. 12). Therefore, the ZnATH-(O) NPs exhibited efficient catalytic capabilities and could be utilized repeatedly. These ZnATH-(O) NPs have the potential to degrade the CV dye more efficiently as compared to those reported in the literature because of their enhanced surface area, reduced size and optimised stable structure as mentioned in Table 3.

4. Plausible mechanism of photocatalytic degradation of CV dye

The possible mechanism for the photocatalytic degradation of CV dye molecules involves three main steps: (i) demethylation leading to photogeneration of electron-hole pairs (e^-- h^+) between the valence bands (VB) and the conduction bands (CB) due to excitation of ZnATH-(O) NPs, (ii) fragmentation of the central aromatic ring of CV dye molecules followed by the aromatic side rings, and (iii) conversion of the resulting fragments into intermediate molecules, such as R-NH_3^+, phenol, and aniline.

The process involves the formation of aldehydic/carboxylate species, followed by the conversion of these intermediates into end products, including CO_2, H_2O, NO_3^{2-}, and NH_4^+. The majority of the reaction intermediates are generated from the fragmentation of the aromatic ring of the CV dye. The dye fragments undergo degradation, resulting in the formation of further chemical intermediates such as aldehydes, carboxylic acids, phenols, and amines [31]. These intermediates are

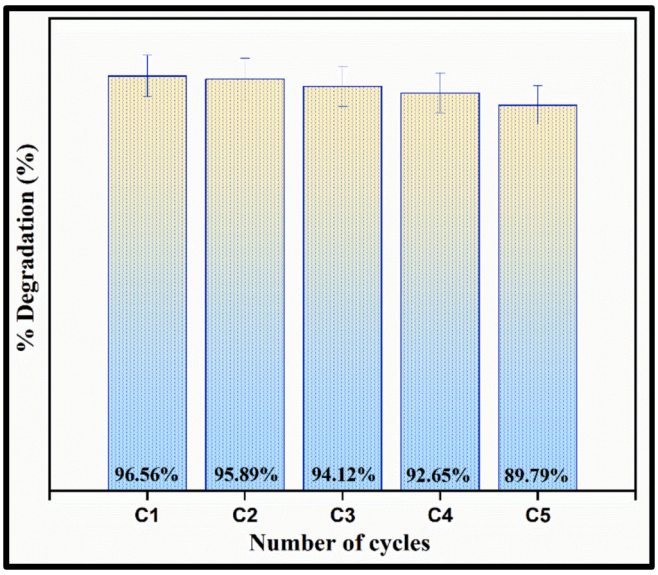


Fig. 12. Reusability of ZnATH-(O) NPs as catalyst for upto five cycles.

eventually transformed into CO_2, H_2O, NO_3^{2-}, NH_4^+ and CO_3^{2-} ions. The first stage entails the creation of an e^- and h^+ pair in the conduction and valence bands of ZnATH-(O) NPs when exposed to xenon light, as explained in Eq. (11).

$$\boldsymbol{ZnATH-(O)NPs + h\nu \rightarrow e^-_{CB} + h^+_{VB}} \quad (11)$$

The generated e^- and h^+ are transported to the surface of the photocatalyst. The rate of recombination of charge carriers slows down, leading to the separation of the generated e^--h^+ pairs to the surface. Secondly, the surface electrons (e^-) reduce oxygen (O_2) molecules to produce superoxide radicals ($\bullet O_2^-$), which further recombine with surface electrons to form peroxide molecules [4,6]. In addition, the surface holes/ proton (h^+) either oxidize water (H_2O) molecules to produce hydroxyl radicals (•OH) or directly oxidize the CV dye. These reactions are mentioned in Eqs. (12)–(15).

$$\boldsymbol{O_2 + e^- \rightarrow O_2^{\bullet} + H^+} \quad (12)$$

$$\boldsymbol{H^+ \rightarrow HOO^{\bullet}} \quad (13)$$

$$\boldsymbol{HOO^{\bullet} + e^- + H^+ \rightarrow H_2O_2} \quad (14)$$

$$\boldsymbol{H_2O + h^+ \rightarrow OH^{\bullet} + H^+} \quad (15)$$

The reactive species and intermediates ($\bullet O_2^-$, •OH, and h^+) are highly unstable, so they readily activate the redox processes and break down the CV dye into CO_2, H_2O, or inorganic ions, respectively, as presented in Eq. (16).

$$\boldsymbol{CV\ dye\ molecule + (O_2^{\bullet -}, OH^{\cdot}, and\ h^+) \rightarrow CO_2 + H_2O + inorganic\ ions} \quad (16)$$

Therefore, the aromatic rings decompose, resulting in the

Table 3
Different photocatalysts reported in the literature for the degradation of CV dye from wastewater.

Photocatalyst	Biosynthesis source	Initial dye concentration	Catalyst dosage	Reaction time	pH	% Degradation	References
ZnO NPs	*Ficus religiosa* leaf extract	10 mg/L	0.05 g	60 min	7	95 %	[34]
ZnO NPs	*Delonix Elata* leaf extract	–	0.02 g	90 min	–	86 %	[5]
ZnO NPs	*Schrebera swietenioides Roxb.* leaf extract	10 mg/L	1 g	120 min	–	85 %	[39]
NiO NPs	*Camellia sinensis* leaf extract	100 mg/L	0.03 g	42 h	–	99.5 %	[40]
CuO NPs	*Solanum lycopersicum* leaf extract	100 mg/L	0.025 g	60 min	–	97 %	[41]
ZnO NPs	*Brassica oleracea* var. *capitata*	**20 mg/L**	**0.1 g**	**90 min**	**7**	**97.66 %**	**Present Work**

degradation of colour of CV dye solution. The reaction steps are presented in Fig. 13. The photodegradation mechanism of CV dye can be comprehended by referring to Fig. 14, which illustrates the degradation of CV dye utilizing ZnATH-(O) NPs as a photocatalyst.

5. Conclusions

The present work explores the innovative synthesis and application of phytofabricated ZnO NPs using ATH extract from RC for the photocatalytic degradation of CV dye. The phytofabricated ZnO NPs show remarkable properties like a reduced particle size of 27.51 nm, irregular morphology, enhanced surface area, and improved surface charge. These characteristics starkly contrasted with those of ZnO NPs synthesized via the chemical co-precipitation method, which exhibited a larger size of 31.9 nm and a plate-like morphology with low surface area as compared to ZnATH-(O) NPs. The synthesized ZnATH-(O) NPs can be effectively utilised as catalysts in the photocatalytic degradation of CV dye using 100 ml CV dye solution of 20 mg/L dye concentration at neutral pH, under visible light irradiation for 90 min. The results revealed a superior degradation efficiency of the CV dye by the phytofabricated ZnO NPs, attributed to their smaller size and increased surface area, which facilitated more effective interaction with the dye molecules. In addition, ZnATH-(O) NPs show satisfactory reusability and a reasonable degradation efficiency of 89.79 % after five consecutive regeneration cycles.

This study improves the basic comprehension of green nanotechnology by minimizing chemical pollution and emphasizing its practical applications in environmental rehabilitation and environmental sustainability. By harnessing natural plant extracts like ATH from RC for NPs synthesis, reducing reliance on hazardous chemicals and minimising the environmental impact associated with conventional synthesis methods. The successful application of phytofabricated ZnO NPs for the degradation of harmful dyes not only highlights their potential in wastewater treatment but also emphasizes the broader implications of adopting sustainable and eco-friendly approaches in various industrial and environmental contexts. In conclusion, the integration of phytofabrication techniques with advanced nanotechnology presents a promising pathway for addressing environmental challenges. This study highlights the potential of using green-synthesized ZnO NPs to achieve efficient photocatalytic degradation of pollutants, offering a glimpse into a future where sustainable solutions are at the forefront of technological advancements for a cleaner and healthier planet.

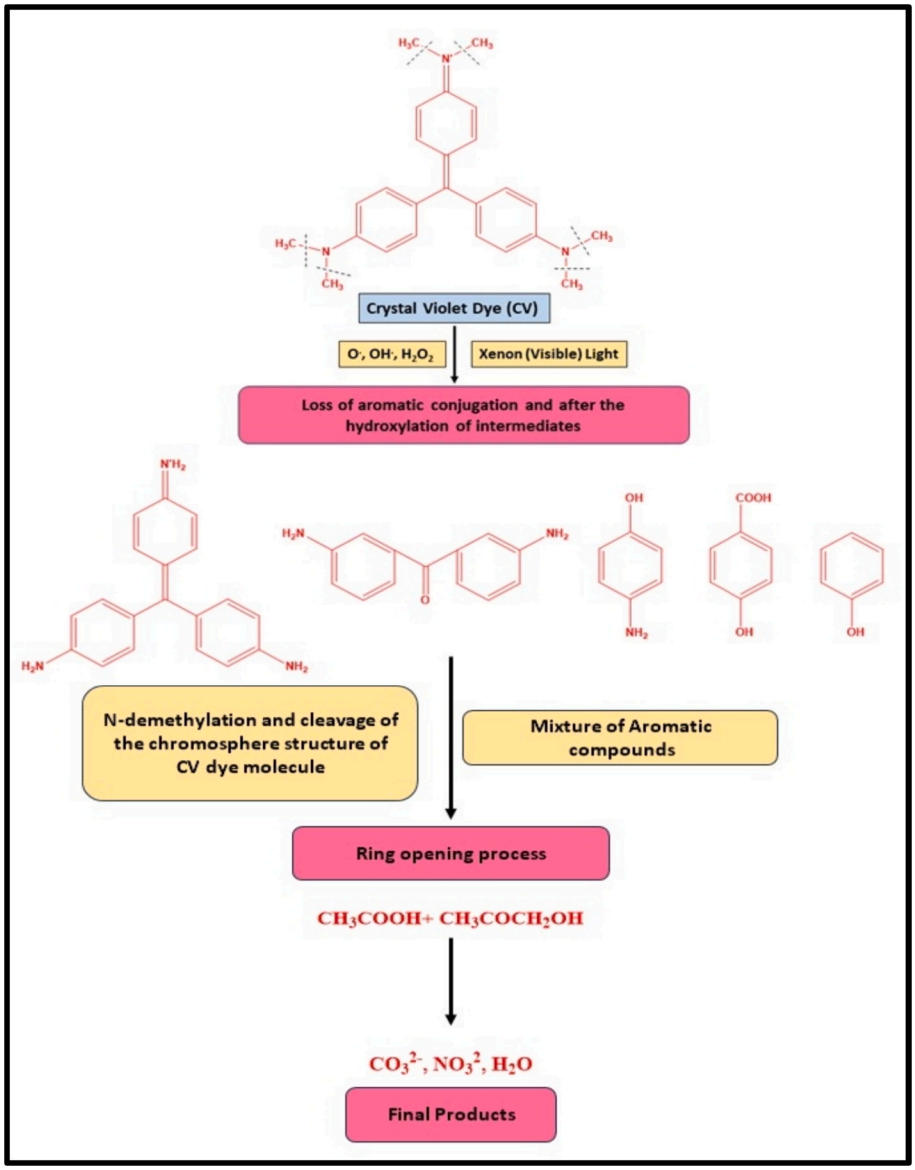


Fig. 13. Plausible reaction pathways for the photocatalytic degradation of CV dye molecules using ZnATH-(O) NPs.

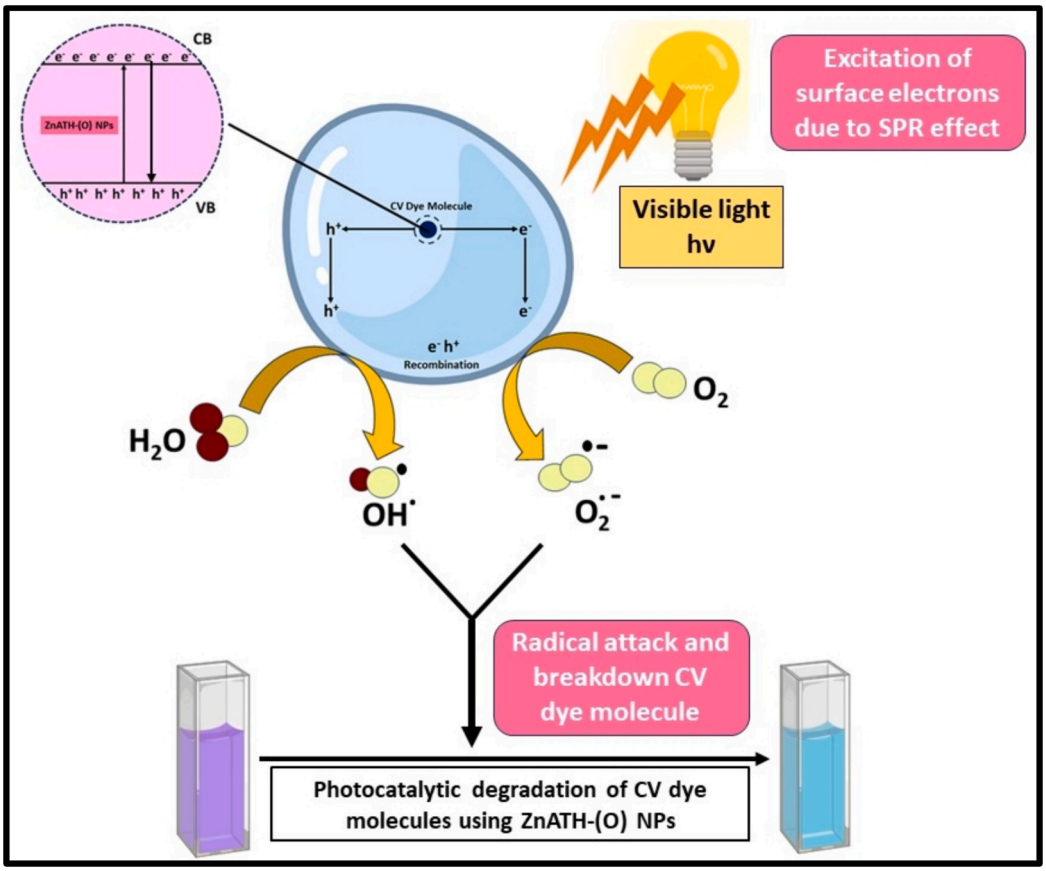


Fig. 14. Plausible photocatalytic degradation mechanism of CV dye molecules using ZnATH-(O) NPs as photocatalysts.

6. Declaration of generative AI in scientific writing

All the authors declared that they had not used any artificial intelligence (AI) or AI assistant technologies in the writing process.

CRediT authorship contribution statement

Sarla Yadav: Writing – original draft, Investigation, Formal analysis, Data curation. **Pooja Singh:** Writing – original draft, Formal analysis. **Gunjan Varshney:** Writing – original draft, Investigation, Formal analysis. **Saurabh:** Writing – original draft, Investigation, Formal analysis, Data curation. **Poonam Singh:** Formal analysis, Writing – original draft. **Raminder Kaur:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: DR. RAMINDER KAUR reports administrative support and equipment, drugs, or supplies were provided by DELHI TECHNOLOGICAL UNIVERSITY. DR. RAMINDER KAUR reports a relationship with Delhi Technological University that includes: employment. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank Delhi Technological University for providing technical and financial assistance.

Data availability

Data will be made available on request.

References

[1] H. Kolya, C.W. Kang, Biogenic synthesis of silver-iron oxide nanoparticles using kulekhara leaves extract for removing crystal violet and malachite green dyes from water, Sustain. 14 (2022), https://doi.org/10.3390/su142315800.
[2] P. Karthik, P.A. Jose, A. Chellakannu, S. Gurusamy, P. Ananthappan, R. Karuppathevan, V.S. Vasantha, J. Rajesh, S. Ravichandran, M. Sankarganesh, Green synthesis of MnO2 nanoparticles from Psidium guajava leaf extract: Morphological characterization, photocatalytic and DNA/BSA interaction studies, Int. J. Biol. Macromol. 258 (2024) 128869. https://doi.org/10.1016/j.ijbiomac.2023.128869.
[3] P. Karthik, S. Ravichandran, V. Sasikala, A. Mukkannan, J. Rajesh, Evaluation of MnO2 incorporated cellulose acetate membranes and their potential photocatalytic studies using Rhodamine-B dye, Process Saf. Environ. Prot. 179 (2023) 691–699. https://doi.org/10.1016/j.psep.2023.09.002.
[4] A. Jadhav, R. Chavan, S. Sonawane, P. Kamble, S. Mahajan, B. Vhankhande, R. Ghorpade, A. Chougale, N.M. Abd El-Salam, H. Fouad, R. Patil, Photocatalytic degradation of crystal violet dye using iron oxide nanoparticles, J. Nanoelectron. Optoelectron. 19 (2024) 272–277, https://doi.org/10.1166/jno.2024.3571.
[5] P. Karthik, S. Ravichandran, A. Mukkannan, J. Rajesh, Plant-mediated biosynthesis of zinc oxide nanoparticles from Delonix Elata: A promising photocatalyst for crystal violet degradation, Inorg. Chem. Commun. 146 (2022) 110122, https://doi.org/10.1016/j.inoche.2022.110122.
[6] K. Kanagamani, P. Muthukrishnan, K. Saravanakumar, K. Shankar, A. Kathiresan, Photocatalytic degradation of environmental perilous gentian violet dye using leucaena-mediated zinc oxide nanoparticle and its anticancer activity, Rare Met. 38 (2019) 277–286, https://doi.org/10.1007/s12598-018-1189-5.
[7] A.V. Kulkarni1, A. Chavhan, A. Bappakhane, J. Chimmankar, ZnO nanoparticles as adsorbent for removal of methylene blue dye, Rjces 4 (2016) 158–163, www.aelsindia.com/rjces.htm%0Ahttp://aelsindia.com/rjcesconfrenceproceeding/28.pdf.
[8] J. Pathak, P. Singh, Ultrafast catalytic reduction of organic pollutants using ternary zinc–copper–nickel layered double hydroxide, Appl. Organomet. Chem. (2024) 1–17, https://doi.org/10.1002/aoc.7507.
[9] A. Fouda, E. Saied, A.M. Eid, F. Kouadri, A.M. Alemam, M.F. Hamza, M. Alharbi, A. Elkelish, S.E.D. Hassan, Green synthesis of zinc oxide nanoparticles using an

aqueous extract of punica granatum for antimicrobial and catalytic activity, J. Funct. Biomater. 14 (2023), https://doi.org/10.3390/jfb14040205.
[10] P. Karthik, J. Rajesh, S. Ravichandran, Enhancement photocatalytic activity of CeO2 infused CA@Gelatin nanocomposites for the degradation of organic dyes under UV-light irradiation, Surf. Interfaces 50 (2024) 104474. https://doi.org/10.1016/j.surfin.2024.104474.
[11] K. Kanagamani, P. Muthukrishnan, A. Kathiresan, K. Shankar, P. Sakthivel, M. Ilayaraja, Detoxication and theranostic aspects of biosynthesised zinc oxide nanoparticles for drug delivery, Acta Metall. Sin. (english Lett. 34 (2021) 729–740, https://doi.org/10.1007/s40195-020-01116-x.
[12] W.M.C. Baldez, J.D.J.P. Santos, W.D.C. Santos, J. Aguilar-Pliego, N. Martín, A. A. Cabral, A.E.M. Paiva, S.F. Rodrigues, M.M. Teixeira, A.C.S. Alcântara, A. Rojas, Enhanced photodegradation of ciprofloxacin antibiotic using ZnO@FAU composite: A promising material for contaminant removal, Desalin. Water Treat. 318 (2024) 100356. https://doi.org/10.1016/j.dwt.2024.100356.
[13] Z. Mirzaeifard, Z. Shariatinia, M. Jourshabani, S.M. Rezaei Darvishi, ZnO photocatalyst revisited: effective photocatalytic degradation of emerging contaminants using S-doped ZnO nanoparticles under visible light radiation, Ind. Eng. Chem. Res. 59 (2020) 15894–15911, https://doi.org/10.1021/acs.iecr.0c03192.
[14] P. Karthik, S. Ravichandran, N. Prakash, A. Mukkannan, J. Rajesh, Evaluation of ZnO infused CA/PCL nanocomposites using potential wastewater treatment and invitro anticancer activity, Water Cycle 5 (2024) 121–130, https://doi.org/10.1016/j.watcyc.2024.03.002.
[15] J. Pathak, B. Pandey, P. Singh, R. Kumar, S. Kaushik, I.P. Sahu, T.K. Thakur, A. Kumar, Exploring the paradigm of phyto-nanofabricated metal oxide nanoparticles: recent advancements, applications, and challenges, Mol. Biotechnol. (2023), https://doi.org/10.1007/s12033-023-00799-8.
[16] S. Yadav, J. Pathak, P. Majhi, S. Kaushik, A.K. Shukla, R. Kaur, R. Kumar, P. Singh, Lac-based-biosynthesis of zinc–copper mixed metal oxide nanoparticles and evaluation of their antifungal activity against A.alternata and F. oxysporum, Mater. Chem. Phys. 330 (2025) 130152, https://doi.org/10.1016/j.matchemphys.2024.130152.
[17] S. Palamthodi, D. Kadam, S.S. Lele, Physicochemical and functional properties of ash gourd/bottle gourd beverages blended with jamun, J. Food Sci. Technol. 56 (2019) 473–482, https://doi.org/10.1007/s13197-018-3509-z.
[18] S. Netramai, T. Kijchavengkul, H. Samsudin, Enhanced Extraction of Anthocyanins from Red Cabbage (Brassica oleraces) Using Microwave Assisted Extraction Development of Colorimetric Indicators for pH and CO2 from Bio-based Materials and Plant Extract, for Short Shelf-life Foods View project, 2020. https://www.researchgate.net/publication/342242524.
[19] R. Mattioli, A. Francioso, L. Mosca, P. Silva, Anthocyanins: a comprehensive review of their chemical properties and health effects on cardiovascular and neurodegenerative diseases, Molecules 25 (2020), https://doi.org/10.3390/molecules25173809.
[20] N.L.W. Septiani, B. Yuliarto, M. Iqbal, Nugraha, synthesis of zinc oxide nanoparticles using anthocyanin as a capping agent, IOP Conf. Ser. Mater. Sci. Eng. 202 (2017), https://doi.org/10.1088/1757-899X/202/1/012070.
[21] S. Patil, R. Chandrasekaran, Biogenic nanoparticles: a comprehensive perspective in synthesis, characterization, application and its challenges, J. Genet. Eng. Biotechnol. 18 (2020), https://doi.org/10.1186/s43141-020-00081-3.
[22] R. Ravanfar, M. Moein, M. Niakousari, A. Tamaddon, Extraction and fractionation of anthocyanins from red cabbage: ultrasonic-assisted extraction and conventional percolation method, J. Food Meas. Charact. 12 (2018) 2271–2277, https://doi.org/10.1007/s11694-018-9844-y.
[23] J. Chandrasekhar, M.C. Madhusudhan, K.S.M.S. Raghavarao, Extraction of anthocyanins from red cabbage and purification using adsorption, Food Bioprod. Process. 90 (2012) 615–623, https://doi.org/10.1016/j.fbp.2012.07.004.
[24] N. Ghareaghajlou, S. Hallaj-Nezhadi, Z. Ghasempour, Red cabbage anthocyanins: Stability, extraction, biological activities and applications in food systems, Food Chem. 365 (2021) 130482, https://doi.org/10.1016/j.foodchem.2021.130482.
[25] U. Wijesinghe, G. Thiripuranathar, H. Iqbal, F. Menaa, Biomimetic synthesis, characterization, and evaluation of fluorescence resonance energy transfer, photoluminescence, and photocatalytic activity of zinc oxide nanoparticles, Sustain. 13 (2021) 1–22, https://doi.org/10.3390/su13042004.
[26] L.M. Jose, R.S.A. Raj, D. Sajan, A. Aravind, Adsorption and photocatalytic activity of biosynthesised ZnO nanoparticles using Aloe Vera leaf extract, Nano Exp. 2 (2021) 10039, https://doi.org/10.1088/2632-959X/abeec6.
[27] M.M. Chikkanna, S.E. Neelagund, K.K. Rajashekarappa, Green synthesis of Zinc oxide nanoparticles (ZnO NPs) and their biological activity, SN Appl. Sci. 1 (2019) 1–10, https://doi.org/10.1007/s42452-018-0095-7.
[28] H. Agarwal, V.K. Shanmugam, A review on anti-inflammatory activity of green synthesized zinc oxide nanoparticle: Mechanism-based approach, Bioorg. Chem. 94 (2020) 103423, https://doi.org/10.1016/j.bioorg.2019.103423.
[29] E.F. El-Belely, M.M.S. Farag, H.A. Said, A.S. Amin, E. Azab, A.A. Gobouri, A. Fouda, Green synthesis of zinc oxide nanoparticles (Zno-nps) using arthrospira platensis (class: Cyanophyceae) and evaluation of their biomedical activities, Nanomaterials 11 (2021) 1–18, https://doi.org/10.3390/nano11010095.
[30] J. Sun, H. Jiang, H. Wu, C. Tong, J. Pang, C. Wu, Multifunctional bionanocomposite films based on konjac glucomannan/chitosan with nano-ZnO and mulberry anthocyanin extract for active food packaging, Food Hydrocoll. 107 (2020) 105942, https://doi.org/10.1016/j.foodhyd.2020.105942.
[31] K. Rambabu, G. Bharath, F. Banat, P.L. Show, Green synthesis of zinc oxide nanoparticles using Phoenix dactylifera waste as bioreductant for effective dye degradation and antibacterial performance in wastewater treatment, J. Hazard. Mater. 402 (2021) 123560, https://doi.org/10.1016/j.jhazmat.2020.123560.
[32] M.A. Ismail, K.K. Taha, A. Modwi, L. Khezami, ZnO nanoparticles: Surface and X-ray profile analysis, J. Ovonic Res. 14 (2018) 381–393.
[33] M. Thommes, K. Kaneko, A. V Neimark, J.P. Olivier, F. Rodriguez-reinoso, J. Rouquerol, K.S.W. Sing, Physisorption of gases , with special reference to the evaluation of surface area and pore size distribution (IUPAC Technical Report), (2015). https://doi.org/10.1515/pac-2014-1117.
[34] B. Pandey, H. Tewatia, D. Meel, P. Singh, Phyto-fabrication of zinc oxide nanoparticles using Ficus religiosa and its potential application in degradation of crystal violet dye, Appl. Chem. Eng. 6 (2023) 1–11. https://doi.org/10.24294/ace.v6i3.2024.
[35] G. Kathiresan, K. Vijayakumar, A.P. Sundarrajan, H.S. Kim, K. Adaikalam, Photocatalytic degradation efficiency of ZnO, GO and PVA nanoadsorbents for crystal violet, methylene blue and trypan blue dyes, Optik (Stuttg). 238 (2021) 166671, https://doi.org/10.1016/j.ijleo.2021.166671.
[36] A. Sharma, D. Mangla, A. Choudhry, M. Sajid, S. Ali Chaudhry, Facile synthesis, physico-chemical studies of Ocimum sanctum magnetic nanocomposite and its adsorptive application against Methylene blue, J. Mol. Liq. 362 (2022) 119752. https://doi.org/10.1016/j.molliq.2022.119752.
[37] V. Parimelazhagan, K. Natarajan, S. Shanbhag, S. Madivada, H.S. Kumar, Effective adsorptive removal of coomassie violet dye from aqueous solutions using green synthesized zinc hydroxide nanoparticles prepared from calotropis gigantea leaf extract, ChemEngineering 7 (2023), https://doi.org/10.3390/chemengineering7020031.
[38] T. Islam, Effect of pH , Ions and Ionic Strength on Ti0 2 -mediated Photodegradation of Brilliant Orange Effect of pH , Ions and Ionic Strength on Ti0 2 -mediated Photodegradation of Brilliant Orange, (2009).
[39] S.L. Tulasi, A.V.V.S. Swamy, P. Pavani, V. Subhashini, Green synthesis of nanostructured zinc particles using aqueous leaf extract of schrebera swietenioides roxb. and their catalytic application in degradation of methyl orange, crystal violet dyes and chromium metal, Int. J. Appl. Pharm. 14 (2022) 308–314. https://doi.org/10.22159/ijap.2022v14i2.43697.
[40] M.T. Yassin, F.O. Al-Otibi, A.A. Al-Askar, Photocatalytic removal of crystal violet dye utilizing greenly synthesized iron oxide nanoparticles, Separations 10 (2023), https://doi.org/10.3390/separations10090513.
[41] D. Vaidehi, V. Bhuvaneshwari, D. Bharathi, B.P. Sheetal, Antibacterial and photocatalytic activity of copper oxide nanoparticles synthesized using Solanum lycopersicum leaf extract, Mater. Res. Express 5 (2018) 85403, https://doi.org/10.1088/2053-1591/aad426.

Photonic Network Communications (2025) 49:3
https://doi.org/10.1007/s11107-025-01028-9

ORIGINAL PAPER

Pointing error influence on the performance of dual-hop radio frequency/wireless optical communication system over Nakagami-m/ Fisher-Snedecor turbulence channels

Rekha Rani[1] · M. Lakshmanan[2] · Anup Kumar Mandpura[3] · N. Jayanthi[1]

Received: 17 January 2024 / Accepted: 10 March 2025
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2025

Abstract
Radio-frequency/Wireless Optical Communication (RF/WOC) systems are used for long-distance transmission where the optical signals are severely impacted by atmospheric turbulence. This article provides the performance analysis of a dual-hop RF/WOC system where RF and wireless optical channels undergo Nakagami-m and Fisher-Snedecor ($\mathcal{F}$) distributions with approximate Beckmann distribution respectively. This system can operate with either coherent heterodyne detection (HD) or non-coherent intensity modulation with direct detection (IM/DD) scheme. The mathematical formulation of the cumulative distribution function (CDF) is derived for instantaneous overall received signal-to-noise ratio (SNR) using the fixed gain amplify and forward (AF) relaying technique. Its CDF is being utilized to create new analytical formulas for outage probability (OP), average bit error rate (ABER), and ergodic capacity (EC). Moreover, asymptotic analysis for the performance metrics is performed to get more insights. Additionally, a single WOC system that was previously studied and modeled by $\mathcal{F}$ fading distribution with pointing errors is compared with the proposed system. The analytical findings demonstrate that the performance of the system is exacerbated by essential factors, including atmospheric turbulence and pointing errors and the HD scheme is preferable to a IM/DD scheme. Moreover, ABER with CBPSK modulation is found better than other binary modulation schemes. Additionally, results reveal that the considered system has remarkable performance improvement over a single WOC system.

Keywords Nakagami-m distribution · Fisher-Snedecor distribution · Beckmann distribution · Cumulative distribution function

1 Introduction

Wireless optical communication (WOC) has served as an effective technology for huge demands of 5G and beyond communication to support high data rates for coverage areas ranging from a few nanometres to more than a 1000 km. Due to the shortage of licensed radio spectrum, WOC accommodates the rapidly rising bandwidth demands, higher data rate, immunity to electromagnetic interference, better security, and unrestricted access to the spectrum [1]. Terrestrial Free space optical (FSO) technology is beneficial for last-mile problems in future-generation networks, optical fiber backup, inter-city backhauls, emergency recovery, offshore communication, and military applications. It also served as an alternative solution to optical fiber communication systems as it offers rapid deployment and fast setup time, more flexibility, and low maintenance costs [2]. Due to the benefits indicated above, FSO has gained enormous attention in which a laser beam is used to transmit the data wirelessly. The performance of the FSO system is hindered by atmospheric turbulence and pointing errors and also limits the communication range to a few hundred meters [3, 4]. These constraints can be reduced by a dual-hop radio frequency/ wireless optical communication (RF/WOC), considered as RF/ FSO systems with different relaying techniques, where

✉ Rekha Rani
rekhaaggarwalece@gmail.com

[1] Department of ECE, Delhi Technological University, New Delhi, Delhi, India

[2] Department of ECE, Sri Eshwar College of Engineering and Technology, Coimbatore, Tamil Nadu, India

[3] Department of EE, Delhi Technological University, New Delhi, Delhi, India

Published online: 21 April 2025

the relay is positioned as an intermediary node between the source and the destination to enhance the long-range transmission. This approach combines the robustness of RF with the high-speed, high-capacity benefits of optical communication, making it suitable for diverse applications like long-distance communication, disaster recovery, and secure data transmission. In such systems, data is transferred from source to relay through the RF channel during the first hop. Afterward, the relay sends it in a subsequent hop across the FSO channel to the destination. Different relaying techniques are used at the relay such as decode and forward (DF) [5, 6] and amplify and forward (AF) [7]. The relay in the DF relaying technique decodes the input signal, reduces the noise, and forwards it to the destination whereas, in the AF-based relaying technique, the relay boosts the incoming signal using either variable gain or constant gain and then transmits it to the destination [8]. AF relay-based system has advantages over the DF due to its less complexity and simpler implementation [9]. Recently many researchers have examined the performance metrics for DF and AF-based dual-hop RF/FSO systems with pointing error using different detection schemes as enumerated in Table 1.

Various fading models have been used such as Rician, Rayleigh, Nakagami-m, $\eta-\mu$, and $k-\mu$ shadowed distributions have been used for RF links. Some distributions like $\eta-\mu$ and $k-\mu$ shadowed distributions have more mathematical complexity due to their complex confluent hypergeometric function in contrast to simple distributions like Rayleigh, Nakagami-m, and Rician [10]. Additionally, the Nakagami-m is generalized distribution, and other RF channels such as the single-sided Gaussian channel, Rayleigh channel, Rician channel, and AWGN channel are obtained as its special case.

For the FSO link, several fading models such as Gamma–Gamma ($\Gamma\Gamma$) distribution [11], Malaga ($\mathcal{M}$) distribution [12], Exponentiated Weibull (EW)[13], Double Generalized Gamma (DGG) distribution [14], and Fisher-Snedecor ($\mathcal{F}$) distribution [15] have been employed to describe the environment turbulence regimes. Although the $\Gamma\Gamma$ distribution is widely used only for strong turbulence conditions whereas $\mathcal{M}$ distribution, EW distribution, DGG distribution, and $\mathcal{F}$ distribution are acceptable for all turbulence conditions. But $\mathcal{M}$ distribution, EW distribution, and DGG distribution are mathematically more intractable than $\mathcal{F}$ distribution [15].

In addition to the atmospheric turbulence, pointing errors are another risk factor for FSO links which are caused by the transmitter and receiver being out of alignment [16]. To characterize this misalignment, many statistical models including Rayleigh [17], Hoyt [18], lognormal-Rician [19], Beckmann, and approximation Beckmann distribution [20] have been presented. Beckmann distribution is the generalized pointing error model. Since several models are included as a special case. Due to the intricacy of its mathematics, Beckmann distribution does not have a closed-form solution. Then, a modified Rayleigh distribution is used to approximate the Beckmann distribution to create the approximation Beckmann distribution [20].

However, different digital binary modulation schemes, aperture averaging, and different detection schemes are just a few methods utilized to improve performance. Various digital binary modulation methods are used to assess the effectiveness of dual-hop RF/FSO systems in literature [33, 34]. Additionally, the aperture averaging technique has been

Table 1 Literature summary related to dual-hop RF/FSO system considering pointing error

References	RF channel	FSO channel with pointing error	Relaying techniques	Detection techniques	Performance parameters
[21]	Rayleigh	$\Gamma\Gamma$	Fixed Gain AF	IM/DD	OP, ABER, and EC
[7]	Nakagami-m	$\Gamma\Gamma$	Fixed Gain AF	IM/DD	OP, ABER, and EC
[22]	Nakagami-m	$\Gamma\Gamma$	Fixed and Variable Gain AF	HD, IM/DD	OP, ABER, and EC
[23]	$\eta-\mu$	$\Gamma\Gamma$	DF	HD, IM/DD	OP, ABER, and EC
[24]	Rayleigh	$\mathcal{M}$	Fixed Gain AF	IM/DD	OP, ABER, and EC
[25]	Rician	$\mathcal{M}$	Fixed Gain AF	IM/DD	OP, ABER, and EC
[26]	Nakagami-m	$\mathcal{M}$	DF	HD, IM/DD	OP, ABER
[8]	$k-\mu$	$\mathcal{M}$	Fixed Gain AF	HD, IM/DD	OP, ABER, and EC
[27]	Nakagami-m	EW	DF	IM/DD	OP, ABER
[28]	Nakagami-m	EW	DF	IM/DD	ABER
[29]	Rayleigh	DGG	Fixed Gain AF	HD, IM/DD	OP, ABER, and EC
[30]	Rician	DGG	Fixed Gain AF	IM/DD	OP
[31]	Rayleigh	$\mathcal{F}$	Fixed and Variable Gain AF	HD, IM/DD	OP, ABER, and EC
[32]	$k-\mu$	$\mathcal{F}$	Fixed Gain AF	HD, IM/DD	OP, ABER, and EC

employed by some authors to minimize the effects of fading by averaging out intensity fluctuations to enhance the system performance [35]. Moreover, different optical detection techniques have been used at the receiver in the literature such as coherent heterodyne detection (HD) scheme, and non-coherent intensity modulation/ direct detection (IM/DD) scheme [36]. For scenarios requiring high spectral efficiency and long-distance communication, HD is advantageous despite its complexity. In contrast, for short-range, cost-sensitive applications, IM/DD is typically preferred due to its simplicity and lower cost. The choice between these schemes depends on factors such as distance, required data rates, security needs, and budget constraints [37].

Inspired by the literature mentioned above, we have considered the fixed gain AF-based dual-hop RF/FSO system where the RF link is followed by the Nakagami-m distribution and the FSO link is experienced by $\mathcal{F}$ distribution with a pointing error that follows the approximation Beckmann distribution in this paper. According to the author's knowledge, no prior work was explored for this kind of investigation. Moreover, Nakagami-m is a generalized distribution, and its special case for varying RF fading parameters yields different RF channel models including the Rayleigh, Rician, and single-sided Gaussian channels. It also reveals the statistical modeling of multipath fading channels (caused by reflections, refraction, diffraction, and scattering by surrounding objects, buildings, and others) and its mathematical form is more analytically tractable as compared to other distributions. Additionally, $\mathcal{F}$ distribution is more popular and applicable to all atmospheric turbulence conditions. Therefore, this work can serve as the benchmark for research findings.

The key aspects of this study are as follows:

- A dual-hop RF/FSO system employing a fixed-gain AF relaying technique is used to generate a closed-form formulation of CDF for instantaneous end-to-end (E2E) received signal-to-noise ratio under coherent HD and non-coherent IM/DD detection schemes.
- Based on its CDF, exact novel unified analytical expressions such as outage probability (OP), average bit error rate (ABER) for different digital binary modulation schemes, and ergodic capacity (EC) are evaluated where OP and ABER are presented in Meijer's G function, and EC is gained in bivariate Meijer's G function.
- Additionally, the resultant expressions are formulated asymptotically to acquire a deeper understanding.
- Exact and asymptotic expressions are simulated on MATLAB, and their performance is examined while accounting for the effects of fading caused by atmospheric turbulence and misalignment.
- Lastly, the performance of the considered system is contrasted to a single WOC system that was the subject of earlier research and is regarded as an FSO system represented by $\mathcal{F}$ distribution.

This paper is divided into different sections. The system model under consideration is covered in Sect. 2. Section 3 provides the CDF for instantaneous received SNR under both detection schemes. The exact and asymptotic mathematical frameworks for performance metrics are presented in Sect. 4. The simulated results are shown and then compared to an existing single FSO system in Sect. 5. The paper's conclusions are presented at the last.

2 System model of the proposed system

Figure 1 shows a proposed system that comprises a source, (S), a relay, (R), and a destination, (D). The data is passed from S to R via the RF link. The received signal at R is written as:

$$y_{SR} = h_{RF}x + n_{RF} \tag{1}$$

where x denotes the information sent by the S, h_{RF} is the channel coefficients represented by Nakagami-m distribution, and $n_{RF} \sim \mathcal{N}(0, \sigma_{RF}^2)$ is an additive white gaussian noise (AWGN) containing zero mean and variance, σ_{RF}^2.

Afterward, the RF signal is converted into an optical signal at R, which is then amplified with gain G and transmitted to D via the FSO link. The signal received at D is expressed as:

$$y_{RD} = \eta_{FSO}h_{FSO}Gy_{SR} + n_{FSO} \tag{2}$$

Substituting Eq. (1) into Eq. (2), it becomes:

$$y_{RD} = \eta_{FSO}h_{FSO}G\left(h_{RF}x + n_{RF}\right) + n_{FSO} \tag{3}$$

where η_{FSO} is the conversion coefficient,$n_{FSO} \sim \mathcal{N}(0, \sigma_{FSO}^2)$, h_{FSO} is the channel gain of the FSO link which is represented as $h_{FSO} = h_a h_p$, where h_a is the atmospheric turbulence modelled by $\mathcal{F}$ distribution and h_p is the pointing loss characterized by approximation Beckmann distribution. The instantaneous received SNR,$\gamma_{RF} = \frac{|h_{RF}|^2}{\sigma_{RF}^2}$, and average received SNR, $\overline{\gamma}_{RF} = \frac{\mathbb{E}^2|h_{RF}|}{\sigma_{RF}^2}$ are represented for the RF link.

The probability density function (PDF) and CDF of the RF link that experiences the Nakagami-m distribution are given by Siddharth et al. [38]

$$f\left(\gamma_{RF}\right) = \frac{m^m {\gamma_{RF}}^{m-1}}{\Gamma(m)\overline{\gamma}_{RF}{}^m} \exp\left(\frac{-m\gamma_{RF}}{\overline{\gamma}_{RF}}\right) \tag{4}$$

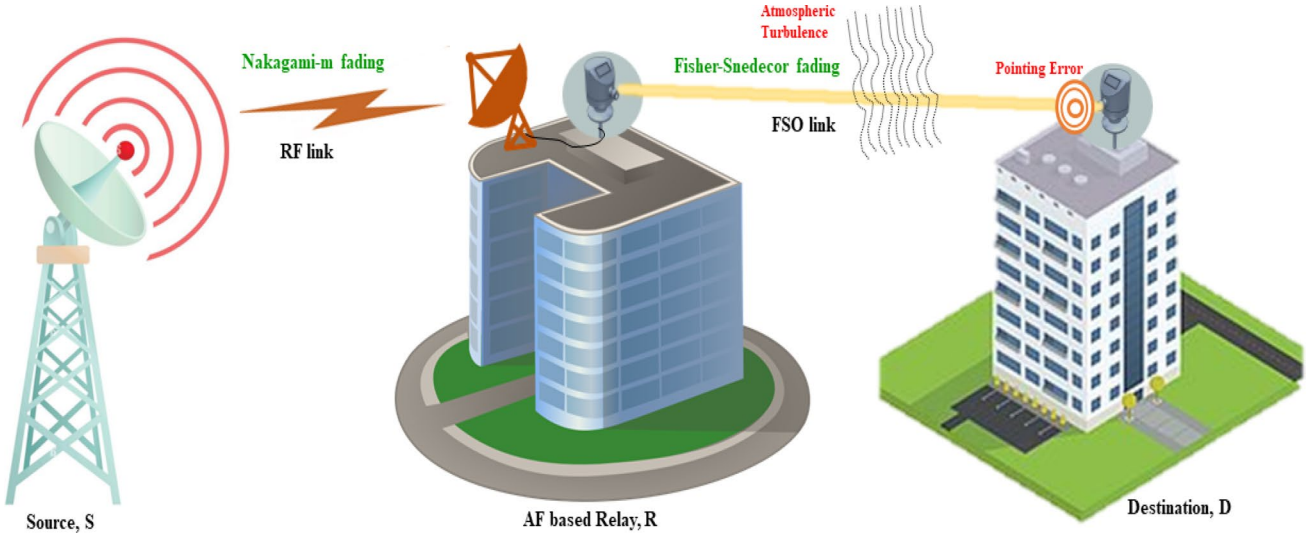


Fig. 1 System model of the dual-hop RF/FSO communication system under the influence of pointing error

$$F(\gamma_{RF}) = \frac{\gamma\left(m, \frac{m}{\overline{\gamma}_{RF}}\gamma_{RF}\right)}{\Gamma m} \tag{5}$$

where, $\gamma(-,-)$ denotes the lower incomplete Gamma function, and $\Gamma(.)$ represents the Gamma function respectively, and RF fading parameter, $m \geq 0.5$.

The instantaneous received electrical SNR, γ_{FSO} and average received electrical SNR, $\overline{\gamma}_{FSO}$ for the coherent HD and non-coherent IM/DD detection schemes are written as $\gamma_{FSO} = \frac{|\eta h_{FSO}|^t}{\sigma_{FSO}^2}$ and $\overline{\gamma}_{FSO} = \frac{\eta^t \mathbb{E}^t |h_{FSO}|}{\sigma_{FSO}^2}$, where t = 1, and 2 are represented for coherent HD and non-coherent IM/DD detection schemes respectively.

The nth moment of h_{FSO} is given by Ding et al. [32]

$$\mathbb{E}|h_{FSO}|^n = \left(\frac{A_{mode}(b-1)}{a}\right)^n \left(\frac{{\xi_{mode}}^2 \Gamma(b-n)\Gamma(a+n)}{({\xi_{mode}}^2 + n)\Gamma(b)\Gamma(a)}\right) \tag{6}$$

The PDF and CDF of γ_{FSO} that follow the $\mathcal{F}$ distribution are given by Badarneh et al. [34]

where, $G_{p,q}^{m,n}[.]$ is the Meijer's G function and values of a and b are determined as follows:

$$a = \frac{1}{\exp(\sigma_{lnS}^2) - 1}; b = \frac{1}{\exp(\sigma_{lnL}^2) - 1} + 2 \tag{9}$$

where σ_{lnS}^2 and σ_{lnL}^2 represent the small-scale and large-scale log-irradiance variances respectively, $\xi_{mode} = \frac{w_{eq}}{2\sigma_{mode}}$, $w_{eq} = \sqrt{\frac{\sqrt{\pi} {w_b}^2 \mathrm{erf}(v)}{2v\exp(-v^2)}}$, $\sigma_{mode} = \left(\frac{3\mu_x^2\sigma_x^4 + 3\mu_y^2\sigma_y^4 + \sigma_x^6 + \sigma_y^6}{2}\right)^{\frac{1}{6}}$, $v = \sqrt{\frac{\pi}{2}}\left(\frac{g}{w_b}\right)$, $A_{mode} = A_0\Omega$, $A_0 = (\mathrm{erf}(v))^2$, $\Omega = e^{\left(\frac{1}{{\xi_{mode}}^2} - \frac{1}{2{\xi_x}^2} - \frac{1}{2{\xi_y}^2} - \frac{\mu_x^2}{2{\xi_x}^2\sigma_x^2} - \frac{\mu_y^2}{2{\xi_y}^2\sigma_y^2}\right)}$, $\xi_x = \frac{w_{eq}}{2\sigma_x}$ and $\xi_y = \frac{w_{eq}}{2\sigma_y}$, where ξ_{mode} refers to the displacement deviation of pointing error, w_{eq} indicates the receiver's equivalent beam width, σ_{mode} implies the receiver's pointing error displacement standard deviation, w_b means gaussian beamwidth, g signifies the receiver's aperture radius, non-zero boresight errors μ_x, μ_y

$$f(\gamma_{FSO}) = \frac{{\xi_{mode}}^2 {\gamma_{FSO}}^{-1}}{t\Gamma(a)\Gamma(b)} G_{2,2}^{2,1}\left[\frac{a{\xi_{mode}}^2 {\gamma_{FSO}}^{\frac{1}{t}}}{(b-1)\left(1+{\xi_{mode}}^2\right)\overline{\gamma}_{FSO}^{\frac{1}{t}}}\middle| \begin{matrix} 1-b, 1+{\xi_{mode}}^2 \\ a, {\xi_{mode}}^2 \end{matrix}\right] \tag{7}$$

$$F(\gamma_{FSO}) = \frac{{\xi_{mode}}^2}{\Gamma(a)\Gamma(b)} G_{3,3}^{2,2}\left[\frac{a{\xi_{mode}}^2 {\gamma_{FSO}}^{\frac{1}{t}}}{(b-1)\left(1+{\xi_{mode}}^2\right)\overline{\gamma}_{FSO}^{\frac{1}{t}}}\middle| \begin{matrix} 1-b, 1, 1+{\xi_{mode}}^2 \\ a, {\xi_{mode}}^2, 0 \end{matrix}\right] \tag{8}$$

and different jittersσ_x,σ_y represent horizontal and elevation displacements respectively.

3 CDF of the proposed system

When fixed-gain AF relaying strategy is employed, instantaneous E2E received SNR, denoted as $\gamma_{E2E,AF}$ is given by Sharma et al. [23]:

$$\gamma_{E2E,AF} = \frac{\gamma_{FSO}\gamma_{RF}}{\gamma_{FSO} + C} \tag{10}$$

where, $C = \frac{1}{\sigma_{RF}^2 G^2}$

$$F_{\gamma_{E2E,AF}} = \int_0^{\infty} \frac{\gamma\left(m, \frac{m\gamma}{\overline{\gamma}_{RF}}\left(\frac{\gamma_{FSO}+C}{\gamma_{FSO}}\right)\right)}{t\Gamma(a)\Gamma(b)\Gamma(m)} \cdot {\xi_{mode}}^2 {\gamma_{FSO}}^{-1} G_{2,2}^{2,1}\left[\frac{a{\xi_{mode}}^2 \gamma_{FSO}^{\frac{1}{t}}}{(b-1)\left(1+{\xi_{mode}}^2\right)\overline{\gamma}_{FSO}^{\frac{1}{t}}}\middle| \begin{matrix} 1-b, 1+{\xi_{mode}}^2 \\ a, {\xi_{mode}}^2 \end{matrix}\right] d\gamma_{FSO} \tag{13}$$

Afterward, expand $\gamma(n,x)$ in the form of series as given by Gradshteyn and Ryzhik [39]

$$\gamma(n,x) = (n-1)!\left[1 - exp(-x)\sum_{m=0}^{n-1}\frac{x^m}{m!}\right] \tag{14}$$

On the substitution of Eq. (14) into Eq. (13) and after simplification, the CDF is reduced to

$$F_{\gamma_{E2E,AF}} = \underbrace{\frac{\xi_{mode}^2}{t\lceil(a)\lceil(b)}\int_0^{\infty}\gamma_{FSO}^{-1}G_{2,2}^{2,1}\left[\frac{a\xi_{mode}^2\gamma_{FSO}^{\frac{1}{t}}}{(b-1)\left(1+\xi_{mode}^2\right)\overline{\gamma}_{FSO}^{\frac{1}{t}}}\middle|\begin{matrix}1-b,1+\xi_{mode}^2\\ a,\xi_{mode}^2\end{matrix}\right]d\gamma_{FSO}}_{I_1} - \underbrace{\frac{\xi_{mode}^2}{t\lceil(a)\lceil(b)}e^{-B}\sum_{n=0}^{m-1}\frac{(B)^n}{n!}\int_0^{\infty}\left(\frac{\gamma_{FSO}+C}{\gamma_{FSO}}\right)^n exp\left(-\frac{BC}{\gamma_{FSO}}\right)\gamma_{FSO}^{-1}G_{2,2}^{2,1}\left[\frac{a\xi_{mode}^2\gamma_{FSO}^{\frac{1}{t}}}{(b-1)\left(1+\xi_{mode}^2\right)\overline{\gamma}_{FSO}^{\frac{1}{t}}}\middle|\begin{matrix}1-b,1+\xi_{mode}^2\\ a,\xi_{mode}^2\end{matrix}\right]d\gamma_{FSO}}_{I_2} \tag{15}$$

Afterward, CDF for instantaneous E2E received SNR, denoted as $F_{\gamma_{E2E,AF}}$ of the considered system becomes:

$$F_{\gamma_{E2E,AF}} = P_r\left[\frac{\gamma_{FSO}\gamma_{RF}}{\gamma_{FSO}+C} \le \gamma|\gamma_{FSO}\right] \tag{11}$$

where, $P_r[.]$ represents the probability that the instantaneous E2E received SNR, $\gamma_{E2E,AF}$ is less than or equal to γ.

After performing some algebraic manipulations, Eq. (11) is written as:

$$F_{\gamma_{E2E,AF}} = \int_0^{\infty} P_r\left[\gamma_{RF} \le \frac{\gamma\left(\gamma_{FSO}+C\right)}{\gamma_{FSO}}\right] f\left(\gamma_{FSO}\right) d\gamma_{FSO} \tag{12}$$

Substituting Eqs. (5) and (7) into Eq. (12), the CDF for $\gamma_{E2E,AF}$ is written as:

assuming, $B = \frac{m\gamma}{\overline{\gamma}_{RF}}$

After applying Eq. (07.34.21.0009.01) given in [40], the term I_1 becomes:

$$I_1 = 1 \tag{16}$$

Utilizing expression Eq. (1.111) provided by [39] and using the formula $e^x = G_{01}^{10}\left(-x\middle|\begin{matrix}-\\0\end{matrix}\right) = G_{10}^{01}\left(-\frac{1}{x}\middle|\begin{matrix}1\\-\end{matrix}\right)$, the term I_2 is reduced to:

$$I_2 = \frac{\xi_{mode}^2}{t\lceil(a)\lceil(b)}e^{-B}\sum_{n=0}^{m-1}\frac{B^n}{n!}\sum_{k=0}^{n}\binom{n}{k}C^k\int_0^{\infty}\gamma_{FSO}^{-1-k}G_{1,0}^{0,1}\left(\frac{\gamma_{FSO}}{BC}\middle|\begin{matrix}1\\-\end{matrix}\right) G_{2,2}^{2,1}\left[\frac{a\xi_{mode}^2\gamma_{FSO}^{\frac{1}{t}}}{(b-1)\left(1+\xi_{mode}^2\right)\overline{\gamma}_{FSO}^{\frac{1}{t}}}\middle|\begin{matrix}1-b,1+\xi_{mode}^2\\ a,\xi_{mode}^2\end{matrix}\right]d\gamma_{FSO} \tag{17}$$

With the aid of Eq. (07.34.21.0013.01) given in [40], Eq. (17) is obtained in a closed-form as:

$$I_2 = \frac{\xi_{mode}^2}{t\Gamma(a)\,\Gamma(b)} e^{-B} \sum_{n=0}^{m-1} \frac{B^n}{n!} \sum_{k=0}^{n} \binom{n}{k} C^k \frac{t^{a+b-1}}{(2\pi)^{t-1}} \left(\frac{1}{BC}\right)^k G_{2t,2t+1}^{2t+1,t}\left(BC\left(\frac{a\xi_{mode}^2}{(b-1)\left(1+\xi_{mode}^2\right)\frac{1}{\overline{\gamma}_{FSO}}} \right)^t \middle| \begin{matrix} \left(\frac{1-b}{t}\right), \left(\frac{-b+t}{t}\right), \left(\frac{1+\xi_{mode}^2}{t}\right), \left(\frac{\xi_{mode}^2+t}{t}\right) \\ \frac{a}{t}, \left(\frac{a+t-1}{t}\right), \left(\frac{\xi_{mode}^2}{t}\right), \left(\frac{\xi_{mode}^2+t-1}{t}\right), k \end{matrix} \right) \tag{18}$$

Substituting Eqs. (16) and (18) into Eq. (15), the unified expression of CDF for instantaneous E2E received SNR is obtained as:

4 Performance parameters

Novel analytical formulations for the outage probability, average bit error rate, and ergodic capacity of the considered system are evaluated in this section.

$$F_{\gamma_{E2E,AF}}(\gamma) = 1 - \frac{\xi_{mode}^2}{t\Gamma(a)\,\Gamma(b)} \frac{t^{a+b-2} e^{-\frac{m\gamma}{\overline{\gamma}_{RF}}}}{(2\pi)^{t-1}} \sum_{n=0}^{m-1} \frac{\left(\frac{m\gamma}{\overline{\gamma}_{RF}}\right)^n}{n!} \sum_{k=0}^{n} \binom{n}{k} \left(\frac{m\gamma}{\overline{\gamma}_{RF}}\right)^{-k} G_{2t,2t+1}^{2t+1,t}\left(\frac{C m\gamma}{\overline{\gamma}_{RF}\overline{\gamma}_{FSO}} \left(\frac{a\xi_{mode}^2}{(b-1)\left(1+\xi_{mode}^2\right)} \right)^t \middle| \begin{matrix} \mathfrak{B}_1 \\ \mathfrak{B}_2 \end{matrix} \right) \tag{19}$$

where

$$\mathfrak{B}_1 = \left[\mathfrak{B}_{11}, \mathfrak{B}_{12}, \mathfrak{B}_{13}, \mathfrak{B}_{14}\right] = \left[\left(\frac{1-b}{t}\right), \left(\frac{-b+t}{t}\right), \left(\frac{1+\xi_{mode}^2}{t}\right), \left(\frac{\xi_{mode}^2+t}{t}\right) \right]$$

$$\mathfrak{B}_2 = \left[\mathfrak{B}_{21}, \mathfrak{B}_{22}, \mathfrak{B}_{23}, \mathfrak{B}_{24}, \mathfrak{B}_{25}\right] = \left[\frac{a}{t}, \left(\frac{a+t-1}{t}\right), \left(\frac{\xi_{mode}^2}{t}\right), \left(\frac{\xi_{mode}^2+t-1}{t}\right), k \right]$$

Further, by employing DF at the relay, the instantaneous E2E received SNR, $\gamma_{E2E,DF}$ is given as

$$\gamma_{E2E,DF} = \min(\gamma_{RF}, \gamma_{FSO}) \tag{20}$$

Afterward, CDF for instantaneous E2E received SNR, $\gamma_{E2E,DF}$ of the considered system is

$$F_{\gamma_{E2E,DF}}(\gamma) = F\left(\gamma_{RF}\right) + F\left(\gamma_{FSO}\right) - F\left(\gamma_{RF}\right)F\left(\gamma_{FSO}\right) \tag{21}$$

By substituting Eqs. (5) and (8) into Eq. (21), we obtained the CDF of $\gamma_{E2E,DF}$ for DF relaying system.

4.1 Outage probability, $F_{\gamma_{E2E,AF}}(\gamma_{th})$

It plays a vital role in determining how reliable a considered system is. When the instantaneous E2E received SNR is less than the tolerable threshold SNR, it is deemed to be in outage i.e. $F_{\gamma_{E2E,AF}}(\gamma_{th}) = P\left(\gamma_{E2E,AF} \le \gamma_{th}\right)$. The outage probability, $F_{\gamma_{E2E,AF}}(\gamma_{th})$ is obtained by replacing γ with γ_{th}in Eq. (19), and it is found as:

$$F_{\gamma_{E2E,AF}}(\gamma_{th}) = 1 - \frac{\xi_{mode}^2}{\Gamma(a)\,\Gamma(b)} \frac{t^{a+b-2} e^{-\frac{m\gamma_{th}}{\overline{\gamma}_{RF}}}}{(2\pi)^{t-1}} \sum_{n=0}^{m-1} \frac{\left(\frac{m\gamma_{th}}{\overline{\gamma}_{RF}}\right)^n}{n!} \sum_{k=0}^{n} \binom{n}{k} \left(\frac{m\gamma_{th}}{\overline{\gamma}_{RF}}\right)^{-k} G_{2t,2t+1}^{2t+1,t}\left(\frac{C m\gamma_{th}}{\overline{\gamma}_{RF}\overline{\gamma}_{FSO}} \left(\frac{a\xi_{mode}^2}{(b-1)\left(1+\xi_{mode}^2\right)} \right)^t \middle| \begin{matrix} \mathfrak{B}_1 \\ \mathfrak{B}_2 \end{matrix} \right) \tag{22}$$

To get more useful performance insights for the high value of SNR, i.e. $\overline{\gamma}_{RF} \to \infty$ and $\overline{\gamma}_{FSO} \to \infty$, we assume $\lim_{\overline{\gamma}_{RF}\to\infty} e^{-\frac{m\gamma_{th}}{\overline{\gamma}_{RF}}} = 1$ and expand the Meijer-G function in Eq. (22) using Eq. (07.34.06.0040.01) given as [40], By doing so, an asymptotic expression of outage probability is obtained as:

$$F^{\infty}_{\gamma_{E2E,AF}}(\gamma_{th}) = 1 - \frac{\xi^2_{mode}}{\Gamma(a)\,\Gamma(b)} \frac{t^{a+b-2}}{(2\pi)^{t-1}} \sum_{n=0}^{m-1} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} \left(\frac{m\gamma_{th}}{\overline{\gamma}_{RF}}\right)^{n-k} \sum_{j=1}^{2t+1} \left\{ \frac{\prod_{\substack{i=1 \\ i\neq j}}^{2t+1} \Gamma(\mathfrak{B}_{2i} - \mathfrak{B}_{2j}) \prod_{i=1}^{t} \Gamma(1 - \mathfrak{B}_{1i} - \mathfrak{B}_{2j})}{\prod_{i=t+1}^{2t} \Gamma(\mathfrak{B}_{1i} - \mathfrak{B}_{2j})} \left(\frac{Cm\gamma_{th}}{\overline{\gamma}_{RF}\overline{\gamma}_{FSO}} \left(\frac{a\xi^2_{mode}}{(b-1)(1+\xi^2_{mode})} \right)^{t} \right)^{\mathfrak{B}_{2j}} \right\} \tag{23}$$

The asymptotic $F^{\infty}_{\gamma_{E2E,AF}}(\gamma_{th})$ expression can be rewritten in terms of the diversity order, G_d i.e., $F^{\infty}_{\gamma_{E2E,AF}} \propto (\overline{\gamma})^{-G_d}$ where $\overline{\gamma}_{RF} = \overline{\gamma}_{FSO} = \overline{\gamma}$ and $G_d = \min\left(\frac{a}{t}, \left(\frac{a+t-1}{t}\right), \left(\frac{\xi_{mode}{}^2}{t}\right), \left(\frac{\xi_{mode}{}^2+t-1}{t}\right), k\right)$ which determines the slope of the outage probability.

$$p^e = \underbrace{\frac{q^p}{2\lceil p} \int_0^{\infty} e^{-q\gamma}\gamma^{p-1} d\gamma}_{I_3} - \underbrace{\frac{q^p}{2\lceil p} \frac{\xi^2_{mode}}{\Gamma(a)\,\Gamma(b)} \frac{t^{a+b-2}}{(2\pi)^{t-1}} \sum_{n=0}^{m-1} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} \left(\frac{m}{\overline{\gamma}_{RF}}\right)^{-k+n} \int_0^{\infty} \gamma^{n-k+p-1} e^{-\left(q+\frac{m}{\overline{\gamma}_{RF}}\right)\gamma} G^{2t+1,\,t}_{2t,\,2t+1}\left(\frac{Cm\gamma}{\overline{\gamma}_{RF}\overline{\gamma}_{FSO}} \left(\frac{a\xi^2_{mode}}{(b-1)(1+\xi^2_{mode})} \right)^{t} \middle| \begin{matrix} \mathfrak{B}_1 \\ \mathfrak{B}_2 \end{matrix} \right)}_{I_4} \tag{25}$$

4.2 Average bit error rate, p^e

The another crucial parameter is ABER which is expressed for several digital binary modulation schemes as given by Han et al. [31]

$$p^e = \frac{q^p}{2\lceil p} \int_0^{\infty} exp(-q\gamma)\gamma^{p-1} F_{\gamma_{E2E,AF}}(\gamma) d\gamma \tag{24}$$

where, different binary modulation schemes are represented by p and q and achieved by setting the values of both.

Substituting Eq. (19) into Eq. (24), it becomes:

Utilizing the integral Eq. (3.381.4) of [39] and after simplification, the term I_3 becomes

$$I_3 = 0.5 \tag{26}$$

By employing the expression Eq. (07.34.21.0088.01) as given in [40], the term I_4 is obtained as:

$$I_4 = \frac{q^p}{2\lceil p} \frac{\xi^2_{mode} t^{a+b-2}}{\lceil a \lceil b (2\pi)^{t-1}} \sum_{n=0}^{m-1} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} \left(\frac{m}{\overline{\gamma}_{RF}}\right)^{-k+n} \left(q + \frac{m}{\overline{\gamma}_{RF}}\right)^{-n+k-p} G^{2t+1,t+1}_{2t+1,2t+1}\left(\frac{Cm}{\overline{\gamma}_{RF}\overline{\gamma}_{FSO}} \left(\frac{a\xi^2_{mode}}{(b-1)(1+\xi^2_{mode})} \right)^{t} \Big/ \left(q + \frac{m}{\overline{\gamma}_{RF}}\right) \middle| \begin{matrix} (1-n+k-p), \left(\frac{1-b}{t}\right), \left(\frac{-b+t}{t}\right), \left(\frac{1+\xi^2_{mode}}{t}\right), \left(\frac{\xi^2_{mode}+t}{t}\right) \\ \frac{a}{t}, \left(\frac{a+t-1}{t}\right), \left(\frac{\xi^2_{mode}}{t}\right), \left(\frac{\xi^2_{mode}+t-1}{t}\right), k \end{matrix} \right) \tag{27}$$

On the substitution of Eqs. (26) and (27) into Eq. (24) and after some mathematical manipulations, the closed-form expression of ABER for digital binary modulation schemes is obtained as follows:

$$p^{e} = 0.5 - \frac{q^{p}}{2\lceil p}\frac{\xi_{mode}^{2} t^{a+b-2}}{\lceil a \lceil b (2\pi)^{t-1}} \sum_{n=0}^{m-1} \frac{\left(\frac{m}{\bar{\gamma}_{RF}}\right)^{n}}{n!} \sum_{k=0}^{n} \binom{n}{k} \left(\frac{m}{\bar{\gamma}_{RF}}\right)^{-k} \left(q + \frac{m}{\bar{\gamma}_{RF}}\right)^{-n+k-p} G_{2t+1,2t+1}^{2t+1,t+1}\left(\frac{\frac{Cm}{\bar{\gamma}_{RF}\bar{\gamma}_{FSO}} \left(\frac{a\xi_{mode}^{2}}{(b-1)\left(1+\xi_{mode}^{2}\right)}\right)^{t}}{\left(q + \frac{m}{\bar{\gamma}_{RF}}\right)} \middle| \begin{matrix} \mathfrak{B}_3 \\ \mathfrak{B}_2 \end{matrix} \right) \tag{28}$$

where

$$\mathfrak{B}_3 = \left[\mathfrak{B}_{31}, \mathfrak{B}_{32}, \mathfrak{B}_{33}, \mathfrak{B}_{34}, \mathfrak{B}_{35}\right] = \left[(1-n+k-p), \left(\frac{1-b}{t}\right), \left(\frac{-b+t}{t}\right), \left(\frac{1+{\xi_{mode}}^{2}}{t}\right), \left(\frac{{\xi_{mode}}^{2}+t}{t}\right)\right]$$

Similar to the asymptotic OP expression, an asymptotic expression for the ABER is obtained by using Eq. (07.34.06.0040.01) provided by Wolfram [40] in Eq. (28) as follows:

$$p^{e\infty} = 0.5 - \frac{q^{p}}{2\lceil p}\frac{\xi_{mode}^{2} t^{a+b-2}}{\lceil a \lceil b\ (2\pi)^{t-1}} \sum_{n=0}^{m-1} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} \left(\frac{m}{\bar{\gamma}_{RF}}\right)^{n-k} \left(q + \frac{m}{\bar{\gamma}_{RF}}\right)^{k-n-p} \sum_{j=1}^{2t+1} \left\{ \frac{\prod_{\substack{i=1 \\ i\neq j}}^{2t+1} \Gamma\left(\mathfrak{B}_{2i} - \mathfrak{B}_{2j}\right) \prod_{i=1}^{t+1} \Gamma\left(1 - \mathfrak{B}_{3i} - \mathfrak{B}_{2j}\right)}{\prod_{i=t+2}^{2t+1} \Gamma\left(\mathfrak{B}_{3i} - \mathfrak{B}_{2j}\right)} \left(\frac{\frac{Cm}{\bar{\gamma}_{RF}\bar{\gamma}_{FSO}} \left(\frac{a\xi_{mode}^{2}}{(b-1)\left(1+\xi_{mode}^{2}\right)}\right)^{t}}{\left(q + \frac{m}{\bar{\gamma}_{RF}}\right)} \right)^{\mathfrak{B}_{2j}} \right\} \tag{29}$$

where, $G_d = \min\left(\frac{a}{t}, \left(\frac{a+t-1}{t}\right), \left(\frac{{\xi_{mode}}^{2}}{t}\right), \left(\frac{{\xi_{mode}}^{2}+t-1}{t}\right), k\right)$ is obtained.

4.3 Ergodic capacity, $\overline{C}_t$

The upper bound on the EC of the considered system related to complementary CDF is given as [21]

$$\overline{C}_t = \frac{\alpha_t}{\ln(2)} \int_0^{\infty} \left(1 + \alpha_t \gamma\right)^{-1} F^{c}{}_{\gamma_{E2E,AF}}(\gamma) d\gamma \tag{30}$$

where $\alpha_t = 1$ and $\frac{e}{2\pi}$ signifies the coherent HD and non-coherent IM/DD detection schemes. $F^{c}(\gamma)$ is the complementary CDF represented as

$$F^{c}{}_{\gamma_{E2E,AF}}(\gamma) = 1 - F_{\gamma_{E2E,AF}}(\gamma) \tag{31}$$

On the substitution of Eq. (19) into Eq. (31), the complementary CDF is obtained as follows:

$$F^{c}_{\gamma_{end-to-end}}(\gamma) = \frac{\xi_{mode}^{2}}{\lceil(a)\lceil(b)} \frac{t^{a+b-2} e^{-\frac{m\gamma}{\bar{\gamma}_{RF}}}}{(2\pi)^{t-1}} \sum_{n=0}^{m-1} \frac{\left(\frac{m\gamma}{\bar{\gamma}_{RF}}\right)^{n}}{n!} \sum_{k=0}^{n} \binom{n}{k} \left(\frac{m\gamma}{\bar{\gamma}_{RF}}\right)^{-k} G_{2t,2t+1}^{2t+1,t}\left(\frac{Cm\gamma_{th}}{\bar{\gamma}_{RF}\bar{\gamma}_{FSO}} \left(\frac{a\xi_{mode}^{2}}{(b-1)\left(1+\xi_{mode}^{2}\right)}\right)^{t} \middle| \begin{matrix} \mathfrak{B}_1 \\ \mathfrak{B}_2 \end{matrix} \right) \tag{32}$$

Further, substituting Eq. (32) into Eq. (30), it can be expressed as:

$$\overline{C}_t = \frac{\alpha_t}{\ln(2)} \frac{\xi_{mode}^{2}}{(a)(b)} \frac{t^{a+b-2}}{(2\pi)^{t-1}} \sum_{n=0}^{m-1} \frac{\left(\frac{m}{\bar{\gamma}_{RF}}\right)^{n}}{n!} \sum_{k=0}^{n} \binom{n}{k} \left(\frac{m}{\bar{\gamma}_{RF}}\right)^{-k} \int_0^{\infty} \left(1 + \alpha_t \gamma\right)^{-1} e^{-\frac{m\gamma}{\bar{\gamma}_{RF}}} \gamma^{n-k} G_{2t,2t+1}^{2t+1,t} \left(\frac{Cm\gamma_{th}}{\bar{\gamma}_{RF}\bar{\gamma}_{FSO}} \left(\frac{a\xi_{mode}^{2}}{(b-1)\left(1+\xi_{mode}^{2}\right)}\right)^{t} \middle| \begin{matrix} \mathfrak{B}_1 \\ \mathfrak{B}_2 \end{matrix} \right) d\gamma \tag{33}$$

Exploiting the two expressions of Eqs. (8.4.3.1) and (8.4.2.5) given as [41] in Eq. (33), it is written as:

$$\overline{C}_t = \frac{\alpha_t}{\ln(2)} \frac{\xi^2_{mode}}{\Gamma(a)\Gamma(b)} \frac{t^{a+b-2}}{(2\pi)^{t-1}} \sum_{n=0}^{m-1} \frac{\left(\frac{m}{\overline{\gamma}_{RF}}\right)^n}{n!} \sum_{k=0}^{n} \binom{n}{k} \left(\frac{m}{\overline{\gamma}_{RF}}\right)^{-k} \gamma^{n-k} G^{1,0}_{0,1}\left(\frac{m\gamma}{\overline{\gamma}_{RF}} \middle| \begin{matrix} - \\ 0 \end{matrix}\right) G^{1,1}_{1,1}\left(\alpha_t\gamma \middle| \begin{matrix} 0 \\ 0 \end{matrix}\right) \int_0^{\infty} G^{2t+1,t}_{2t,2t+1}\left(\frac{C m\gamma}{\overline{\gamma}_{RF}\overline{\gamma}_{FSO}} \left(\frac{a\xi^2_{mode}}{(1+\xi^2_{mode})(b-1)}\right)^t \middle| \begin{matrix} \left(\frac{1-b}{t}\right), \left(\frac{t-b}{t}\right), \left(\frac{1+\xi^2_{mode}}{t}\right), \left(\frac{\xi^2_{mode}+t}{t}\right) \\ \frac{a}{t}, \left(\frac{a+t-1}{t}\right), \left(\frac{\xi^2_{mode}}{t}\right), \left(\frac{\xi^2_{mode}+t-1}{t}\right), k \end{matrix}\right) d\gamma \tag{34}$$

With the aid of Eq. (07.34.21.0081.01) of [40] in Eq. (34) which results the ergodic capacity in the closed-form expression as follows:

$$\overline{C}_t = \frac{\alpha_t}{\ln(2)} \frac{\xi^2_{mode}}{\Gamma(a)\Gamma(b)} \frac{t^{a+b-2}}{(2\pi)^{t-1}} \sum_{n=0}^{m-1} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} \left(\frac{m}{\overline{\gamma}_{RF}}\right)^{-1} G^{0,1:1,1:2t+1,t}_{1,0:1,1:2t,2t+1}\left(\begin{matrix} -n+k \\ - \end{matrix} \middle| \begin{matrix} 0 \\ 0 \end{matrix} \middle| \begin{matrix} \left(\frac{1-b}{t}\right), \left(\frac{t-b}{t}\right), \left(\frac{1+\xi^2_{mode}}{t}\right), \left(\frac{\xi^2_{mode}+t}{t}\right) \\ \frac{a}{t}, \left(\frac{a+t-1}{t}\right), \left(\frac{\xi^2_{mode}}{t}\right), \left(\frac{\xi^2_{mode}+t-1}{t}\right), k \end{matrix} \middle| \frac{\alpha_t\overline{\gamma}_{RF}}{m}, C\left(\frac{a\xi^2_{mode}}{(1+\xi^2_{mode})(b-1)\overline{\gamma}^{1/t}_{FSO}}\right)^t \right) \tag{35}$$

5 Numerical results and discussion

In this section, the performance of the considered system is analyzed for both detection schemes under the influence of atmospheric turbulence and pointing errors. For this purpose, the following parameters are considered in accordance [32, 34]. For the RF channel, $m = 2$ has been taken. The transmission range, $L = 3.5$ km, and optical wavelength, $\lambda = 1550$ nm are chosen for the FSO link. The refraction index structure parameter,$C_n^2 = 5 \times 10^{-15}$ $m^{-\frac{2}{3}}$, 2×10^{-14} $m^{-\frac{2}{3}}$ and 8×10^{-14} $m^{-\frac{2}{3}}$ are considered for weak, moderate, and strong turbulent environments respectively. The Rytov variance,$\sigma_1^2 = 0.4023$, 1.6092, and 6.4369 are obtained using $\sigma_1^2 = 0.5C_n^2 k^{7/6} L^{11/6}$, where$k = \frac{2\pi}{\lambda}$. The corresponding values of $(a, b) = (4.5916, 7.0941)$, (a, b) = (1.8276, 3.9363), and (a, b) = (1.4262, 3.5297) are calculated from Eq. (9) for weak, moderate, and strong turbulent environments respectively, taking into account the inner scale of turbulence, $l_0 = 5.98$ mm, outer scale of turbulence,$L_0 = 0.8$ m, and receiver aperture diameter$D = 2g = 10$ mm . The average SNR for the different hop, i.e. $\overline{\gamma}_{RF} = \overline{\gamma}_{FSO} = \overline{\gamma}$ and fixed gain, C = 1 are considered. Moreover, the agreement between asymptotic and exact analytic expressions offers valuable insights into the system's performance. The results of the considered system are compared with those available in the literature [34].

The exact and asymptotic expressions for OP given by Eqs. (22) and (23) are demonstrated in Figs. 2, 3, and 4 and are plotted against average SNR, $\overline{\gamma}$ for both detection schemes at $\gamma_{th} = 0$dB.

Figure 2 shows how turbulence conditions change the OP performance while keeping other parameters fixed at $w_b/g = 10$, $(\sigma_x/g, \sigma_y/g) = (8, 1)$, $(\mu_x/g, \mu_y/g) = (3, 1)$ and $\xi_{mode} = 0.6650$. The findings demonstrate that atmospheric turbulence severity put a negative impact on the system performance for both detection schemes. Additionally, it finds that OP with $t = 1$ outperforms with $t = 2$ indicating that HD detection scheme performs better than IM/DD detection scheme. Further, it can also be seen that AF relaying scheme is better as compared to DF relaying scheme.

How the behaviour of jitters affect the OP performance, Fig. 3 shows the impacts of $(\sigma_x/g, \sigma_y/g)$ with fixed values of $w_b/g = 10$, and $(\mu_x/g, \mu_y/g) = (3, 1)$ under strong

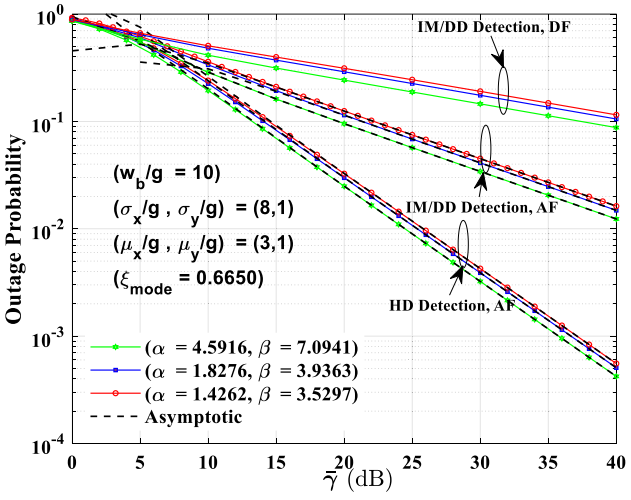
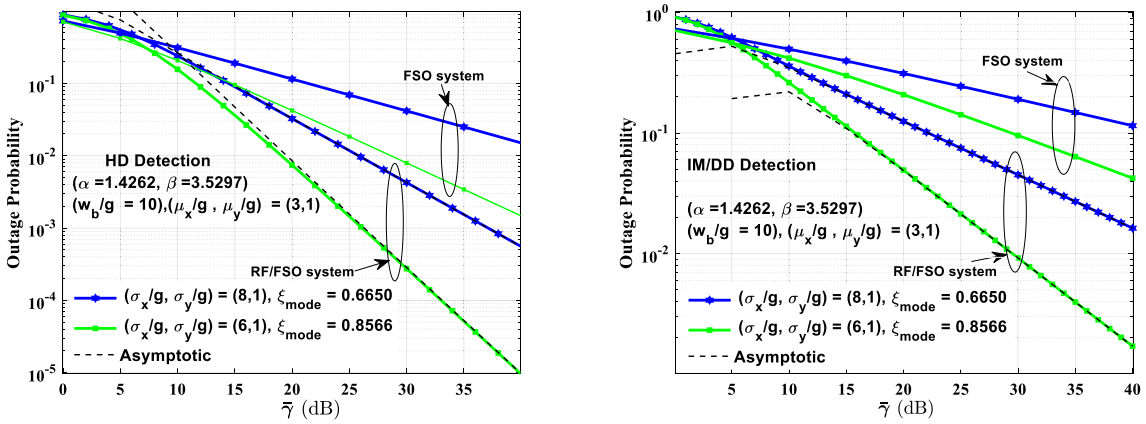


Fig. 2 OP for different atmospheric turbulence regimes considering both detection schemes using AF and DF relaying techniques



(a) HD detection scheme (b) IM/DD detection scheme

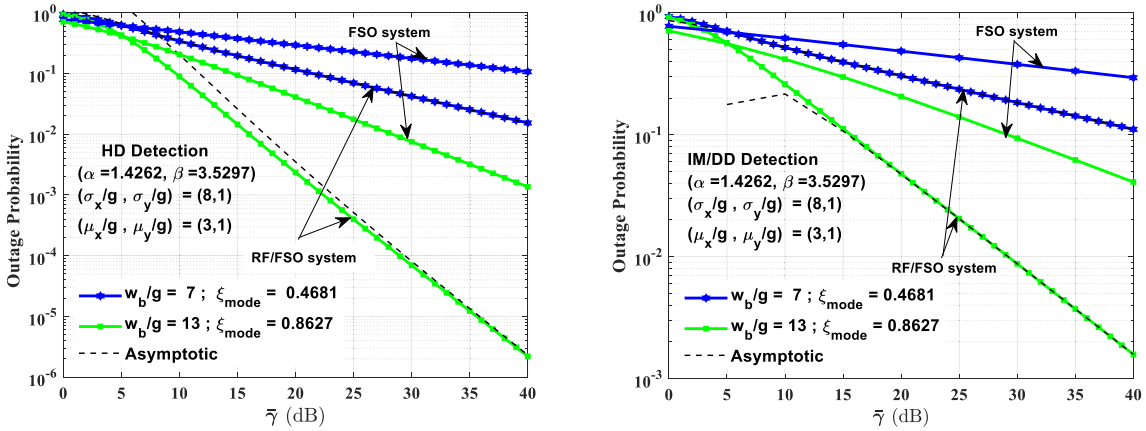
Fig. 3 OP for different values of $(\sigma_x/g, \sigma_y/g)$ under strong turbulence conditions

turbulence for both detection schemes. The results indicate that higher the value of jitters corresponding value of ξ_{mode} decreases which deteriorates the OP performance as shown in Table 2. It also shows that the impact of jitters is less on the OP of dual hop RF/FSO system than that of a single FSO system given in literature [34].

The effect of beamwidth, w_b on the OP at settings of $(\sigma_x/g, \sigma_y/g) = (8,1)$, and $(\mu_x/g, \mu_y/g) = (3,1)$ presents in Fig. 4 under strong turbulence for both detection schemes. The results indicate that a smaller value of w_b/g, corresponding value of ξ_{mode} decreases which depreciates the OP performance as presented in Table 3.

Exact and asymptotic expressions for ABER given in Eqs. (28) and (29) are plotted against average SNR, $\overline{\gamma}$ for both detection schemes as presented in Figs. 5, 6, 7, and 8. The ABER is demonstrated in Fig. 5 for the setting of $w_b/g = 10$, $(\sigma_x/g, \sigma_y/g) = (8,1)$, $(\mu_x/g, \mu_y/g) = (3,1)$ and $\xi_{mode} = 0.6650$ under different atmospheric turbulence conditions using the CBPSK modulation scheme. It exhibits that the system bit error rate performance becomes poor as the turbulence condition turns stronger for both detection schemes. It is also seen that HD improves ABER performance in comparison to IM/DD.

Figure 6 portrays the ABER for $w_b/g = 10$, $(\sigma_x/g, \sigma_y/g) = (8,1)$, $(\mu_x/g, \mu_y/g) = (3,1)$ and $\xi_{mode} = 0.6650$ using CBPSK, CBFSK, DBPSK, RZ-OOK, NBFSK, and NRZ-OOK modulations scheme under strong atmospheric turbulence conditions in both detection schemes. It is also reported that ABER with CBPSK modulation outperforms other modulations for both detection schemes. For instance, at $\overline{\gamma} = 30$ dB, ABER is 6.960×10^{-4} ,9.490×10^{-4}, 1.056×10^{-3}, 1.301×10^{-3}, $1.443 \times 10^{-3}$and 1.808×10^{-3} for CBPSK, CBFSK, DPSK, RZ-OOK, NBFSK, and NRZ-OOK respectively for HD and



(a) HD detection scheme (b) IM/DD detection scheme

Fig. 4 OP for different values of w_b/g under strong turbulence conditions

Table 2 OP under various values of jitters, $(\sigma_x/g, \sigma_y/g)$.

Parameter values	System			
	OP for HD detection scheme, Fig. 3a		OP for IM/DD detection scheme, Fig. 3b	
$(\mu_x/g, \mu_y/g) = (3,1)$, $w_b/g = 10, \overline{\gamma} = 30$ dB, Strong Turbulence	Dual-hop system	Single FSO system	Dual-hop system	Single FSO system
$(\sigma_x/g, \sigma_y/g) = (8,1)$, $\xi_{mode} = 0.6650$	2.379×10^{-3}	4.166×10^{-2}	4.499×10^{-2}	1.906×10^{-1}
$(\sigma_x/g, \sigma_y/g) = (6,1)$, $\xi_{mode} = 0.8566$	7.780×10^{-5}	7.944×10^{-3}	9.197×10^{-3}	9.553×10^{-2}

Table 3 OP under various values of beamwidth, w_b/g

Parameter values	System			
	OP for HD detection scheme, Fig. 4a		OP for IM/DD detection scheme, Fig. 4b	
$(\sigma_x/g, \sigma_y/g) = (8,1)$,$(\mu_x/g, \mu_y/g) = (3,1)$, $\overline{\gamma} = 30$ dB, Strong Turbulence	Dual-hop system	Single FSO system	Dual-hop system	Single FSO system
$w_b/g = 7, \xi_{mode} = 0.4681$	4.210×10^{-2}	1.772×10^{-1}	1.833×10^{-1}	3.775×10^{-1}
$w_b/g = 13, \xi_{mode} = 0.8627$	6.943×10^{-5}	7.521×10^{-3}	8.731×10^{-3}	9.351×10^{-2}

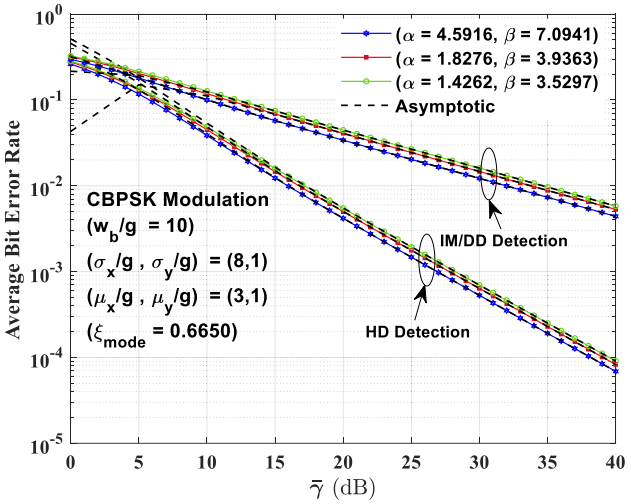


Fig. 5 ABER for different atmospheric turbulence conditions using HD and IM/DD detection schemes with CBPSK modulation

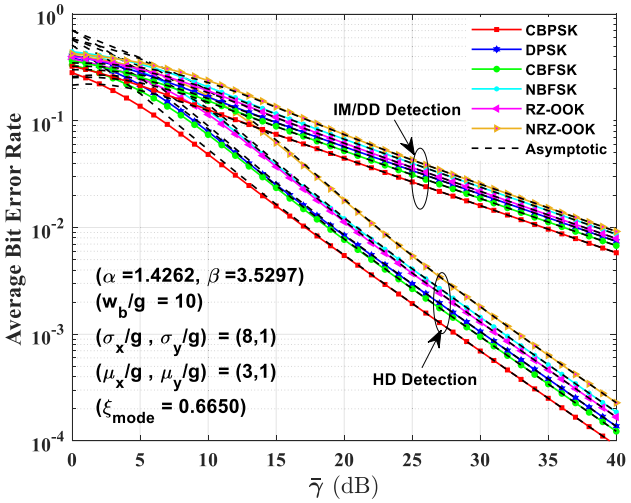


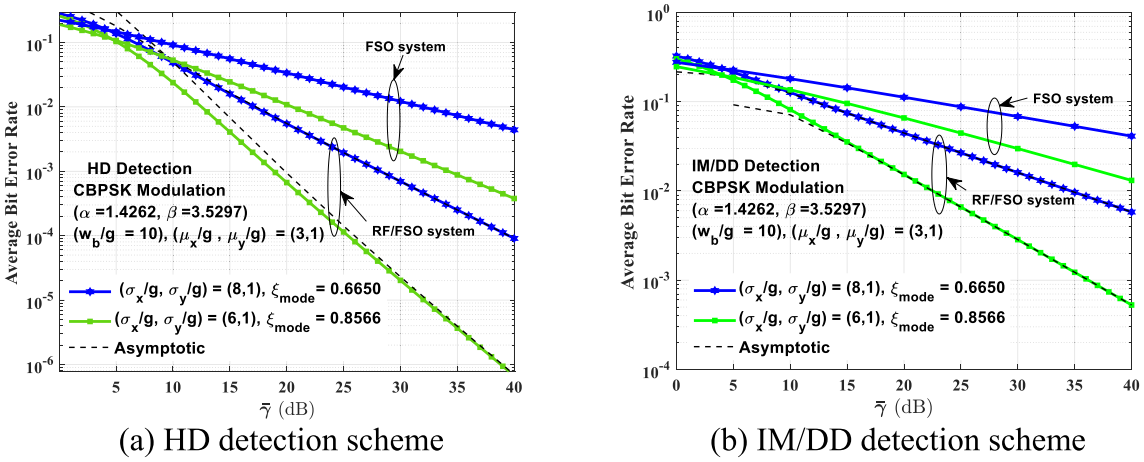
Fig. 6 ABER for various CBPSK, DPSK, CBFSK, NBFSK, RZ-OOK, and NRZ- OOK modulations under strong turbulence conditions with both detection schemes

$1.606 \times 10^{-2}, 1.873 \times 10^{-2}, 2.054 \times 10^{-2}, 2.186 \times 10^{-2}$, $2.396 \times 10^{-2}$and 2.554×10^{-2} for CBPSK, CBFSK, DPSK, RZ-OOK, NBFSK, and NRZ-OOK respectively for IM/DD detection scheme.

Figure 7 presents the influence of $(\sigma_x/g, \sigma_y/g)$ on the ABER at $w_b/g = 10$, and $(\mu_x/g, \mu_y/g) = (3,1)$ under strong turbulence for both detection schemes with CBPSK modulation. It indicates that the value of jitters corresponds smaller value of ξ_{mode} result in the degradation of ABER performance. However, as highlighted in Table 4, the impact of pointing error for a single FSO link is larger than the considered system.

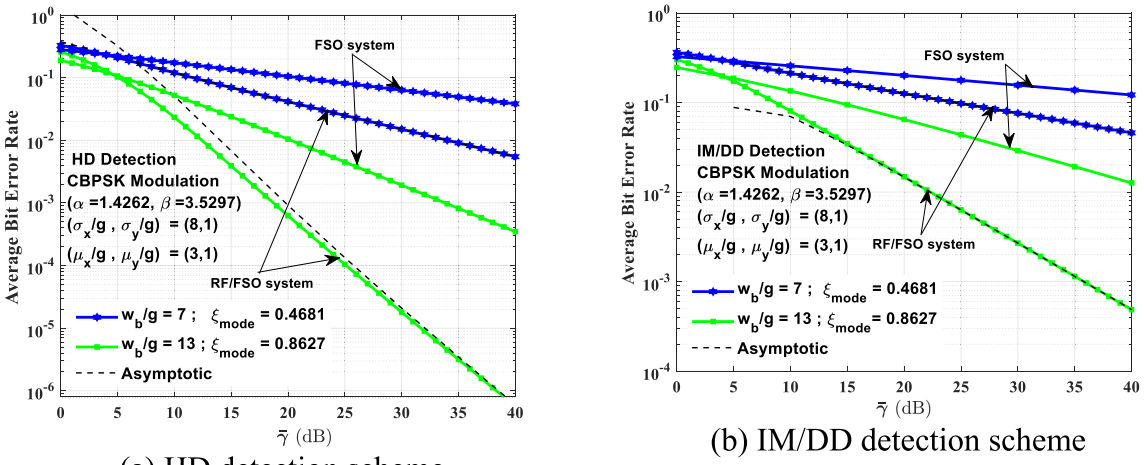
Figure 8 depicts the impact of w_b on the ABER for fixed values of $(\sigma_x/g, \sigma_y/g) = (8,1)$ and $(\mu_x/g, \mu_y/g) = (3,1)$ with CBPSK modulation. The result implies that the ABER performance goes worse with the small value of w_b/g corresponding ξ_{mode} decreases. ABER performance for the single FSO system with a pointing error is less than the concerned system as provided in Table 5.

EC of the considered system given in Eq. (35) is plotted against average electrical SNR, $\overline{\gamma}$ for both detection schemes as displayed in Figs. 9, 10, and 11. The EC of the concerned system is shown in Fig. 9 at fixed



(a) HD detection scheme (b) IM/DD detection scheme

Fig. 7 ABER for different values of $(\sigma_x/g, \sigma_y/g)$ under strong turbulence conditions with CBPSK modulation



(a) HD detection scheme (b) IM/DD detection scheme

Fig. 8 ABER for different values of w_b/g under strong turbulence conditions

Table 4 ABER under various values of $(\sigma_x/g, \sigma_y/g)$.

Parameter values	System with CBPSK modulation			
	ABER for HD detection scheme, Fig. 7a		ABER for IM/DD detection scheme, Fig. 7b	
$(\mu_x/g, \mu_y/g) = (3,1)$, $w_b/g = 10, \bar{\gamma} = 30$ dB, Strong Turbulence	Dual-hop system	Single FSO system	Dual-hop system	Single FSO system
$(\sigma_x/g, \sigma_y/g) = (8,1), \xi_{mode} = 0.6650$	6.960×10^{-4}	1.219×10^{-2}	1.606×10^{-2}	6.819×10^{-2}
$(\sigma_x/g, \sigma_y/g) = (6,1), \xi_{mode} = 0.8566$	2.023×10^{-5}	2.039×10^{-3}	2.848×10^{-3}	2.982×10^{-2}

$w_b/g = 10$, $(\sigma_x/g, \sigma_y/g) = (8,1)$, $(\mu_x/g, \mu_y/g) = (3,1)$ and $\xi_{mode} = 0.6650$ by varying atmospheric turbulence conditions. It tells that as the turbulence level increases, EC performance decreases. Furthermore, the concerned system for the HD detection scheme performs much better than IM/DD detection scheme.

Figure 10 depicts the influence of $\left(\sigma_x/g, \sigma_y/g\right)$ on the ergodic capacity for $w_b/g = 10$ and $(\mu_x/g, \mu_y/g) = (3,1)$

Table 5 ABER under various values of w_b/g.

Parameter values	System with CBPSK modulation			
	ABER for HD detection scheme, Fig. 8a		ABER for IM/DD detection scheme, Fig. 8b	
$(\sigma_x/g, \sigma_y/g) = (8,1)$,$(\mu_x/g, \mu_y/g) = (3,1)$, $\overline{\gamma} = 30$ dB, Strong Turbulence	Dual-hop system	Single FSO system	Dual-hop system	Single FSO system
$w_b/g = 7, \xi_{mode} = 0.4681$	1.507×10^{-2}	6.344×10^{-2}	7.590×10^{-2}	1.563×10^{-1}
$w_b/g = 13, \xi_{mode} = 0.8627$	1.803×10^{-5}	1.922×10^{-3}	2.693×10^{-3}	2.908×10^{-2}

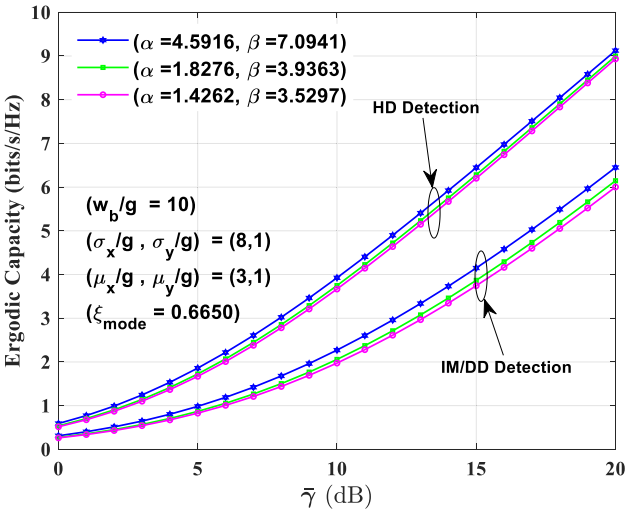


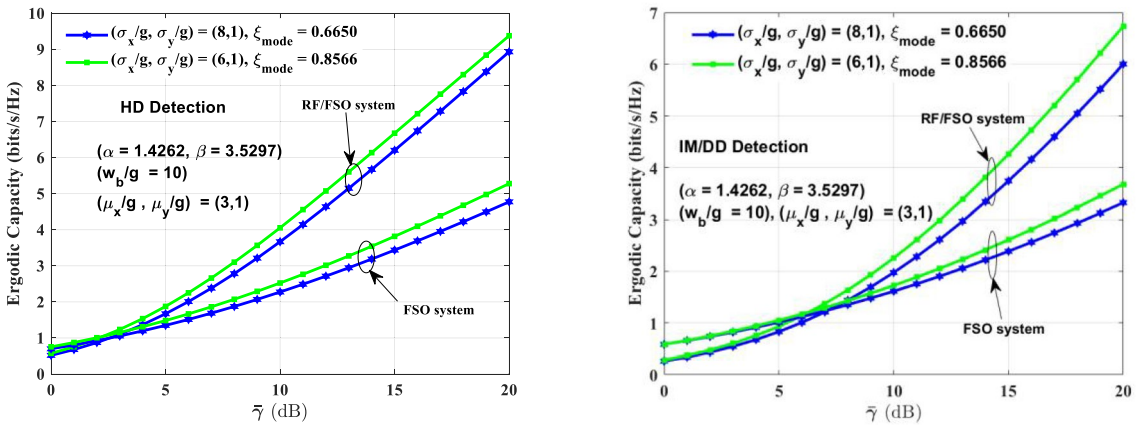
Fig. 9 EC under different atmospheric turbulence conditions for both detection schemes

under strong turbulence for both detection schemes. Similar to the OP and ABER performance, EC is significantly impacted by jitters. Furthermore, as reported in Table 6, It indicates that a single FSO system is more influenced than the concerned system.

Figure 11 demonstrates the impact of w_b on the system's EC for the fixed values of $(\sigma_x/g, \sigma_y/g) = (8,1)$ and $(\mu_x/g, \mu_y/g) = (3,1)$ under strong turbulence for both detection schemes. It suggests that the EC performance goes better with the increase of w_b/g and ξ_{mode} in both detection schemes. Also, EC for a single FSO system with pointing error is inferior to the concerned system as outlined in Table 7.

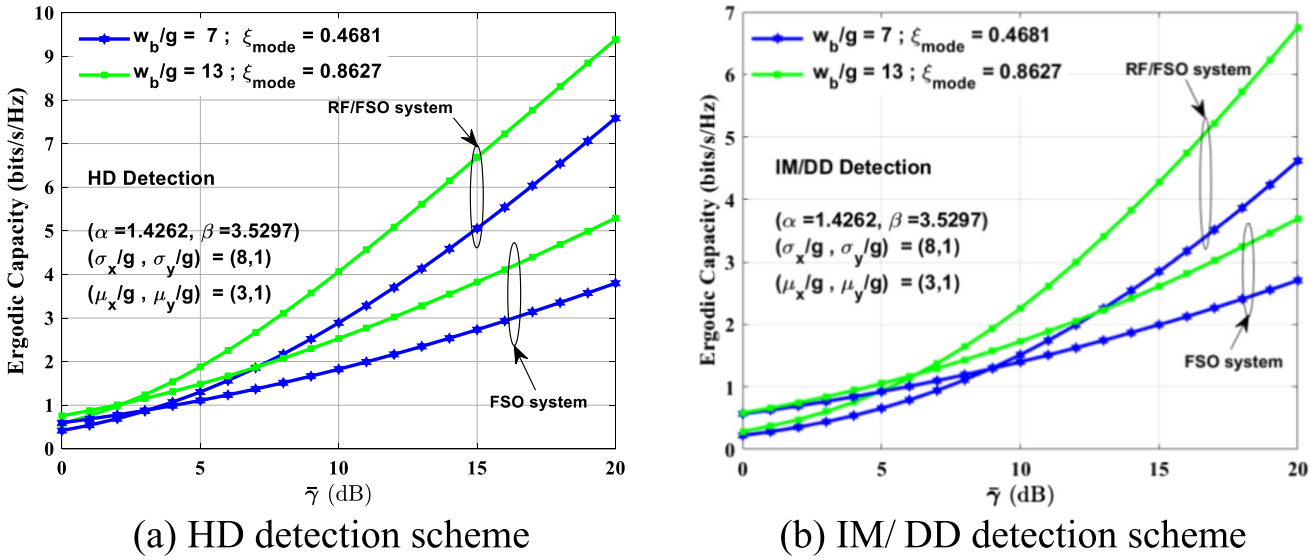
6 Conclusion

In this paper, analysis was done on the effectiveness of an AF-based relay dual-hop RF/WOC system for different detection schemes where RF and FSO links are distinguished by the Nakagami-m and $\mathcal{F}$ distributions with pointing errors respectively. An exact closed-form mathematical representation of the CDF for overall instantaneous received SNR was formulated. Subsequently, novel exact and asymptotic expressions for performance metrics were obtained. The numerical results indicated that the turbulent severity in the environment



(a) HD detection scheme (b) IM/DD detection scheme

Fig. 10 EC under different values of $(\sigma_x/g, \sigma_y/g)$ for strong turbulence conditions



(a) HD detection scheme (b) IM/ DD detection scheme

Fig. 11 EC for different values of w_b/g under strong turbulence conditions

Table 6 EC for various values of jitters $(\sigma_x/g, \sigma_y/g)$.

Parameter values	System			
	EC (bits/s/Hz) for HD detection scheme, Fig. 10a		EC (bits/s/Hz) for IM/DD detection scheme, Fig. 10b	
$(\mu_x/g, \mu_y/g) = (3, 1)$, $w_b/g = 10, \overline{\gamma} = 15$ dB, Strong Turbulence	Dual-hop system	Single FSO system	Dual-hop system	Single FSO system
$(\sigma_x/g, \sigma_y/g) = (8,1)$, $\xi_{mode} = 0.6650$	6.202	3.436	3.744	2.387
$(\sigma_x/g, \sigma_y/g) = (6,1)$, $\xi_{mode} = 0.8566$	6.676	3.817	4.263	2.608

Table 7 EC for various values of w_b/g.

Parameter values	Systems			
	EC (bits/s/Hz) for HD detection scheme, Fig. 11a		EC (bits/s/Hz) for IM/DD detection scheme, Fig. 11b	
$(\sigma_x/g, \sigma_y/g) = (8,1)$, $(\mu_x/g, \mu_y/g) = (3, 1)$, $\overline{\gamma} = 15$ dB, Strong Turbulence	Dual-hop system	Single FSO system	Dual-hop system	Single FSO system
$w_b/g = 7, \xi_{mode} = 0.4681$	5.055	2.731	2.849	1.995
$w_b/g = 13, \xi_{mode} = 0.8627$	6.685	3.826	4.275	2.613

and the pointing error degraded the performance. Moreover, ABER with CBPSK modulation was found better than other binary modulation schemes. Furthermore, it was also seen that a coherent HD detection scheme-based system exhibits better system performance than a non-coherent IM/DD detection scheme. Additionally, this work was contrasted with a single FSO system using the same parameters. In comparison to a single FSO system, it was found that the proposed system tends to ameliorate the consequences of atmospheric turbulence and pointing error.

Author contributions Rekha Rani, M. Lakshmanan, Anup Mandpura, N. Jayanthi have analyzed the system performance by using Matlab and Mathematica and prepared the manuscript.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

Ethical approval NA.

References

1. Khalighi, M.A., Uysal, M.: Survey on free space optical communication: a communication theory perspective. IEEE Commun. Surv. Tutorials **16**(4), 2231–2258 (2014). https://doi.org/10.1109/COMST.2014.2329501
2. Rani, R., Kaur, G. Design and analysis of MIMO FSO system and WDM FSO system for Leh (Ladakh), India under worst weather conditions. In 2021 7th International Conference on Signal Processing and Communication (ICSC), pp. 31–36. https://doi.org/10.1109/ICSC53193.2021.9673278 (2021)
3. Son, I.K., Mao, S.: A survey of free space optical networks. Digit. Commun. Networks **3**(2), 67–77 (2017). https://doi.org/10.1016/j.dcan.2016.11.002
4. Mansour, A., Mesleh, R., Abaza, M.: New challenges in wireless and free space optical communications. Opt. Lasers Eng. **89**, 95–108 (2016). https://doi.org/10.1016/j.optlaseng.2016.03.027
5. Odeyemi, K.O., Aiyetoro, G., Owolawi, P.A., Olakanmi, O.O.: Performance analysis of reconfigurable intelligent surface in a dual-hop DF relay empowered asymmetric RF/FSO networks. Opt. Quantum Electron. **53**(11), 621 (2021). https://doi.org/10.1007/s11082-021-03277-0
6. Goel, A., Bhatia, R.: Joint impact of interference and hardware impairments on the performance of mixed RF/FSO cooperative relay networks. Opt. Quantum Electron. **53**(9), 530 (2021). https://doi.org/10.1007/s11082-021-03064-x
7. Anees, S., Bhatnagar, M.: Performance of an amplify-and-forward dual-hop asymmetric RF-FSO communication system. J. Opt. Commun. Netw. **7**(2), 124–135 (2015). https://doi.org/10.1364/JOCN.7.000124
8. Xu, G., Song, Z.: Performance analysis for mixed κ-μ fading and M-distribution dual-hop radio frequency/free space optical communication systems. IEEE Trans. Wirel. Commun. **20**(3), 1517–1528 (2021). https://doi.org/10.1109/TWC.2020.3034104
9. Nguyen, B.C., Le, X.H., Nguyen, V.D., Dung, L.T.: On the capacity of full-duplex AF/DF relay system with energy harvesting for vehicle-to-vehicle communications. Wirel. Commun. Mob. Comput. (2021). https://doi.org/10.1155/2021/8865615
10. Yacoub, M.D.: The κ-μ distribution and the η-μ distribution. IEEE Antennas Propag. Mag. **49**(1), 68–81 (2007)
11. Anbarasi, K., Hemanth, C., Sangeetha, R.G.: A review on channel models in free space optical communication systems. Opt. Laser Technol. **97**, 161–171 (2017). https://doi.org/10.1016/j.optlastec.2017.06.018
12. Ansari, I.S., Yilmaz, F., Alouini, M.S.: Performance analysis of free-space optical links over Málaga (M) turbulence channels with pointing errors. IEEE Trans. Wirel. Commun. **15**(1), 91–102 (2016). https://doi.org/10.1109/TWC.2015.2467386
13. Barrios Porras, R. Exponentiated Weibull fading channel model in free-space optical communications under atmospheric turbulence. (2013).
14. Kashani, M.A., Uysal, M., Kavehrad, M.: A novel statistical channel model for turbulence-induced fading in free-space optical systems. J. Light. Technol. **33**(11), 2303–2312 (2015). https://doi.org/10.1109/JLT.2015.2410695
15. Peppas, K.P., Alexandropoulos, G.C., Xenos, E.D., Maras, A.: The fischer-Snedecor F-distribution model for turbulence-induced fading in free-space optical systems. J. Light. Technol. **38**(6), 1286–1295 (2020). https://doi.org/10.1109/JLT.2019.2957327
16. Odeyemi, K.O., Owolawi, P.A.: Partial relay selection in mixed RF/FSO dual-hop system over unified M-distributed fading channel with non-zero boresight pointing errors. Opt. Quantum Electron. **51**(5), 141 (2019). https://doi.org/10.1007/s11082-019-1863-3
17. Farid, A.A., Hranilovic, S.: Outage capacity optimization for free-space optical links with pointing errors. J. Light. Technol. **25**(7), 1702–1710 (2007). https://doi.org/10.1109/JLT.2007.899174
18. Gappmair, W., Hranilovic, S., Leitgeb, E.: OOK performance for terrestrial FSO links in turbulent atmosphere with pointing errors modeled by Hoyt distributions. IEEE Commun. Lett. **15**(8), 875–877 (2011). https://doi.org/10.1109/LCOMM.2011.062911.102083
19. Yang, F., Cheng, J., Tsiftsis, T.A.: Free-space optical communication with nonzero boresight pointing errors. IEEE Trans. Commun. **62**(2), 713–725 (2014)
20. Boluda-Ruiz, R., García-Zambrana, A., Castillo-Vázquez, C., Castillo-Vázquez, B.: Novel approximation of misalignment fading modeled by Beckmann distribution on free-space optical links. Opt. Express **24**(20), 22635 (2016). https://doi.org/10.1364/oe.24.022635
21. Ansari, I.S., Yilmaz, F., Alouini, M.S.: Impact of pointing errors on the performance of mixed RF/FSO dual-hop transmission systems. IEEE Wirel. Commun. Lett. **2**(3), 351–354 (2013). https://doi.org/10.1109/WCL.2013.042313.130138
22. Zedini, E., Member, S., Imran Shafique Ansari, I., Mohamed-Slim Alouini, I., Shafique Ansari, I., Alouini, M.-S.: Performance analysis of mixed Nakagami-m and gamma-gamma dual-hop FSO transmission systems. IEEE Photonics J. **7**(1), 1–20 (2015). https://doi.org/10.1109/JPHOT.2014.2381657
23. Sharma, N., Bansal, A., Garg, P.: Decode-and-forward relaying in mixed η−μ and gamma–gamma dual hop transmission system. IET Commun. **10**(14), 1769–1776 (2016)
24. Samimi, H., Uysal, M.: End-to-end performance of mixed RF/FSO transmission systems. J. Opt. Commun. Netw. **5**(11), 1139–1144 (2013)
25. Trinh, P.V., Cong Thang, T., Pham, A.T.: Mixed mmwave RF/FSO relaying systems over generalized fading channels with pointing errors. IEEE Photonics J. **9**(1), 1–14 (2017). https://doi.org/10.1109/JPHOT.2016.2644964
26. Palliyembil, V., Vellakudiyan, J., Muthuchidambaranathan, P.: Performance analysis of RF-FSO communication systems over the Málaga distribution channel with pointing error. Optik (Stuttg) **247**(August), 167891 (2021). https://doi.org/10.1016/j.ijleo.2021.167891
27. Jing, Z., Shang-hong, Z., Wei-hu, Z., Ke-fan, C.: Performance analysis for mixed FSO/RF Nakagami-m and exponentiated Weibull dual-hop airborne systems. Opt. Commun. **392**(December 2016), 294–299 (2017). https://doi.org/10.1016/j.optcom.2017.01.033
28. Wang, Y., Wang, P., Liu, X., Cao, T.: On the performance of dual-hop mixed RF/FSO wireless communication system in urban area over aggregated exponentiated Weibull fading channels with pointing errors. Opt. Commun. **410**, 609–616 (2018). https://doi.org/10.1016/J.OPTCOM.2017.10.062
29. AlQuwaiee, H., Ansari, I.S., Alouini, M.S.: On the performance of free-space optical communication systems over double generalized gamma channel. IEEE J. Sel. Areas Commun. **33**(9), 1829–1840 (2015). https://doi.org/10.1109/JSAC.2015.2432529
30. Samimi, H.: Outage analysis of mixed dual-hop RF-FSO communication system over fading channels with pointing errors. Wirel. Pers. Commun. **109**(3), 1557–1569 (2019). https://doi.org/10.1007/s11277-019-06627-0

31. Han, L., Liu, X., Wang, Y., Li, B.: Performance analysis of mixed Rayleigh and F distribution RF-FSO cooperative systems with AF relaying. Electronics **11**(15), 2299 (2022)
32. Ding, J., Xie, X., Tan, L., Ma, J., Kang, D.: Dual-Hop RF/FSO systems over κ-μ shadowed and fisher-snedecor f fading channels with non-zero boresight pointing errors. J. Light. Technol. **40**(3), 708–719 (2022). https://doi.org/10.1109/JLT.2021.3120767
33. Prabu, K., Kumar, D.S., Srinivas, T.: Performance analysis of FSO links under strong atmospheric turbulence conditions using various modulation schemes. Optik (Stuttg) **125**(19), 5573–5581 (2014)
34. Badarneh, O.S., Derbas, R., Almehmadi, F.S., El Bouanani, F., Muhaidat, S.: Performance analysis of FSO communications over F turbulence channels with pointing errors. IEEE Commun. Lett. **25**(3), 926–930 (2021). https://doi.org/10.1109/LCOMM.2020.3042489
35. Khalighi, M.A., Schwartz, N., Aitamer, N., Bourennane, S.: Fading reduction by aperture averaging and spatial diversity in optical wireless systems. J. Opt. Commun. Netw. **1**(6), 580–593 (2009). https://doi.org/10.1364/JOCN.1.000580
36. Pattanayak, D.R., Rai, S., Dwivedi, V.K., Singh, G.: A statistical channel model for a decode-and-forward based dual hop mixed RF/FSO relay network. Opt. Quantum Electron. **50**(6), 1–17 (2018). https://doi.org/10.1007/s11082-018-1486-0
37. Ansari, I.S., Alouini, M.S., Cheng, J.: Ergodic capacity analysis of free-space optical links with nonzero boresight pointing errors. IEEE Trans. Wirel. Commun. **14**(8), 4248–4264 (2015). https://doi.org/10.1109/TWC.2015.2418285
38. Siddharth, M., Shah, S., Vishwakarma, N., Swaminathan, R.: Performance analysis of adaptive combining based hybrid FSO/RF terrestrial communication. IET Commun. **14**(22), 4057–4068 (2020). https://doi.org/10.1049/iet-com.2020.0598
39. Gradshteyn, I.S., Ryzhik, I.M.: Table of Contents for Table of Integrals, Series, and Products. Academic press, Cambridge (2015)
40. Wolfram, The Wolfram Functions Site. Online https://functions.wolfram.com/ (2004)
41. Prudnikov, A.P., Brychkov, I. A., Marichev, O. I.: Integrals and series: special functions. vol. 2, CRC press (1986)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.



Rekha Rani is currently pursuing the Ph.D. degree in Electronics and Communication Engineering from Delhi Technological University, Delhi. She has 8 years of experience in academics and industrial research. She has number of research publications and book chapters. Her research interests include the Wireless Communication, Channel Modeling, Optimal Power Allocation, Channel Estimation and Free Space Optical Wireless Communication.



Dr. M. Lakshmanan is a highly accomplished academician and researcher with over 22 years of teaching experience and 15 years of research expertise. He holds a B.E. in Electronics and Communication Engineering from Madras University, an M.E. in Communication Systems from Anna University, and a Ph.D. from VIT University. Currently, he serves as Professor and Dean IQAC at Sri Eshwar College of Engineering, Coimbatore, Tamilnadu, India. Dr. Lakshmanan is a recognized expert in NBA, NAAC & ABET Accreditation, and NIRF & QS Ranking. He has mentored numerous professional institutions and universities across the nation to achieve the highest grades and ranks in accreditation and rankings respectively. With a Senior Member status in IEEE and UACEE, and fellow in IETE and IEI, he also holds Life Memberships in ISTE, ACM, IACSIT, and IAENG. Dr. Lakshmanan is an Executive Committee Member of IETE Noida Section since July 2022. His extensive research contributions include over 100 papers published in reputed SCI/SCIE/ESCI and Scopus indexed journals, along with 4 patents published in India and 2 patents granted in Australia. He is a sought-after speaker, having delivered keynote talks and invited lectures in various FDPs, workshops, and seminars. Actively involved as a reviewer for SCI/SCIE and Scopus indexed journals, he has reviewed more than 200 papers. Dr. Lakshmanan has received multiple prestigious awards, including the Best Administrator Award and Best Researcher Award for 2024 awarded by The Innovative and Advanced Multidisciplinary Research Foundation (IAMRF) in collaboration with the Prerna Education Research Society, Award of Excellence in Research for 2022-23 in National Faculty Award, Researcher of the Year Award (Engineering, Male) for 2021 in 3rd International Business & Academic Excellence Awards, Best Researcher Award for 2021 in International Excellence Award, and the Award of Emerging Leader for 2018 in Higher Education Leadership Meet – 2018 (HELA 2018). Additionally, he has been honored with the Active Researcher award five times from VIT University. His current research interests span Wireless Communication and Networks, Security Algorithms, Device to Device Communications, Digital Communication, Information Theory and Coding, Signal and Image Processing, Artificial Intelligence, Deep Learning, and Machine Learning.



Anup Kumar Mandpura (Member IEEE) received the B.Tech degree in Electronics Engineering from Banaras Hindu University, Varanasi, India in 2006,M. Tech degree in Digital Signal Processing from the Indian Institute of Technology, Guwahati, India in 2009 and the Ph.D. degree in wireless communication, from the Indian Institute of Technology, Delhi, India in 2018. He is currently working with the Department of Electrical Engineering, Delhi Technological University, Delhi, India as an Assistant Professor. His research interests include performance analysis of 5G and beyond wireless communication systems, energy harvesting communication systems, microwave/millimeter-wave systems and signal processing and machine learning with application to Control Systems, PV systems and Seismic data.



Dr. N. Jayanthi is an Assistant Professor in the department of Electronics and Communication Engineering at Delhi Technological University, since 2007. She has done B.E. in Electronics and Communication Engineering from Bharathidasan University, Trichy, Tamil Nadu and M. Tech. in Communication Systems from National Institute of Technology, Trichy. She has completed her Ph.D. thesis titled "Analysis of Digitized Historical Inscriptions and Manuscripts at Delhi Technological University. Her area of research interest includes image processing, Pattern recognition, machine learning and Communication systems. She has published a number of technical papers in International Conferences and reputed Journals. She has completed two DST sponsored projects titled "Development of OGC standards-based sensor network for intelligent traffic management" and "Development of frame work for imaging, restoring and archiving inscription and manuscript" in the capacity of co-principal investigator. Currently she is working on the Project titled "Digitizing manuscripts at the Indira Gandhi National Centre for the arts and facilitating online access to digitized documents ", in collaboration with iHub Drishti Foundation, IIT Jodhpur. She is member in professional societies such as IEEE and Intelligent Transport system (ITS) and also life member in Pattern Recognition and Artificial Intelligence (IUPARI). She is secretary of IEEE intelligent transportation systems society for Delhi chapter. Also, she is Co-Chair of technical council DTU.

Asia-Pacific Financial Markets
https://doi.org/10.1007/s10690-025-09525-7

ORIGINAL RESEARCH

Predicting Financial Distress Using Machine Learning Techniques

Pallavi Sethi[1] · **Archana Singh**[1] · **Vikas Gupta**[1]

Accepted: 5 March 2025
© The Author(s), under exclusive licence to Springer Japan KK, part of Springer Nature 2025

Abstract

Bankruptcy in future would lead to heavy losses, and attempts should be made to reduce it and prevent such a loss in advance. Potential misclassification of potential and futurist bankruptcy can be referred to as an audit failure. Predicting bankruptcy for different users, including investors, auditors, creditors and regulators, is essential. Various prediction models have been considered in other studies, and this research study includes multiple techniques like random forest, ANN, and logistic regression. The main reason behind the prediction of financial distress is that it will help ensure an increase in compatibility with the decision-making process. In this study, 1111 companies were considered liquidated and restructured by NCLT since its inception, i.e., 2016. Of these, 342 companies had their resolution plan approved from 2017 to 18 till December 2023, while 769 companies were liquidated. The selected companies' financial information has been considered for the last five years, and machine learning (Random forest, Artificial neural network) and traditional models (Logistic regression) have been used to predict the outcome of the firm (in terms of bankruptcy or restructuring) with their respective level of accuracy. The study, firstly, reveals that machine learning models have better predictive accuracy than statistical models. Secondly, the model's predictive accuracy is highest near the year/event of distress; as we move further from the year of liquidation/restructuring, the accuracy declines. Thirdly, financial variables and firm-specific characteristics have better precision and accuracy than only assessing the firms based on economic parameters.

Keywords Financial distress · Liquidation · Machine learning · Prediction · Restructuring

✉ Pallavi Sethi
pallavi.kalra@dtu.ac.in

Archana Singh
archanasingh@dtu.ac.in

Vikas Gupta
vikasguptadtu@gmail.com

[1] Delhi Technological University, Delhi 110089, India

Published online: 24 April 2025

1 Introduction

Bankruptcy exerts a detrimental impact on any corporation. Various methodologies can predict bankruptcy, which is crucial for diverse users, including investors, auditors, creditors, and regulators. Future bankruptcy could result in substantial losses, necessitating proactive measures to mitigate and preempt such losses. Misclassification of potential and future bankruptcy could lead to allocating scarce financial resources to inefficient firms, further leading to the erosion of public finance. Consequently, there is a need to enhance the accuracy of bankruptcy prediction. Techniques need to be developed to ensure stakeholders derive certain benefits.

Business failure is not an abrupt outcome or event; it is a process where firms undergo various stages of financial distress before liquidation or being declared bankrupt. This underscores the importance of early detection signs, enabling management to take corrective actions and avert failure. Intuitive credit risk management benefits all stakeholders: lenders, management, shareholders, investors, and regulators. Basel III norms re-emphasise the importance of credit risk management by developing credit risk models, which would, in turn, lead to decreased delinquencies in bank loans and Non-performing assets (NPA). Early warning models are widely used to evaluate a company's financial health; however, this analysis should be contextualized within value creation, which is the primary objective of firms operating in capital markets. Conservative corporate strategies are often viewed as safer, lowering both risk and returns, whereas aggressive strategies are associated with higher levels of risk and returns, as indicated by models like the Altman Z-Score (Bolek & Gniadkowska-Szymańska, 2023).

While there is existing literature on predicting financial distress, there remains a need to reevaluate the models used for such predictions. Credit risk models, which have been used interchangeably for insolvency and bankruptcy prediction and financial distress prediction (Gepp et al., 2010), initially employed Logistic regression (LR) for distress prediction. However, with Artificial Intelligence techniques such as Artificial neural networks and Support vector machines, predictive modelling has become a crucial tool for decision-making (Abu-Mostafa, 2001).

This research contributes to developing financial distress prediction models by utilising parsimonious information, including financial ratios derived from the economic data provided by firms. The quantitative variables considered include leverage, liquidity, and corporate profitability. Various prediction models have been explored in different studies; this research incorporates techniques like Decision trees, Random forests, and Artificial neural networks (ANN) and conducts a comparative analysis with the traditional method of Logistic Regression. The study has been performed on 1111 firms that were under financial distress and were liquidated or underwent restructuring through NCLT from 2017 to 2023.

The trade-off theory posits that firms must balance the costs and benefits associated with debt usage. Key advantages of debt financing include the interest tax shield, reduced financing costs, and preservation of ownership structure, while the associated costs encompass bankruptcy expenses and the likelihood of

financial distress. This necessitates robust credit risk management by lenders and financial institutions (Suresh et al., 2023). Credit risk management is of utmost concern in India, also due to the alarming rate at which the NPAs have risen over the last several years. This led the Central bank, i.e. The Reserve Bank of India, to develop various schemes to reduce the exposure of commercial banks to ever-escalating issues of NPAs. Corporate debt defaults were governed by different laws and regulations before 2016, like the Sick Industrial Companies Act (SICA), 1985, Debt Recovery Act, 1993, Corporate Debt Restructuring (CDR), Strategic Debt Restructuring (SDR), Securitization and Reconstruction of Financial Assets and Enforcement of Security Interest Act (SARFAESI), 2002 and Company Law, 2013 (amendments to Companies Act). These laws and regulations failed to solve the banks' twin balance-sheet problem, which led to the conversion of standard assets (loans) of the banks to sub-standard assets and finally into NPAs. The continuing increase in the number of NPAs has advocated the implementation of a uniform code to recover and resolve stressed assets. Insolvency and Bankruptcy Code (IBC) was implemented in India in 2016. As per IBC 2016, companies undergoing financial distress can approach the National Company Law Tribunal (NCLT), the legal body that would appoint insolvency professionals (IP) to manage the process. The result would be either restructuring the stressed assets or liquidating the company under trial. (Gupta, 2022).

This study evaluates the predictive ability of Logistic Regression and two machine learning models (Random Forest and Artificial neural network). The purpose is to find a suitable model for predicting whether a company can be restructured or needs to be liquidated so that the misclassification can be eliminated. The main reason behind this is that stakeholders need to decide while classifying firms that are financially distressed and are potential candidates that could face bankruptcy in future. This would ensure an increase in compatibility towards the decision-making process.

The analysis indicates that predictive accuracy is enhanced when firm-specific characteristics, such as age, industry, and size, are incorporated alongside financial variables. Among the models evaluated, Random Forest demonstrated the highest efficacy in predicting financial distress, followed by neural networks and logistic regression, suggesting that machine learning models generally surpass traditional statistical models for this application. Additionally, company size was identified as a significant determinant of distress outcomes, with smaller firms being more susceptible to liquidation, a finding consistent with prior studies by Almaskati et al., (2021) and Sehgal et al., (2021). Furthermore, the predictive accuracy of all models declined as the forecasting horizon increased, a trend similarly reported by Charalambakis and Garrett (2019) and Sehgal et al., (2021). Key financial variables contributing to distress prediction include Profit After Tax as a percentage of total income, Return on Assets, and Interest Coverage.

The rest of the paper proceeds as follows: Sect. 2 reviews the literature, Sect. 3 describes the study's objective, Sect. 4 highlights the data and methodology, and Sect. 5 provides the results and discussion. Finally, Sect. 6 concludes and provides the scope for future research.

2 Review of Literature

2.1 Studies Employing Statistical Model for Predicting Distress

Since the 1960s, many types of research have focused on companies that have become bankrupt. "Insolvency" refers to a company's financial situation when it finds it difficult to generate successful performance to satisfy the liabilities. The companies conduct financial analyses to ensure long-term viability. The study warns about potential risk and solvency issues by analysing a company's financial performance through its liquidity, financial stability and creditability. Many studies have been conducted in different fields to adapt the insolvency prediction model for businesses in other sectors. Research has been conducted to predict the solvency of various firms. The previous researchers incorporate both financial and non-financial aspects in computation. Studies mainly focus on financial ratios derived from financial reports instead of non-financial elements like a firm's age or management expertise.

In many published studies, financial ratios have been used to forecast solvency. Beaver (1966) analysed thirty ratios split into six groups. The ratio was chosen from different groups, which specified various solvency problems like current ratio, total debt/ cash flow, etc. Altman (1968) used the Z-score model to predict financial bankruptcy. The study used different variables estimated using a balance sheet and income statement. Z-score models indicated the probability of avoiding failure. The variables considered for evaluating the firms' financial conditions were coverage, activity, leverage, profitability and liquidity (Altman et al., 2010). Zmijewski (1984) developed a weighted probit model for predicting bankruptcy. Hopwood et al. (1994) noted that financial distress is related to a probability of going bankrupt. Rather than using a multiplicity of various financial ratios used in previous research, this study is based on specific models, including logistic regression, ANN, etc. Some earlier studies (DeFond and Jiambalvo, 1994) have considered a model incorporating various ratios that helped measure liquidity, solvency, and profitability. The author selected the model based on a quantitative measure named the Zmijewski score.

Various research studies have indicated the use of qualitative measures of financial distress. As per evidence provided in research studies by Gajpal et al., (1994), Gilson et al. (1990), Giroux and Wiggins (1984) and Turetsky (1997), there is use of negative operating cash flow, omitting or reducing dividend payment, debt default includes default in loan payments or technical default. Troubled debt restructuring has also been used to indicate financial distress before a firm goes bankrupt. The study has used qualitative variables as dummy variables.

John and John (1993) suggested that financial distress could progressively result in negative operating cash flows. Giroux and Wiggins (1984) consider negative cash flows signalling a decline in corporate financial status. In the study, Foster et al., (1998) mentions that any firm's cash flow provides vital information to the investors and helps evaluate the chances of economic distress. As per the author's opinion, negative operating cash flow implies a tendency towards

insufficient working capital and future cash position. This could lead to the survival of the firm in the long term.

Dielman and Oppenheimer (1984) noted that a firm's dividend decisions could be vital in identifying financially distressed firms. The authors mentioned in the study that firms generally follow a stable dividend policy when the firm is financially healthy. DeAngelo et al., (1994) contended that any firm facing financial problems could think of omitting or cutting dividends. Hence, dividend changes, mainly reductions in the amount of dividends, might predict financial distress. Dividend omissions or reductions were measured as dummy variables.

Chen & Church (1996) have mentioned that when a firm is about to fail, its financial ratios deteriorate. However, a clear indication of potential problems could lead to bankruptcy, and it becomes challenging to encounter debt obligations. It is not easy to comply with lending agreements and make scheduled payments. Additionally, when certain firms violate the debt covenants, the creditors might accelerate debt maturity and impose higher re-contracting costs on the lenders (Anandarajan et al., 2001). Further, any technical defaults would constrain the freedom of action while operating in the firm (Smith & Warner, 1979). When the firm defaults on the debt agreements and cannot provide a remedy through re-negotiation or doesn't get a waiver from the creditors, there is a higher chance of becoming bankrupt when the company defaults (Varadraj & Abhay, 2014).

Previous studies have regarded Troubled debt restructurings (TDR) as a distinct stage in the financial distress continuum (as mentioned by Giroux & Wiggins, 1983 and Turetsky, 1997). TDR results from various private debt renegotiations to resolve or remedy debt defaults (Anandarajan et al., 2001). As a result of TDR, current and future cash flows related to debt service might improve. TDR firms that are unsuccessful might end their bankruptcy. TDR could be considered a clear signal about financial distress, which could indicate bankruptcy.

As per Ooghe and Prijcker, (2008), lower financial independence, weaker profitability, increased expenses and lower cash flows were considered the critical solvency predictors. The study observed increased expenditure at constant turnover, which might help predict financial problems and inadequate professional competence for implementing decisions arising from various non–financial predictors. Bhimani et al. (2010) applied eleven financial ratios: financial coverage, capital/ total asset, investment ratio, days in payables and receivables, interest cost, financial coverage, solidity, gross income and return on equity. They have highlighted days and payables in receivables that determine the payment behaviour of creditors and debtors. In a study of Taiwan electronics market by Su and Huang (2010), firm failure prediction for Taiwan's electronics companies is conducted using the Logit regression model, the KMV option-based asset valuation model, and the ZPP GARCH autoregressive time series simulation model. It was concluded that the predicted failure probabilities are much larger using the Logit model than the other models.

Al-Kassar and Soileau (2014) selected specific ratios based on assets and liabilities, and there was no credit interval for failure in prediction. The study indicated that the higher the profitability and other ratios, the lower the solvency risk. When this ratio is lower, the risk of solvency is high. Mironiuc et al., (2015) indicated that the current ratio, average collection, financial leverage, liquidity, financial expenses,

and other ratios indicated that factors were significant statistically in predicting the firm's solvency position. Liang et al., (2016) have used ratios like cash flow, growth, turnover, profitability, solvency, and capital structure to assess bankruptcy. The study observed profitable and solvency ratios as the most significant ratios to predict solvency.

2.2 Studies Employing Machine Learning to Predict Distress

In the current times, machine learning (ML) is considered to be a cross between Computer Science and Statistics. Artificial INTELLIGENCE (AI) has been constantly improved, and big data is easily accessible in different sectors. The computation cost has been reduced, and people can use AI in the best way possible. AI's evolution directly impacts the technology market, mainly the firms related to AI (Lu et al., 2018).

Network structures with more robust mapping capacity could manage the interaction between non-linear associations and variables. Artificial neural networks (ANNs) are a solution for different Machine learning (ML) models. They are used in the financial sector to predict exchange rates, financial distress, credit scoring, and portfolio management.

Among the early studies, Henley and Hand (1996) used the k-nearest neighbour (k-NN) algorithm to assess the customer's ability to pay loans. Galindo and Tamayo (2000) performed a comparative examination of machine learning and statistics to determine mortgage loan defaults. They concluded that the best estimates were provided by CART Decision trees, followed by neural networks and k-NN.

Anandarajan et al., (2001) used an artificial neural network to predict the bankruptcy of financially distressed firms. They used the Zmijewski distress score, which incorporates ratios measuring profitability, liquidity, and solvency.

Yeh and Lien (2009) examined the forecasting accuracy of a few data mining techniques and observed that artificial neural networks (ANN) estimated the default more accurately than the other models.

Decision trees were used by Gepp et al., (2010) to predict business failure, and they observed that decision trees were far superior to traditional discriminant analysis. Similarly, Bhandari and Iyer (2013) used ratios on 50 failed and 50 non-failed US firms. They concluded that discriminant analysis was a better predictor than the cross-validation approach adopted under the data mining technique. Li et al. (2014) evaluated distress and the effectiveness of financial ratios using data envelopment analysis and logistic regression on a dataset of Chinese enterprises. Their results confirmed the theory that using efficiency information improves prediction ability.

Morum and Roy (2013) employed seven financial ratios to predict corporate sickness in the Indian Steel industry. They developed the model using stepwise logistic regression and cluster analysis on 40 companies classified as sick and non-sick. The authors believed that stepwise logistic regression has a better predictive ability than the other methods.

In the study, Bapat & Nagale (2014) compared the predictive powers of three distinct distress models for Indian listed companies: discriminant, logistic, and neural

network. They opined that neural networks were the most accurate regarding precise classification.

Behr and Weinblat (2017) conducted another study using machine learning tools, logit regression, decision trees, and random forests. The study included German, French, Italian, British, Spanish, and Portuguese firms. They observed that distressed firms had a spiralling debt ratio, liquidity ratio, return on sales and return on assets, which were significant bankruptcy predictors.

Khemakhem and Boujelbene (2018) integrated financial and non-financial data to create a credit score model. Their research revealed that profitability ratios, repayment capability, and solvency are essential default indicators.

Addo et al., (2018) worked on binary classifiers using machine learning models to forecast loan default probabilities. They concluded that decision trees are more accurate than neural networks for predicting distress.

Shrivastava et al. (2018) predicted a corporate crisis for enterprises in India that did not have bankruptcy specifics. Based on their data, they argued that the Bayesian technique performed better in accuracy and predictive ability than the standard logistic model. In another study in the same year by Nair and Sachdeva (2016) on 574 Indian manufacturing companies classified as distressed and non-distressed, it was established that profitability and efficiency ratios were of utmost importance while predicting distress. Tang et al., (2019) used the random forest model to ascertain the credit card default risk for the firms involved in China's energy business. Their results indicated that the overdraft ratio and the amount of credit card costs significantly affect credit risk. De Paula et al., (2020) analysed data from a Brazilian credit union's loan transactions for credit risk. They established the superiority of the random forest technique in estimating credit scoring compared to logit regression. Shrivastava et al., (2020) employed logistic regression, random forest, and Ada Boost to predict bank failure. Ada Boost provided the highest level of accuracy when compared to the other approaches.

Golbayani et al. (2020) used four machine learning techniques deemed helpful in previous studies (bagged decision trees, random forest, support vector machine, and multilayer perceptron) on a dataset, and the results show that decision tree-based models perform better on the chosen datasets.

Recently, a study of 12,000 small and Medium enterprises concluded that the Random Forest algorithm had the highest prediction accuracy compared to Logistic regression and Artificial Neural networks (Malakauskas & Lakštutienė, 2021).

Gupta (2022) evaluated four machine learning models on the companies under IBC and revealed that random forest outperforms the other three techniques: Support Vector Machine, Decision tree and Logit model.

A study was performed to predict financial distress amongst Fintech companies by employing an artificial neural network model. The study pinpointed Return on capital, current ratio, quick ratio and debt-to-equity ratio as significant predictors of financial distress. (Halteh et al., 2024).

Alvi and Arif, (2024) undertook a thorough longitudinal analysis of non-financial firms in Pakistan and an extensive assessment of machine learning algorithms for predicting default risk. The lack of early warning indicators within the non-financial

sector contributes to insolvency and financial distress, which in turn adversely affects the economy and drives higher unemployment rates.

Most studies employing data mining techniques opine that machine learning models outperform statistical models in terms of discriminatory power and precision. It was also discovered that machine learning algorithms for credit distribution led to allocation to safer and larger borrowers, reducing lenders' losses.

The current study intends to bridge the gap in the existing literature by comparing the predictive accuracy of statistical models and machine learning models on companies under IBC to predict bankruptcy or reorganisation. It is imperative to develop specific models based on assessing the likelihood of defaulting on a counterparty. The financial institutions deal with counterparty default now and then, a tragic issue they face. The lenders, financial market players, and government should reduce counterparty defaults to protect the creditor's interest and restore a firm's sustainability. Financial Distress Prediction intends to solve problems by classification through financial ratios, which could describe a firm's current situation for a particular period. Also, there is a need to identify the economic and firm-specific factors that can serve as an early indicator of default.

3 Objectives of the Study

The current study caters to the following objectives-

- To predict whether the firm will become bankrupt or be restructured by NCLT based on the analysis of financial information in the form of financial ratios using Machine Learning Techniques and Statistical Techniques.
- To understand the key ratios that define the impending default
- To identify the impact of non-financial variables like age, size of the firm and industry group on the predicting default.
- To assess the predictive ability of model(s) years before the restructuring/liquidation.

4 Data and Research Methodology

This study is analytical in nature. It focused on analysing five years of data from 1111 companies.

Most published studies have used healthy firms as the control group. In this study, 769 companies were considered liquidated by NCLT since its inception, i.e., 2016, and 342 companies had their resolution plan approved from 2017 to 18 till December 2023. Table 1 shows the details of the companies included in the study's sample. The selected companies' financial information has been considered for five years before being liquidated or restructured. We used financially stressed firms admitted to NCLT as the sample of study to ensure increased compatibility to a realistic decision-making process as adopted by Anandarajan et al., 2001. This is because

Table 1 Year-wise data of restructured and liquidated companies (source-IBBI)

Year	No. of cases admitted in IBBI for restructuring	No. of restructured cases available in CMIE prowess	No. of cases admitted in IBBI for liquidation	No. of liquidated cases available in CMIE prowess	Total Number of distressed firms for which the data is available	Financial and Non-Financial data collected for the years
2016–17	Nil	**Nil**	23	**16**	16	2011–12 to 2015–16
2017–18	19	**7**	376	**158**	165	2012–13 to 2016–17
2018–19	75	**31**	550	**195**	226	2013–14 to 2017–18
2019–20	132	**46**	751	**214**	260	2014–15 to 2018–19
2020–21	119	**42**	161	**43**	85	2015–16 to 2019–20
2021–22	144	**46**	231	**58**	104	2016–17 to 2020–21
2022–23	189	**72**	155	**20**	92	2017–18 to 2021–22
2023–24 (Till December 2023)	208	**98**	354	**65**	163	2018–19 to 2022–23

The bold text signifies the best out of all the values signifying the superiority of a model over the others based on that specific parameter

the Resolution Professionals have to decide to classify a financially distressed firm (rather than a financially healthy firm) as a potential candidate for bankruptcy.

The observations regarding 1111 companies for the past five years gave us 5555 firm observations, which were used to identify the key financial ratios that predict the outcome of economic distress.

Table 2 gives a snapshot of sample characteristics from various industries.

The industry classification is based on the National Industrial Classification (NIC) 2008, adopted by the Government of India to disseminate industry-wise economic data.

Table 3 Based on an extensive literature review, the financial ratios and the independent variables employed in the study.

The literature review shows that machine learning has been used to predict bankruptcy and credit risk management since the 1990s. Statistical tools, which relied more on assumptions of normality and linearity of data, gave way to machine learning techniques, which were data-driven, robust, and distribution-free (Gupta, 2022).

Statistical Techniques were predominant from the 1960s until the 1990s, with Multivariate discriminant analysis (MDA), proposed by Altman (1968), and Logit analysis (LA) being the prominent ones (Bellovary et al., 2007). Altman used six financial ratios from the companies' financial statements to measure risk, and since then, numerous studies have used the Altman MDA model. (Martin, 1977) Developed an early warning model predicting bank failure using Logit Analysis and compared it with Linear Discriminant Analysis.

In addition to these, Ohlson (1980), Jones and Hensher (2004), and Canbas et al., (2005) used more improved and innovative modelling techniques. Statistical

Table 2 Various industries to which the sample companies belong

S. no	Industry	No. of companies restructured	No. of companies liquidated	Total
1	Mining and Quarrying	01	Nil	01
2	Manufacturing	157	361	518
3	Electricity, Gas, Steam and Air conditioning Supply	15	24	39
4	Construction	37	62	99
5	Wholesale and Retail Trade: Repair of Motor Vehicles and Motorcycles	44	164	208
6	Transportation and Storage	10	26	36
7	Accommodation and Food Services Act	8	5	13
8	Information and Communication	16	18	34
9	Financial and Insurance Activities	36	67	103
10	Professional, Scientific and Technical Activity	1	12	13
11	Administrative and Support Service Activities	13	18	31
12	Education	1	1	2
13	Human Health and Social Work Activities	3	6	9
14	Arts, Entertainment and Recreation	0	3	3
15	Other Service Activities	0	2	2

Table 3 Financial variables and firm-specific characteristics included in the study

Financial variables	
Measure	Ratio/Metric
Efficiency	Total asset utilisation ratio Finished goods turnover ratio Debtors Turnover ratio Creditors Turnover Ratio Gross Fixed Asset Utilisation Ratio Return on Total Assets Return on Capital Employed Return on Net Worth
Profitability	Profit before tax as % of total Income Profit After tax as % of total Income Net profit margin Operating Profit margin Cash profit as % of Total Income
Coverage	Interest Cover Debt Service Coverage ratio
Leverage	Debt equity ratio Total Operating Liabilities to Tangible Net Worth (TOL/TNW)
Liquidity	Current ratio Quick ratio
Firm-specific characteristics	
Size	Size Decile
Age	Year of incorporation
Industry	Industry Group of the firm

techniques required fulfilling the assumptions about the data concerning normality and linearity. The researchers were required to define the structure of the model a priori and then estimate its parameters to fit in the data (Gupta, 2022).

Machine learning techniques appeared in the late 1980s as an alternative and an improvement over the statistical models (Bellovary et al., 2007). Machine Learning techniques are data-driven and tend to be more computationally based. Numerous papers have been published using various machine-learning techniques for distress prediction. They are non-parametric, and despite requiring more computational power, they have become the primary bankruptcy prediction method, outperforming statistical models (Clement, 2020). Machine learning models are good at classifying almost any variable selection method and are more robust to multicollinearity (Liang et al., 2015).

Since both statistical and Machine learning techniques have their own merits and limitations, the study employs one statistical technique and two machine learning techniques to identify the accuracy and performance of each method.

The study applies random forest, artificial neural network, and logistic regression models, which are used to predict the firm's outcome (in terms of bankruptcy or reorganisation)[1] With a certain level of accuracy.

The following techniques were used for data analysis:

- Random Forest (Machine Learning model)—This model is a supervised algorithm that creates a random forest. Many decision trees are combined to make a random forest. It is used in regression and classification problems. This structure is easier to understand, and its interpretation is straightforward. It has certain advantages: the tree structure is visualised, and there is no need for extensive data preparation for analysis (Aker & Karavardar, 2023). Also, the diversity is created by incorporating the benefit of double randomisation, where each tree is estimated based on a synthetic sample drawn randomly from an estimation sample (Klaus-Peter Hellwig, 2021)
- Artificial Neural network (Machine Learning model)—This model adopts the human brain system and can generalise knowledge for predicting future events. This model is used widely for optimisation support and prediction with a back-propagation algorithm (Alamsyah et al., 2021).
- Logistic Regression (Statistical technique)–This model emerged as the most suitable for handling situations wherein the dependent variable is classified or categorical. It helps in determining cause-effect relationships with certain explanatory variables. The prediction in this model is made using equations. This model is focused on finding an equation that would minimise the difference between the dependent variable's actual value and the independent variable's predicted value (Özdamar, 2002). Cross-validation and stratified sampling techniques were adapted to conduct the modelling as specified by Japkowicz and Shah (2011).

4.1 Data Pre-processing

Machine learning models follow the Garbage in, Garbage out (GIGO) principle (Arora & Saurabh, 2022). Therefore, the data must be preprocessed before the modelling stage for accurate results. Preprocessing required dealing with an imbalanced dataset, as the liquidated firms were dominant in number, and spreading the sample to enable random selection of instances. SMOTE, spread subsample, and randomise filter were used to deal with this.

Ten sets were randomly created using 10-fold cross-validation, with 100 iterations performed to get the best results. Cross-validation is a technique for evaluating model performance on unseen data.

[1] Bankruptcy and Liquidation have been used interchangeably.
Restructuring and reorganisation have been used interchangeably as reorganisation is financial restructuring.

4.2 Optimisation Parameters for Model

The study is based on machine learning models that must be tuned for optimum results. Optimizing the hyperparameters to achieve the best possible model architecture is often necessary to implement machine learning algorithms successfully (Beniwal et al., 2023). The hyperparameter optimisation for the machine learning models is as follows-

1. Random Forest—The default batch size was 100, the number of iterations remained the same, and the seed's initial value remained 1.
2. Neural Network—The neural network's learning rate was 0.3, and the batch size was 100. The hidden layers were set to half the sum of attributes and class.

4.3 Performance Evaluation Metrics

The study applied the performance evaluation metrics to evaluate the model's predictive and classification ability: the Confusion Matrix and its indicators, ROC (Receiver Operating Characteristics), and the Gini coefficient.

4.3.1 Prediction Performance (Confusion Matrix) and its Indicators

A confusion matrix gives valuable information about the classifier's performance by comparing actual and predicted classes (Balasubramanian et al., 2019).

A sample confusion matrix is provided in Table 4.

4.3.1.1 Sample Confusion Matrix The reorganised cases have been referred to as positive, and bankrupt cases have been referred to as unfavourable.

Classification accuracy rates calculated from the confusion Matrix as follows-

1. *Accuracy ratio (AR)*: It is the proportion of cases that have been correctly classified (positive or negative).

$$\text{Accuracy Ratio} = (\text{Npp} + \text{Nnn})/\text{Total Sample}$$

   where, Npp(Nnn) is the number of positive(negative) cases classified as positive(negative)
2. *Sensitivity*: It is the proportion of negative cases correctly classified as harmful.

Table 4 Sample Confusion Matrix

Predicted			
Actual		Positive	Negative
	Positive	TP	FN
	Negative	FP	TN

$$\text{Sensitivity} = \text{Nnn}/(\text{Nnn} + \text{Nnp})$$

where Nnn (Nnp) is the number of negative (negative) cases classified as negative(positive)

3. *Specificity*: It is the proportion of positive cases correctly classified as positive and is measured as follows-

$$\text{Specificity} = \text{Npp}/(\text{Npp} + \text{Npn})$$

where, Npp(Npn) is the number of positive(positive) cases classified as positive(negative)

4. *Precision*: is the proportion of correctly classified negative cases out of all those classified as harmful. It is measured as.

$$\text{Precision} = \text{Nnn}/(\text{Nnn} + \text{Npn})$$

where, Nnn(Npn) id number of negative(positive) cases classified as negative(negative)

5. *Type I and Type II errors*: Type I error is the proportion of negative cases classified as positive. In contrast, type II error is the proportion of Positive cases classified as unfavourable.

$$\text{Type I error} = 1 - \text{Sensitivity}$$
$$\text{Type II error} = 1 - \text{Specificity}$$

6. *Gini Coefficient*: The Gini coefficient is an essential tool for assessing the consistency in the results of a machine learning model. It is also the widely used index for validating credit scoring models. (Gupta, 2022).
   It is calculated as G=2(AUROC)–1

Classification rates calculated from the confusion Matrix provide information about assessing a classification model's accuracy and appropriateness. Out of all the parameters discussed above, the most important one from a decision-maker's perspective is the type I error. The reason is that the cost of classifying a negative case as positive far exceeds that of a positive case as unfavourable. For example, the losses to a bank from lending to a bankrupt firm (Type I error) are expected to be much higher than the losses from refusing to lend to a typical going concern, which are the foregone interest payments (Type II error) (Almaskati et al., 2021). It is worth mentioning, however, that the costs to society and the economy from Type II errors might be significant when suitable opportunities are systematically declined due to this error (Almaskati et al., 2021). Thus, error cost should be an essential criterion for model evaluation (Adnan Aziz & Dar, 2006).

4.3.2 Classification Power (Receiver Operating Characteristic (ROC) Curve) and Gini Coefficient

Apart from the confusion matrix and its accuracy rates for prediction ability, the ROC curve or Area under the curve (AUC) and the Gini coefficient are used as classification ability indicators.

The analysis uses the sensitivity (the percentage of default positive outcomes that have been correctly classified) v/s 1- specificity (the percentage of default adverse outcomes that have been correctly classified) to make the curve. The false positive rate is plotted on the horizontal axis, with the actual positive rate on the vertical axis. With this, we get the area under the ROC curve (AUROC) based on a particular model. The area ranges from 0.5 (for a worthless model) to 1 (for a perfect classifier). The larger the area under the curve, the better the model's classification ability.

The Gini Coefficient highlights the difference between a base classifier (taken as the default 0.5 for each class) and the classifier model in question.

5 Results and Discussion

The results indicate the output of a Machine Learning or statistical technique in the form of Accuracy Rate, Precision, Type I error and AUC. The output was generated by first including only financial variables and later incorporating firm-specific characteristics with the economic variables.

The analysis was spread over five years before the date of Liquidation/restructuring. Also, the study was carried out each year by including financial variables only and later on by including financial variables and firm-specific characteristics. It was observed that the accuracy rate was higher when firm-specific characteristics were included along with the economic variables to predict the financial distress outcome of the firm. Also, the accuracy declines as we move further away from the date of liquidation/restructuring. However, accuracy cannot be considered as the criteria for measuring the prediction accuracy of a model. According to (Hsieh, 1993). The cost linked with Type I error is higher for the stakeholders than the cost related to Type II error, so the best model is the one with the lowest Type I error.

The results of the various models one year before bankruptcy/reorganisation are given in Table 3. As we compare Panels A and B, we observe that the accuracy rate, precision, area under the curve (AUC), and Gini coefficient provide a better classification and prediction measure when the firm-specific characteristics are included with the financial variables. Out of the three models employed in the study, the Random Forest is the best bet in the distress prediction models as it has the lowest Type-I error. As per Anandarajan et al., (2001), "The objective of learning classifications from sample data is to classify and predict successfully on new data. A classifier's error rate is the most commonly used measure of success or failure. The error rate is statistically defined as the error rate of the classifier on an asymptotically large number of new cases that converge in the limit to the actual population distribution." For the analysis results, one year before the actual event of bankruptcy/restructuring, as shown in Table 5. It can be seen that type-I error is

Table 5 One year before Bankruptcy/reorganisation

Model	Accuracy rate	Precision	Type I error (1-sensitivity)	AUC	Gini coefficient
Panel A: Only financial variables					
Random Forest	**0.613**	0.586	**0.087**	**0.631**	**0.262**
Neural network	0.557	0.559	0.233	0.565	0.13
Logistic	0.543	0.543	0.146	0.518	0.036
Panel B: Financial variables and firm-specific characteristics					
Random Forest	**0.735**	0.711	**0.160**	**0.784**	**0.568**
Neural network	0.727	0.736	0.244	0.740	0.48
Logistic	0.718	0.708	0.204	0.764	0.528

The bold text signifies the best out of all the values signifying the superiority of a model over the others based on that specific parameter

lower in the analysis done when considering only financial variables. One reason for that could be that as the firm approaches bankruptcy or restructuring, the financial position of the firm, as exhibited in the financial statements, has already deteriorated to a great extent. An analysis of attributes using the attribute selector indicates that size, Return on Total Assets, and Interest cover are the significant attributes impacting the firm's failure. Thus, as the firm nears failure, the ability to pay the financial creditors becomes a major deciding factor in terms of the fate of the firm.

As we move further away from the date of liquidation and restructuring, we see (Table 6) that the accuracy rate of the model declines. Also, the probability of type-I error increases. This highlights that distress prediction is more accurate one year before liquidation/restructuring. Also, the attributes impacting the final decision concerning the firm's fate are size, and Gross fixed asset utilisation ratio. The size of the firm plays a significant role, and this follows the results as defined by Malakauskas and Lakštutienė, (2021) and Arora and Saurabh, (2022).

Table 6 Two years before Bankruptcy/Reorganisation

Model	Accuracy rate	Precision	Type I error (1-sensitivity)	AUC	Gini coefficient
Panel A: Only financial variables					
Random Forest	**0.577**	0.568	**0.164**	**0.617**	**0.234**
Neural network	0.557	0.581	0.414	0.604	0.208
Logistic	0.522	0.531	0.166	0.546	0.092
Panel B: Financial variables and firm-specific characteristics					
Random Forest	**0.736**	0.715	**0.166**	**0.789**	**0.578**
Neural network	0.723	0.729	0.241	0.745	0.490
Logistic	0.723	0.711	0.196	0.754	0.508

The bold text signifies the best out of all the values signifying the superiority of a model over the others based on that specific parameter

(1994) opine that a firm is in financial trouble if its Earnings before interest, tax, depreciation, and amortisation (EBITDA) are less than its expenses. Therefore, the current study examines three years before the firm is adjudicated to be liquidated or restructured and calculates the accuracy rate and other prediction and classification indicators.

As we look at the results of Table 7, we see better prediction accuracy and a lower type-I error. This signifies that the prediction based on financial variables and firm-specific characteristics was more accurate three years before the firm was declared bankrupt or eligible for restructuring.

Thus, it can be analysed that to predict a firm's financial distress, it is imperative to look into its performance as soon as it starts showing warning signals regarding low income and inability to pay operational creditors. Three years before the judgement of the distressed firms as bankrupt or restructured, the significant attributes/variables that are important are size and Profit after tax as a percentage of total income.

Early detection of a firm's movement towards failure in the early stages benefits business stakeholders. (Ashraf et al., 2019) In all the years of analysis, random forest emerged as a clear winner in predicting early distress and had the lowest Type I error, indicating its superiority over the traditional Logit model. This can be due to the fundamental difference between traditional statistical and Machine Learning models in handling data and making predictions. Statistical models, such as linear and logistic regression, are based on predefined assumptions about data distribution and variable relationships. These models are typically parametric, assuming a specific functional form (e.g., linearity) to define the relationship between input variables and the outcome. Statistical models are often effective with smaller datasets and perform well when the relationships among variables are well-understood. However, their predictive accuracy can be limited if data distribution assumptions are unmet. Here, the dataset comprises 1111 firms, which is a large dataset in terms of statistical models. On the other hand, Machine-learning models focus on flexible learning from complex, high-dimensional data, resulting in enhanced predictive capabilities. ANN

Table 7 Three years before Bankruptcy/Reorganisation

Model	Accuracy rate	Precision	Type I error (1-sensitivity)	AUC	Gini coefficient
Panel A: Only financial variables					
Random Forest	**0.607**	0.590	**0.153**	**0.665**	0.33
Neural network	0.591	**0.627**	0.437	0.653	**0.306**
Logistic	0.546	0.554	0.270	0.567	0.134
Panel B: Financial variables and firm-specific characteristics					
Random Forest	**0.729**	**0.709**	**0.174**	**0.782**	**0.564**
Neural network	0.701	0.707	0.259	0.739	0.478
Logistic	0.707	0.698	0.215	0.745	0.490

The bold text signifies the best out of all the values signifying the superiority of a model over the others based on that specific parameter

follows Random Forest, indicating the superiority of machine learning models over conventional prediction models. The study's results are similar to those reported by Gupta, (2022), Balasubramanian et al., (2019), and Sehgal et al., (2021).

6 Summary of Findings

The analysis indicates that predicting a firm's financial distress necessitates early detection through evaluating performance indicators, particularly when warning signs such as declining income and an inability to meet obligations to operational creditors become apparent. This would enable the firm to move towards sustainable measures in terms of innovation and creativity so as to deal with the ever-growing and dynamic business environment (Jain & Gupta, 2018; Sinha et al., 2009). The predictive accuracy of distress models improves considerably when firm-specific characteristics, including age, industry, and size, are incorporated alongside financial variables. Among the models assessed, Random Forest demonstrated the highest predictive efficacy, followed by neural networks and logistic regression, underscoring the superior performance of machine learning algorithms compared to traditional statistical models in forecasting financial distress. Firm size emerged as a particularly influential determinant, with smaller firms exhibiting a greater vulnerability to liquidation, consistent with prior findings by Almaskati et al., (2021) and Sehgal et al., (2021). Furthermore, the analysis revealed a decline in predictive accuracy as the forecasting horizon extended, aligning with observations by Charalambakis and Garrett (2019) and Sehgal et al., (2021). Key financial variables influencing distress prediction were profit after tax as a percentage of total income, return on assets and interest coverage. Notably, the cost associated with Type I errors exceeded that of Type II errors, emphasising the importance of models with minimised Type I errors. Finally, while statistical models offer interpretability through predefined relationships, machine learning models leverage flexibility in learning from complex, high-dimensional data, enhancing predictive capabilities.

7 Conclusion

Financial distress leads to significant economic losses for all stakeholders, making it imperative to identify early distress signals. It encompasses several adverse consequences, including reduced market share and reputational standing, increased capital costs, a withdrawal of favourable trade credit terms from suppliers, and losses associated with asset sales conducted under distressed conditions (Li, 2024). Additionally, financial distress often leads to the departure of valuable human capital and heightened exposure to aggressive strategies from competitors. This paper aimed to determine the predictive and classification ability of the traditional logit model with the machine learning models. It also aimed at determining the relative importance of the various attributes in predicting distress. Classic financial ratios proposed by Turetsky and McEwen, (2007), Altman et al., (2017) and Almaskati et al., (2021) were used, along

with the firm's specific characteristics, such as its age, size, and industry. A comparative analysis was done to identify the impact of including non-financial variables in the model.

The traditional logit model was compared to two machine learning models: Random forest and Artificial Neural network. The predictive ability of the three models was compared using the AUC under the ROC curve and the type-I error. The Random Forest model emerged as the clear winner, followed by ANN and the logit model, thereby highlighting the supremacy of machine learning models over traditional models. The analysis also identified the highest predictive ability years before the firm failed. It was found that three years before the liquidation/restructuring of the firm, the Random Forest model had the highest predictive accuracy and lowest error rate. It is also observed that all the models had a predictive accuracy greater than 75%, thus confirming their robustness.

The study involved the firms that were referred to NCLT, and to the best of our knowledge, it is the first paper using the database of only distressed firms as compared to an equivalent number of solvent firms in line with the study undertaken by Balasubramanian et al., (2019).

To predict financial distress, it is imperative to look into the firm's size and significant financial variables like Profit after tax as a % of total income, Return on assets and Interest coverage. This can aid the managers in assessing the economic health of their firms and serve as a significant indicator for the investors who are willing to invest in such firms. It can significantly contribute to the banking sector, helping them identify the default firms and save their balance sheets from the Non-Performing Asset (NPA) problem.

Further, the analysis can be extended by incorporating macroeconomic factors such as inflation, GDP growth, etc. Machine learning algorithms like Naïve Bayes can be used to perform sentiment analysis of other components of Annual reports like Management discussion and analysis (MD&A). A Naïve Bayes classifier assigns a probability to the word or text, whether positive or negative (Arora & Saurabh, 2022). A detailed industry-based study can also be undertaken to predict specific patterns of financial distress in the Indian economy. This study included all companies restructured or liquidated by the NCLT, regardless of their industry. An industry-specific analysis, however, could provide more comprehensive insights and yield more robust results. The study is based on the financial ratios derived from CMIE Prowess. Also, the missing ratios, which are irrelevant and can't be computed for a particular industry, were imputed by the machine-learning software WEKA based on the median values. Such adversities can be handled when the study concerns a particular industry.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

Abu-Mostafa, Y. S., Atiya, A. F., Magdon-Ismail, M., & White, H. (2001). Introduction to the special issue on neural networks in financial engineering. *IEEE Transactions on Neural Networks, 12*(4), 653–656. https://doi.org/10.1109/TNN.2001.935079

Addo, P. M., Guegan, D., & Hassani, B. (2018). Credit risk analysis using machine and deep learning models. *Risks, 6*(2), 38. https://doi.org/10.3390/risks6020038

Adnan Aziz, M., & Dar, H. A. (2006). Predicting corporate bankruptcy: Where we stand? *Corporate Governance: THe International Journal of Business in Society, 6*(1), 18–33. https://doi.org/10.1108/14720700610649436

Aker, Y., & Karavardar, A. (2023). Using Machine Learning Methods in Financial Distress Prediction: Sample of Small and Medium Sized Enterprises Operating in Turkey. *Ege Akademik Bakis (Ege Academic Review)*.

Alamsyah, A., Kristanti, N., & Kristanti, F. T. (2021). Early warning model for financial distress using Artificial Neural Network. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1098, No. 5, p. 052103). IOP Publishing.

Al-Kassar, T. A., & Soileau, J. S. (2014). Financial performance evaluation and bankruptcy prediction (failure). *Arab Economic and Business Journal, 9*(2), 147–155. https://doi.org/10.1016/j.aebj.2014.05.010

Almaskati, N., Bird, R., Yeung, D., & Lu, Y. (2021). A horse race of models and estimation methods for predicting bankruptcy. *Advances in Accounting, 52*, 100513. https://doi.org/10.1016/j.adiac.2021.100513

Altman, E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance, 23*, 589–609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2017). Financial distress prediction in an international context: A review and empirical analysis of Altman's Z-score model. *Journal of International Financial Management & Accounting, 28*(2), 131–171. https://doi.org/10.1111/jifm.12053

Altman, E. I., Sabato, G., & Wilson, N. (2010). The value of non-financial information in SME risk management. *Journal of Credit Risk, 6*(2), 95–127.

Alvi, J., & Arif, I. (2024). Credit scorecards & forecasting default events–A novel story of non-financial listed companies in Pakistan. *Asia-Pacific Financial Markets*. https://doi.org/10.1007/s10690-024-09494-3

Anandarajan, M., Lee, P., & Anandarajan, A. (2001). Bankruptcy prediction of financially stressed firms: An examination of the predictive accuracy of artificial neural networks. *Intelligent Systems in Accounting, Finance and Management, 10*(2), 69–81. https://doi.org/10.1002/isaf.199

Arora, P., & Saurabh, S. (2022). Predicting distress: A post Insolvency and Bankruptcy Code 2016 analysis. *Journal of Economics and Finance, 46*(3), 604–622. https://doi.org/10.1007/s12197-022-09582-y

Ashraf, S., Félix, G. S. E., & Serrasqueiro, Z. (2019). Do traditional financial distress prediction models predict the early warning signs of financial distress? *Journal of Risk and Financial Management, 12*(2), 55. https://doi.org/10.3390/jrfm12020055

Balasubramanian, S. A., GS, R., P, S., & Natarajan, T. (2019). Modeling corporate financial distress using financial and non-financial variables: The case of Indian listed companies. *International Journal of Law and Management., 61*(3/4), 457–484. https://doi.org/10.1108/IJLMA-04-2018-0078

Bapat, V., & Nagale, A. (2014). Comparison of bankruptcy prediction models: Evidence from India. *Accounting and Finance Research, 3*(4), 91–98.

Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research, 4*, 71–111. https://doi.org/10.2307/249017

Behr, A., & Weinblat, J. (2017). Default prediction using balance-sheet data: A comparison of models. *Journal of Risk Finance, 18*(5), 523–540. https://doi.org/10.1108/JRF-01-2017-0003

Bellovary, J. L., Giacomino, D. E., & Akers, M. D. (2007). A review of bankruptcy prediction studies: 1930 to present. *Journal of Financial Education, 33*, 1–42.

Beniwal, M., Singh, A., & Kumar, N. (2023). Forecasting long-term stock prices of global indices: A forward-validating Genetic Algorithm optimization approach for Support Vector Regression. *Applied Soft Computing, 145*, 110566. https://doi.org/10.1016/j.asoc.2023.110566

Bhandari, S. B., & Iyer, R. (2013). Predicting business failure using cash flow statement based measures. *Managerial Finance, 39*(7), 667–676. https://doi.org/10.1108/03074351311323455

Bhimani, A., Gulamhussen, M. A., & Lopes, S. D. (2010). Accounting and non-accounting determinants of default: An analysis of privately-held firms. *Journal of Accounting and Public Policy, 29*, 517–532.

Bolek, M., & Gniadkowska-Szymańska, A. (2023). Is the growth of companies influencing their financial condition depending on their size: S&P 500 listed companies example. *Asia-Pacific Financial Markets, 30*(2), 323–337. https://doi.org/10.1007/s10690-022-09376-6

Canbas, S., Cabuk, A., & Kilic, S. B. (2005). Prediction of commercial bank failure via multivariate statistical analysis of financial structures: The Turkish case. *European Journal of Operational Research, 166*(2), 528–546. https://doi.org/10.1016/j.ejor.2004.03.023

Charalambakis, E. C., & Garrett, I. (2019). On corporate financial distress prediction: What can we learn from private firms in a developing economy? Evidence from Greece. *Review of Quantitative Finance and Accounting, 52*, 467–491.

Chen, K. C., & Church, B. K. (1996). Going concern opinions and the market's reaction to bankruptcy filings. *Accounting Review*, 117–128.

Clement, C. (2020). Machine learning in bankruptcy prediction–A review. *Journal of Public Administration, Finance and Law*, (17), 178–196.

DeAngelo, H., DeAngelo, L., & Skinner, D. J. (1994). Accounting choice in troubled companies. *Journal of Accounting and Economics, 17*(1/2), 113–143.

DeFond, M. L., & Jiambalvo, J. (1994). Debt covenant violation and manipulation of accruals. *Journal of Accounting and Economics, 17*(1–2), 145–176. https://doi.org/10.1016/0165-4101(94)90008-6

Dielman, T. E., & Oppenheimer, H. R. (1984). An examination of investor behavior during periods of large dividend changes. *Journal of Financial and Quantitative Analysis, 19*(2), 197–216. https://doi.org/10.2307/2330898

Foster, B. P., Ward, T. J., & Woodroof, J. (1998). An analysis of the usefulness of debt defaults and going concern opinions in bankruptcy risk assessment. *Journal of Accounting, Auditing & Finance, 13*(3), 351–371. https://doi.org/10.1177/0148558X9801300311(Originalworkpublished1998)

Gajpal, P. P., Ganesh, L. S., & Rajendran, C. (1994). Criticality analysis of spare parts using the analytic hierarchy process. *International Journal of Production Economics*, *35*, 293–297.

Galindo, J., & Tamayo, P. (2000). Credit risk assessment using statistical and machine learning: Basic methodology and risk modeling applications. *Computational Economics, 15*, 107–143. https://doi.org/10.1023/A:1008699112516

Gepp, A., Kumar, K., & Bhattacharya, S. (2010). Business failure prediction using decision trees. *Journal of Forecasting, 29*, 536–555. https://doi.org/10.1002/for.1153

Gilson, S. C., John, K., & Lang, L. H. (1990). Troubled debt restructurings: An empirical study of private reorganization of firms in default. *Journal of Financial Economics, 27*(2), 315–353. https://doi.org/10.1016/0304-405X(90)90059-9

Giroux, G. A., & Wiggins, C. E. (1983). Chapter XI and corporate resuscitation. *Financial Executive*, 36-41.

Giroux, G. A., & Wiggins, C. E. (1984). An events approach to corporate bankruptcy. *Journal of Bank Research, 15*(3), 179–187.

Golbayani, P., Florescu, I., & Chatterjee, R. (2020). A comparative study of forecasting corporate credit ratings using neural networks, support vector machines, and decision trees. *The North American Journal of Economics and Finance, 54*, 101251. https://doi.org/10.1016/j.najef.2020.101251

Gupta, V. (2022). Bankruptcy prediction using machine learning techniques: Evidence on Indian companies under insolvency and bankruptcy code. *Journal of Prediction Markets, 16*(2), 77–100.

Halteh, K., Alkhouri, R., Ziadat, S., & Haddad, F. (2024). Fintech Unicorns Forecaster: An AI Approach For Financial Distress Prediction. *Migration Letters*, *21*(S4), Article S4.

Hellwig, K.-P. (2021). Predicting fiscal crises: A machine learning approach. IMF Working Papers. 150.

Henley, W., & Hand, D. J. (1996). Ak-nearest-neighbour classifier for assessing consumer credit risk. *Journal of the Royal Statistical Society Series d: the Statistician, 45*(1), 77–95. https://doi.org/10.2307/2348414

Hsieh, S.-J. (1993). A note on the optimal cutoff point in bankruptcy prediction models. *Journal of Business Finance & Accounting, 20*(3), 457–464. https://doi.org/10.1111/j.1468-5957.1993.tb00268.x

Jain, A. K., & Gupta, B. B. (2018). Towards detection of phishing websites on client-side using machine learning based approach. *Telecommunication Systems, 68*, 687–700. https://doi.org/10.1007/s11235-017-0414-0

Japkowicz, N., & Shah, M. (2011). *Evaluating Learning Algorithms: A Classification Perspective*. Cambridge University Press. https://doi.org/10.1017/CBO9780511921803

John, T. A., & John, K. (1993). Top-management compensation and capital structure. *The Journal of Finance, 48*, 949–974. https://doi.org/10.1111/j.1540-6261.1993.tb04026.x

Jones, S., & Hensher, D. A. (2004). Predicting firm financial distress: A mixed logit model. *The Accounting Review, 79*, 1011–1038. https://doi.org/10.2308/accr.2004.79.4.1011

Khemakhem, S., & Boujelbene, Y. (2018). Predicting credit risk on the basis of financial and non-financial variables and data mining. *Review of Accounting and Finance, 17*(3), 316–340. https://doi.org/10.1108/RAF-07-2017-0143

Li, K. (2024). Does innovation relieve corporate financial distress? *Asia-Pacific Financial Markets*. https://doi.org/10.1007/s10690-023-09445-4

Li, Z., Crook, J., & Andreeva, G. (2014). Chinese companies distress prediction: An application of data envelopment analysis. *The Journal of the Operational Research Society, 65*(3), 466–479.

Liang, C., Berant, J., Le, Q., Forbus, K.D., and Lao, N. (2016). Neural symbolic machines: learning semantic parsers on freebase with weak supervision. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 23–33, Vancouver, Canada. Association for Computational Linguistics.

Liang, D., Tsai, C.-F., & Wu, H.-T. (2015). The effect of feature selection on financial distress prediction. *Knowledge-Based Systems, 73*, 289–297. https://doi.org/10.1016/j.knosys.2014.10.010

Lu, D., Popuri, K., Ding, G. W., et al. (2018). Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer's disease using structural MR and FDG-PET images. *Science and Reports, 8*, 5697. https://doi.org/10.1038/s41598-018-22871-z

Malakauskas, A., & Lakštutienė, A. (2021). Financial distress prediction for small and medium enterprises using machine learning techniques. *Engineering Economics, 32*(1), 4–14. https://doi.org/10.5755/j01.ee.32.1.27382

Martin, D. (1977). Early warning of bank failure: A logit regression approach. *Journal of Banking & Finance, 1*(3), 249–276. https://doi.org/10.1016/0378-4266(77)90022-X

Mironiuc, M., Carp, M., & Chersan, I. C. (2015). The relevance of financial reporting on the performance of quoted romanian companies in the context of adopting the IFRS. *Procedia Economics and Finance, 20*, 404–413. https://doi.org/10.1016/S2212-5671(15)00090-8

Morum, A., & Roy, D. (2013). Financial indicators of corporate sickness: A study of Indian steel industry. *South Asian Journal of Management, 20*(2), 85–101.

Nair, J., & Sachdeva, J. K. (2016). Indicators of financial distress: An empirical study of Indian Textile sector. *Journal of Global Economy, 12*, 101–110. https://doi.org/10.1956/jge.v12i2.418

National Industrial Classification | Ministry of Statistics and Program Implementation | Government Of India (mospi.gov.in)

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research, 18*, 109–131. https://doi.org/10.2307/2490395

Ooghe, H., & De Prijcker, S. (2008). Failure processes and causes of company bankruptcy: A typology. *Management Decision, 46*(2), 223–242. https://doi.org/10.1108/00251740810854131

Özdamar, K. (2002). Statistical data analysis with package programs-1 (4. Print). *Kaan Bookstore*.

Paula, L. F., Fritz, B., Prates, D. (2020). The metamorphosis of external vulnerability from 'original sin' to 'original sin redux': Currency hierarchy and financial globalisation in emerging economies. UFRJ, Instituto de Economia, Discussion Paper 033, November

Sehgal, S., Mishra, R. K., Deisting, F., & Vashisht, R. (2021). On the determinants and prediction of corporate financial distress in India. *Managerial Finance, 47*(10), 1428–1447. https://doi.org/10.1108/MF-06-2020-0332

Shrivastava, A., Kumar, K., & Kumar, N. (2018). Business distress prediction using bayesian logistic model for indian firms. *Risks, 6*(4), 113. https://doi.org/10.3390/risks6040113

Shrivastava, A., Kumar, N., Kumar, K., & Gupta, S. (2020). Corporate distress prediction using random forest and tree net for India. *Journal of Management and Science, 1*(1), 1–11. https://doi.org/10.26524/jms.2020.1

Sinha, N., Kakkar, N. K., & Gupta, V. (2009). Unleash the power of creativity and innovation. *International Journal of Sustainable Strategic Management, 1*, 417.

Smith, C. W., Jr., & Warner, J. B. (1979). On financial contracting: An analysis of bond covenants. *Journal of Financial Economics, 57*(2), 117–161. https://doi.org/10.1016/0304-405X(79)90011-4

Su, E.-D., & Huang, S.-M. (2010). Comparing firm failure predictions between logit, KMV, and ZPP models: Evidence from Taiwan's electronics industry. *Asia-Pacific Financial Markets, 17*(3), 209–239. https://doi.org/10.1007/s10690-010-9113-5

Suresh, K., Saxena, A., & Srikanth, M. (2023). Comparing financial debt choices of existing and new SMEs in Indian manufacturing sector. *Asia-Pacific Financial Markets, 30*(3), 445–456. https://doi.org/10.1007/s10690-022-09382-8

Tang, H., & Zhang, C. (2019). Investment risk, return gap, and financialization of non-listed non-financial firms in China$_*$. *Pacific-Basin Finance Journal*, *58*, 101213.

Turetsky, H. F. (1997). *An empirical investigation of firm longevity: A model of the ex ante predictors of financial distress*. Virginia Commonwealth University.

Turetsky, H. F., & McEwen, R. A. (2001). An empirical investigation of firm longevity: A model of the ex ante predictors of financial distress. *Review of Quantitative Finance and Accounting, 16*, 323–343. https://doi.org/10.1023/A:1011291425075

Varadraj, B., & Abhay, N. (2014). Comparison of bankruptcy prediction models: Evidence from India. *Accounting and Finance Research, 3*, 5743–5743. https://doi.org/10.5430/afr.v3n4p91

www.ibbi.gov.in

Yeh, I. C., & Lien, C. H. (2009). The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. *Expert Systems with Applications, 36*(2), 2473–2480. https://doi.org/10.1016/j.eswa.2007.12.020

Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research, 22*, 59–82. https://doi.org/10.2307/2490859

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Indian Journal of Microbiology
https://doi.org/10.1007/s12088-025-01477-w

ORIGINAL RESEARCH ARTICLE

Production of Phytase by *Penicillium oxalicum* in Solid State Fermentation Exhibiting Hydrolysis of Phytates and Haloperoxidase Activity

Priya[1] · Bijender Singh[2,3] · Ravi Kumar Goswami[4] · Jai Gopal Sharma[1] · Bhoopander Giri[5]

Received: 20 July 2024 / Accepted: 8 April 2025
© Association of Microbiologists of India 2025

Abstract
Penicillium oxalicum PBG30 produced an extracellular phytase in solid-state fermentation and maximum production was obtained (200 ± 6.01 U/g DMR) at 30 °C after 5 days. The optimal temperature and pH for enzyme were 70 °C and 3.0, respectively. The phytase is thermostable with $T_{1/2}$ of 1 h at 70 °C and showed broad-substrate specificity with K_m and V_{max} values of 4.42 mM and 909.1 U/ml, respectively with calcium phytate. Phytase activity was enhanced in the existence of metal ions and surfactants and retarded by SDS, EDTA, sodium molybdate, DTT, and ß-ME. Phytase exhibited resistance against trypsin and pepsin with better storage stability at 4 °C and −20 °C. Insoluble phytates (metal and protein) were efficiently hydrolyzed by fungal phytase showing liberation of inorganic phosphate in a time-dependent manner. Also, the phytic acid reduction was observed in phytase-treated fish feed with the lowest phytic acid occurring in a diet supplemented with 1500 FTU/kg dose of phytase. Furthermore, phytase was converted into vanadium-dependent peroxidase. Fungal phytase, due to its thermostability, protease resistance, broad substrate specificity, and ability to hydrolyze phytate forms, has the potential to serve as an additive for improving nutrient digestibility in the food and feed industry.

Keywords Phytase · *Penicillium oxalicum* · Thermostable · Protease-resistant · Insoluble phytates · Fish feed · Haloperoxidase

Introduction

Phytase is a well-known feed enzyme in today's world that hydrolyses the phytic acid into lower esters of myo-inositol and inorganic phosphate. It also releases other phytate-associated nutrients necessary for the organisms [1]. Phytic acid or phytate is a precursor of organic phosphorus and is predominantly present in the plant's seeds. Phytate, the complex form of phytic acid binds metal ions, amino acids, proteins, digestive enzymes, and other biomolecules. Due to the chelation property, phytate acts as an anti-nutritive factor [2]. Phytic acid can be reduced by physical and chemical methods which decreases the nutritive value, hence enzymatic action (phytase) is required for the degradation of phytic acid [3]. In ruminant animals, phytase is found in the microbes of the gut region while it is absent in monogastric animals [4]. Since monogastric animals lack phytase in the digestive tract, phytate molecules accumulate inside the body and are excreted out in the soil via fecal matter [1]. In the soil, various phytase-containing microbes are present that break phytate molecules and liberate phosphorus which goes to water bodies by run-off or leaching [5]. The released phosphorus forms algal blooms that lead to eutrophication and phosphorus pollution in water. Therefore, supplementation of phytase in monogastric animals as a feed additive is

✉ Bhoopander Giri
bhoopg@yahoo.com

1 Environmental and Industrial Biotechnology Laboratory, Department of Biotechnology, Delhi Technological University, Delhi, India

2 Laboratory of Bioprocess Technology, Department of Microbiology, Maharshi Dayanand University, Rohtak, Haryana, India

3 Department of Biotechnology, Central University of Haryana, Mahendargarh, Haryana, India

4 Department of Zoology, Hindu College, University of Delhi, Delhi, India

5 Department of Botany, Swami Shraddhanand College, University of Delhi, Delhi, India

Published online: 18 April 2025

requisite for their growth and development. Phytase ameliorates the nutritive value of feed and also reduces the level of phosphorus in the environment thus minimizing water pollution. It also limits the need for inorganic phosphorus and therefore restricts the load on non-renewable rock phosphate [4]. Although phytase enzymes have been reported from many microorganisms (bacteria, fungi, yeasts) they are still lacking in fulfilling the properties of an ideal phytase [6]. Thermostable, protease-resistant, broad pH range, substrate-specificity, long shelf life, and cost-effectiveness are the properties of an ideal phytase. The commercially available phytases either lack one or more features of an ideal phytase [4]. Therefore, the necessity of an ideal phytase in the industries is continuously increasing.

Solid-state fermentation (SSF) is the most widely adopted cost-effective method that uses agricultural waste as a substrate and reduces its accumulation. SSF is a simple process to operate and maintain for fungal enzyme production. It requires less water, minimal oxygen supply and less energy, while producing higher yields compared to other fermentation methods. The parameters used in the SSF significantly influence microbial growth and enzyme production, making the selection of appropriate physio-chemical factors a critical aspect [7]. Purification and characterization are important aspects that demonstrate the properties of phytase enzyme and thus need to be studied. Purification includes precipitation and dialysis followed by chromatography techniques [8]. Purification and biochemical characterization of phytase from various fungal species have been reported in the literature [9–15].

Phytic acid or phytate complexes are commonly present in plant-based food and feed ingredients which are insoluble in the GIT pH of monogastric animals. Consequently, monogastric animals are unable to utilize certain nutrients, resulting in nutrient deficiencies that can lead to malnutrition and severe health complications. The addition of phytase to the diet releases bounded nutrients and ameliorate the nutritional quality of the food and feedstuffs [16]. The application of phytase in lowering phytic acid amount has been studied in poultry feed ingredients [15], and wheat flour [12, 15, 17]. Also, the liberation of inorganic phosphate, reducing sugar and protein content was observed along with the reduction of phytic acid [16–18].

Phytase can act as haloperoxidase when vanadate is incorporated with phytase. Vanadium is an analog of phosphate molecules and an inhibitor of acid phosphatase. It binds to the active site of histidine acid phosphatase (HAP)-phytase and converts it into peroxidase. Vanadium haloperoxidase is similar in structure to phytase except for the presence of vanadate instead of phosphate. This semi-synthetic haloperoxidase is involved in various catalytic oxidative reactions [19, 20].

The present study deals with the production of extracellular phytase obtained from *Penicillium oxalicum* PBG30 using agricultural waste during the SSF method followed by the biochemical characterization to determine the properties of phytase. The involvement of phytase in hydrolysing insoluble metal- and protein-phytate complexes commonly found in food and feed ingredients as well as its role in degrading phytic acid in the fish feed, was studied to assess its potential as an additive in the food and feed industry. Additionally, the conversion of phytase enzyme to semi-synthetic haloperoxidase using vanadium was examined.

Materials and Methods

Culture Conditions

Penicillium oxalicum PBG30 was cultured on the potato dextrose agar (PDA) media at 30 °C and maintained at 4 °C on PDA slants, and also at − 20 °C in glycerol stock. The fungal spores were scrapped from 72 h old culture via a sterile loop by adding 25 ml normal saline solution containing 0.1% Tween 80 and the spore count was estimated by a hemocytometer and 1 ml of spore suspension (7.7×10^7 CFU/ml) was used as inoculum for the production of phytase [18, 21].

Phytase Production in SSF

Ten gram of wheat bran was used with 20 ml production medium of pH 7.0 consisting 0.5% Urea, 0.1% $MgSO_4.7H_2O$, 0.1% KCl and 0.1% $FeSO_4.7H_2O$ [22]. The medium was autoclaved for 20 min at 121 °C and then allowed to cool at room temperature. One ml of spore suspension was added and incubated for 5 days at 30 °C. The culture filtrate was extracted and centrifuged at 10,000 rpm for 10 min at 4 °C. The phytase activity was calculated from the supernatant using phytase assay [18, 21].

Phytase Assay

Phytase activity was calculated by estimating the liberation of inorganic phosphate from the calcium phytate substrate at 70 °C with 0.1 M sodium acetate buffer (pH 3.0) through the Fiske and Subbarow [23] method [18]. Each 1.0 ml reaction mixture contained 0.5 ml substrate dissolved in buffer and 0.5 ml enzyme diluted with buffer. Blank and control (substrate and enzyme) were also prepared during the experiment and performed in triplicate. The substrate control consisted of 0.5 mL substrate with 0.5 mL buffer, while the enzyme control consisted of 0.5 mL enzyme with 0.5 mL buffer. The tubes were incubated for 5 min at 70 °C in a water bath. Following this, 1.0 mL of TCA, 1.0 mL of distilled water, and 0.5 mL of 9 N H_2SO_4 were added, and the tubes

were incubated for additional 10 min at room temperature. After incubation, 0.5 ml of ammonium molybdate (6%) was added and the tubes were kept for 15 min in the dark. Then, 0.5 mL of freshly prepared $FeSO_4$ was added and the tubes were kept for 30 min in the dark. The absorbance was measured at 660 nm [23]. Phytase activity was calculated using the regression equation derived from the standard curve of KH_2PO_4 [25]. One unit indicates the amount of phytase used to liberate 1 nmol of inorganic phosphorus per sec under assay conditions [24]. The phytase activity is expressed as a unit per gram of dry mouldy residue (U/g DMR).

Biochemical Characterization of Phytase

For characterization, the enzyme was concentrated using ammonium sulphate precipitation method. The crude enzyme solution was placed in a beaker and stirred using a magnetic stirrer. Ammonium salt was added slowly to the extract until it reached 80% saturation. The entire procedure was conducted at 4 °C, followed by overnight incubation to allow the precipitated enzyme to settle. The sample was centrifuged at 10,000 rpm for 10 min at 4 °C and the pellet was re-suspended in 0.1 M sodium acetate buffer (pH 3.0). Dialysis was then carried out at 4 °C. A suitable length of dialysis membrane was cut and pre-treated to activate it. After cooling, the membrane was filled with the precipitated enzyme and tied at both ends. The dialysis bag was submerged in 20 mM sodium acetate buffer (pH 3.0) and kept on a magnetic stirrer at 4 °C. The buffer was changed at least three times. After dialysis, the sample was centrifuged at 10,000 rpm for 10 min at 4 °C. The supernatant was collected and the pellet discarded. The clear supernatant was used for characterization studies [9].

Effect of Temperature on the Activity and Stability of Phytase

To determine the optimum value of temperature, a phytase assay was carried out from 30 °C to 80 °C in 0.1 M sodium acetate buffer (pH 3.0) with 1 mM calcium phytate. To investigate the thermostability of phytase, the enzyme was incubated at 70° and 80 °C in a buffer of pH 3.0 for 2 h. Samples were collected at regular intervals and the assay was performed at 70 °C (pH 3.0) [9, 18].

Impact of pH on the Activity and Stability of Phytase

To evaluate the optimum value of pH, phytase enzyme activity was studied at 70 °C with a buffer of different pH range 2.5—8.0 (0.1 M glycine–HCl buffer pH 2.5, 0.1 M sodium acetate buffer pH 3.0—6.0, 0.1 M tris–HCl buffer pH 7.0—8.0). For assessing the pH stability of phytase, the enzyme was dissolved in pH (3.0, 5.0) and incubated at 70 °C. Samples were collected at 0, 5, 10, 30, 60, and 120 min interval and the assay was performed at 70 °C (pH 3.0) [9, 18].

Determination of Substrate Specificity and Kinetic Constants

The specificity of phytase towards different substrates, viz, sodium phytate, calcium phytate, glucose-6-phosphate, p-nitrophenyl phosphate, adenosine triphosphate (ATP), adenosine diphosphate (ADP), and adenosine monophosphate (AMP) was examined. The phytase assay was performed at 70 °C with a pH 3.0 and 1 mM substrate concentration, using calcium phytate as the substrate, with its activity considered as 100%. The kinetic parameters (K_m and V_{max}) were evaluated from the Lineweaver–Burk plot by using different concentrations of calcium phytate (0.1—2 mM) and the assay was conducted at 70 °C (pH 3.0) [9, 18].

Effect of Organic Solvents, Detergents, Metal ions, and Inhibitors on Phytase Activity

The experiments were performed to determine the influence of organic solvents (ethanol, methanol, butanol, propan-2-ol, acetone, DMSO) and detergents (SDS, Tween-20, Tween-60, Tween-80, Triton-X) on phytase activity. In the reaction mixture, different concentrations of organic solvents (2 and 5%) and detergents (0.1 and 0.5%) were added [9].

Similarly, the effect of metal ions and inhibitors on phytase activity was analyzed by providing different metal ions (Ca^{2+}, Co^{2+}, Cu^{2+}, Mg^{2+}, Mn^{2+}, Fe^{3+}, Fe^{2+}, Na^{+}) and inhibitors (sodium azide, EDTA, sodium molybdate, ß-ME, DTT) with 1 mM and 5 mM concentrations in the reaction mixture.

The enzymatic assay was performed along with the control. The control value is taken as 100% having no additional chemical [9, 26].

Effect of Proteolytic Enzymes and Storage Conditions on Phytase Activity

The phytase enzyme was treated with 1% (w/v) trypsin and pepsin for 2 h at 37 °C. The samples were taken at the desired interval (0, 30, 60, 120 min) and a phytase assay was performed [18, 25]

Phytase was kept for 6 months in different conditions (room temperature, 4 °C, −20 °C), and activity was determined at the desired interval. Activity on the 0 th day is considered 100% [9].

Hydrolysis of Insoluble Phytates

Metal phytates (Ca^{2+}, Co^{2+}, Mg^{2+}, Mn^{2+}, Fe^{2+}, Cu^{2+}, Zn^{2+}) were formed by mixing each metal ion along with sodium phytate. 100 mM stock solution of metal ions ($CaCl_2{\cdot}2H_2O$, $CoCl_2{\cdot}2H_2O$, $FeSO_4{\cdot}7H_2O$, $MgCl_2{\cdot}6H_2O$, $MnCl_2{\cdot}4H_2O$, $ZnCl_2{\cdot}7H_2O$) and 10 mM of sodium phytate were prepared in distilled water separately. Both were mixed in equal volume and kept at 4 °C for overnight incubation. The precipitated salts were centrifuged at 10,000 g for 5 min at 4 °C. The supernatant was discarded and the pellet was washed three times with distilled water, and finally suspended in 5 ml of 0.1 M Na-acetate buffer (pH 3.0). Each phytate salt was treated with 20 U of phytase and incubated at 37 and 50 °C. Samples were taken at different intervals (3, 6, 12, 24 h) and salts were pelleted by centrifuge. The supernatant was used to determine the amount of released inorganic phosphate at 600 nm [25]. The protein-phytate complex was made by adding an equal amount of lysozyme (2.5 mM) in distilled water and sodium phytate (3 mM) in distilled water and kept overnight at 4ºC. Centrifuge the solution at 10,000 g for 5 min at 4ºC and discard the supernatant. Washed the pellet three times with distilled water, dissolved in buffer and 20 U of phytase was added into it. The reaction was incubated at 30ºC and samples were collected at defined intervals (0, 5, 10, 30, 60, 120, 360 min). Centrifuge the sample and inorganic phosphate was estimated in the supernatant at 600 nm [25].

Phytic Acid Reduction in Fish Feed Using Fungal Phytase

The formulation of the feed sample is mentioned in Table 1 containing 35% protein in the diet. Different doses of phytase were added to the feed ingredients (500 FTU/kg, 1000 FTU/kg, 1500 FTU/kg). The pellet of different diets was formed with the help of pelletizer. The determination of phytic acid in the phytase-treated and untreated feed pellet was performed by Gao et al. [27] method. Firstly, the sample was treated with 10 ml of 2.4% HCl and kept on a shaker for 16 h at 220 rpm. Centrifuge the sample for 20 min with 10,000 rpm at 10 °C and the obtained supernatant was shifted to falcon consisting of 1 g NaCl. Again, the mixture was shaken at 350 rpm for 20 min so that salts dissolved completely and settled by keeping it at 4 °C for 60 min or −20ºC for 20 min. After that, centrifugation was done at 10,000 rpm for 20 min. The supernatant was used to estimate the phytic acid content by adding 500 µl Wade's reagent (containing 0.03% ferric chloride and 0.3% sulfosalicylic acid) and kept for 20 min incubation. The absorbance was taken at 500 nm [20, 28]. The quantity of phytic acid in the diet was calculated through the regression equation of the standard curve of sodium phytate.

Transformation of Phytase into Haloperoxidase

Phytase was mixed with ammonium metavanadate (5 and 10 µM) and incubated at 30 °C for 4 h with 100 rpm. H_2O_2 and guaiacol were added to the solution and absorbance was recorded at 470 nm for each minute up to 10 min [20]. The relative activity of phytase was also calculated during the experiment.

Statistical Analysis

All the experiments were carried out in triplicate. The values were reported as the mean and standard error (mean ± S.E). The significance level was determined using a t-test and is expressed as a p-value. IBM SPSS Statistics software is used for the statistical analysis. Sigma plot (version 10.0) is used for creating the graphs.

Results and Discussion

Phytase Production in SSF

SSF is the commonly used approach nowadays for the production of enzymes. In this process, microbes are grown on a solid substrate in either the absence or a low amount of water [2]. The agricultural waste or residues are commonly utilized as a solid substrate that makes the production cost-effective and also resolves the issues that occur due to their accumulation [29]. SSF has been utilized by various fungi for the production of phytase [10, 16, 24, 30, 31]. The highest amount of phytase production was achieved (200.41 ± 6.01 U/g DMR) with 10 g wheat bran moistened with

Table 1 Composition of the control and experimental diets

Feed ingredients (g)	Diet I (control)	Diet II	Diet III	Diet IV
Soybean flour	181.57	181.57	181.57	181.57
Wheat flour	118.40	118.40	118.40	118.40
Fish oil	3.00	3.00	3.00	3.00
Vitamins and minerals	1.20	1.20	1.20	1.20
Phytase	-	500 FTU/kg	1000 FTU/kg	1500 FTU/kg

media (0.5% Urea, 0.1% $MgSO_4.7H_2O$, 0.1% KCl and 0.1% $FeSO_4.7H_2O$) in 1:2 ratio of pH 7.0, incubated at 30 °C for 5 days. The other fungal strains that produced phytase were *A. oryzae* (506.12 U/g DMR) [20], *A. awamori* NRC- F18 (437.195 ± 7.16 IU/g) [35], *A. niger* NT-7 (208.30 ± 0.22 U/g DMR) [16], *P. adiposa* (24.48 0.3 U/gds) [30]. These variations in phytase levels occurred due to the utilization of different fungi, substrates, production media, and fermentation conditions. Among the various substrates used for the SSF method, wheat bran is commonly used because it is rich in carbon, nitrogen, amino acids, and a high amount of phytic acid [32]. Due to the presence of phytate, it acts as an inducer for the production of phytase [33]. It is easily available, has a low cost, and provides support to the growth of fungi [16, 34]. The involvement of wheat bran in the production of phytase has been mentioned in various studies [16, 31, 35].

Biochemical Characterization of Phytase

Partial purification of *P. oxalicum* phytase was done by ammonium sulphate precipitation followed by the dialysis method. The crude sample expressed the lowest specific activity (0.31 U/mg) with a total enzyme activity of 381.57 U. After 80% ammonium sulphate precipitation, the sample showed 357.48 U activity with 1.23 U/mg specific activity. Further, the dialyzed sample exhibited 1.52 U/mg specific activity with total enzyme activity (211.04 U). The dialyzed enzyme was purified to homogeneity with 4.9-fold and a yield of 55.31%. Further, purification by column chromatography was omitted to make the process cost-effective. The phytase isolated from *A. niger* CFR 335 reported yields of 66.8% with a 4.1-fold and 71.9% with a 2.7 purification fold following the precipitation and dialysis methods, respectively [36]. Similarly, Saxena et al. [12] found 1.42-fold purification after the dialysis of *A aculeatus* APF1 phytase with a 25.54% yield. The phytase obtained from *T. purpureogenus* NSA20 was concentrated by precipitation and dialysis methods exhibiting yields of 39.8% and 11.9% with 2.6 and 5.9 purification folds, respectively [14]. Table 2 demonstrate the comparison of biochemical properties of phytases isolated from different fungi.

Effect of Temperature

During food and feed processing, enzymes have to face extreme temperatures, therefore they must be resistant to high temperature. The *P. oxalicum* phytase exhibited more than 80% activity at the temperature ranging between 55 to 80 °C (Fig. 1a), however, the optimal activity was recorded at 70 °C, showing its survival at high temperature. In other studies, the optimal activity of fungal phytases was observed between 30 and 65 °C [9, 12, 24, 28, 36, 41–45]. Ajith et al. (2019) found maximum activity of *A. foetidus* MTCC 11682 phytase at 37ºC [37]. However, the phytases obtained from *A. niger* BIONCL8 [15], *P. oxalicum* [38], and *A. fumigatus* exhibited optimal activity at 40ºC [11] while the phytases isolated from *R. oligosporus* MTCC 556 exhibited optimal activity at 50ºC [39], *A. fumigatus* at 55ºC [17], *A. niger* 7 A-1 at 56ºC [10] and *A. niger* NT7 at 60ºC [16].

The thermostability of phytase is studied at 70° and 80 °C. The *P. oxalicum* phytase is found thermostable at 70 °C with half-life of 60 min, while at 80 °C the phytase lost 70% activity within 60 min (Fig. 1b). However, the studies carried out on the *A. niger* BIONCL8 showed that it lost 70% phytase activity [15], and *A. fumigatus* lost 86% phytase activity [11] within 1 h at 80 °C. In the case of *A. niger* 7 A-1, only 9.3% activity is left after 5 min of incubation at 80 ºC [10] and *A. oryzae* retains only 4.2% phytase activity after 1 h [40]. A reduction in the phytase activity of *A. fumigatus* was observed after 1 h at 70 ºC [11]. In another study, phytase obtained from *A. fumigatus* showed 90% activity after 20 min while exhibiting more than 50% at 2 h [17]. The thermostable half-life of *A. oryzae* phytase was found at 300 min at 50 °C [9], while at 70 ºC phytase lost its activity within 100 min [24]. In the case of *P. oxalicum,* > 70% of activity was lost within 10 min [38].

Effect of pH

pH is a critical parameter that influences enzyme activity by changing the ionic composition in its microenvironment. The phytase is active in the acidic range (2.5–5.0) and found optimum at pH 3.0 (sodium acetate buffer) during the assay. The stomachs of humans and animals are acidic because their pH range between 1.5–3.5, therefore, the phytases active at this acidic pH are considered useful in the food and feed industry. Likewise, *A. aculeatus* APF1 [12], and *A. niger* [41] exhibited the optimum pH at 3.0. In contrast, neutral pH is mentioned in *P. oxalicum* EUFR-3 [38]. As shown in Fig. 2a, phytase activity declined in alkaline conditions due to changes in the ionic state of amino acids. The reduction of phytase activity in alkaline pH was observed in case of *R. oligosporus* MTCC 556 [39], *A. niger* NT7 [16], *A. niger* BIONCL8 [15], *A. niger* 7 A-1 [10] and *A. oryzae* [9, 24].

The stability of pH was checked with both pH 3.0 and 5.0. In pH 3.0, 50% phytase activity was retained for up to 60 min, while in pH 5.0, 50% activity was found within 40 min (Fig. 2b). In the case of *A. oryzae*, 80% phytase activity was reported with pH 3.0 and 5.0 for 45 min and 100 min respectively [9]. *A. niger* BIONCL8 phytase exhibited pH stability at 2.0 and 8.0. At pH 2.0, phytase showed 80% activity with incubation of 24 h, while at pH 8.0, it retained more than 50% activity [15].

Table 2 Comparison of properties of phytases isolated from various fungi

Fungal source	Optimum temperature (°C)	Optimum pH	Thermostability	Protease resistance after 120 min	Substrate specificity	References
Penicillium oxalicum PBG30	70	3.0	$t_{1/2}$ 60 min at 70 °C and 30% activity left after 1 h at 80 ºC	88% and 80% activity with pepsin and trypsin	sodium phytate, calcium phytate, p-nitrophenyl phosphate, ATP, ADP, Glucose-6-phosphate and AMP	Current study
P. oxalicum EUFR-3	40	7.0	5.35% activity left after 20 min at 80 °C	–	–	[38]
Aspergillus niger BIONCL8	40	2.1	30% activity left after 1 h of incubation at 80 °C	–	Sodium phytate	[15]
A. fumigatus	55	5.5	Retain > 50% activity after 2 h at 70 °C	–	Calcium phytate	[17]
A. niger NT7	60	2.6, 4.8	$t_{1/2} \leq$ 60 min at 70 °C	70.12% and 52.35% activity with pepsin and trypsin	Sodium phytate	[16]
A. tubingensis TEM 37	45	2.0, 5.5	Possess 100% activity after 3 h at 80 °C	Retain 88% and 98% activity with pepsin and trypsin	D-Glucose-6-phosphate, D-Fructose-6-phosphate, ATP, ADP, AMP, Sodium phytate	[13]
A. aculeatus APF1	50	3.0	Only 6–7% activity left after 5 min at 80 °C	Retain 74% and 73% activity with pepsin and trypsin	Calcium phytate	[12]
A. foetidus MTCC 11682	37	3.5, 5.5	56% activity after 30 min at 80 °C	–	–	[37]
A. fumigatus	40	6.0	20% and 14% activity left after 1 h at 70 °C and 80 °C	–	Sodium phytate	[11]
A. niger 7 A-1	56	5.3	9.3% activity left after 5 min at 80 ºC	–	Sodium phytate, Phenyl phosphate, Naphthyl phosphate, p-Nitrophenyl phosphate, Glycerol-2-phosphate, Glucose-1-phosphate, ADP and AMP	[10]
A. oryzae SBS50	50	5.0	$t_{1/2}$ 10 min at 80 °C	Retain 97% and 62% activity with pepsin and trypsin	sodium phytate, calcium phytate, potassium phytate, ADP, ATP, glucose-6-phosphate and p-nitrophenyl phosphate, β-glycerophosphate	[9]

Table 2 (continued)

Fungal source	Optimum temperature (°C)	Optimum pH	Thermostability	Protease resistance after 120 min	Substrate specificity	References
Rhizopus oligosporus MTCC 1116	50	5.5	55% and 30% activity left after 5 min at 70 °C and 80 °C	–	Sodium phytate, glucose-6-phosphate, glucose-1-phosphate, disodium pyrophosphate, ATP and GTP	[39]

Substrate Spectrum of Fungal Phytase

The phytase exhibited broad substrate specificity, a characteristic feature of an enzyme to be used for the animal feed. The highest specificity is observed with sodium phytate and considered as best substrate for *P. oxalicum* phytase. Other studies also demonstrated the maximum specificity of sodium phytate with phytases obtained from *P. oxalicum* KCTC6440 [41], *A. flavus* [28], *R. oligosporus* MTCC 556 [39], *A. niger* 7 A-1 [10]. The order of specificity is sodium phytate > calcium phytate > p-nitrophenyl phosphate > ATP > ADP > Glucose-6-phosphate > AMP (Fig. 3a).

The K_m and V_{max} were measured from the Lineweaver–Burk plot (Fig. 3b) and found to be 4.42 mM and 909.1 U/ml respectively. The K_m represents the enzyme affinity towards the substrate while V_{max} is the value at which all the active sites of enzymes were occupied by the substrate. The K_m and V_{max} values for *A. oryzae* phytase were 1.14 mM and 58.82 U/ml [9], and *A. aculeatus* APF1 phytase were 3.21 mM and 3.78 U/mg [12] with calcium phytate as a substrate. In other studies, the K_m and V_{max} values of phytase with sodium phytate were 0.545 mM and 600 U/mg for *P. oxalicum* PJ3 [43], 0.48 mM and 672 U/mg for *P. oxalicum* KCTC6440 [42], 0.815 mM and 1092 U/mg for *R. mucilaginosa* [44], 3.35 mM and 1.27 U/mg for *A. niger* BIONCL8 [15].

Effect of Organic Solvents, Detergents, Metal Ions and Inhibitors

Organic solvents have a positive impact on phytase activity. It increases in the presence of ethanol, methanol, butanol, propan-2-ol (isopropyl alcohol), acetone, and DMSO (Fig. 4a). Ethanol showed the maximum stimulatory effect on both 2 and 5% concentrations. The response of hydrophobic residues to the activity of phytase is negligible. In the existence of ethanol, methanol, butanol, and acetone, phytase activity was enhanced in *A. oryzae* [9]. The inhibitory effect of ethanol, butanol, isopropanol and acetone was observed in *A. niger* BIONCL8 [15]. In *A. aculeatus* APF1 phytase, butanol reduced the activity, acetone and isopropanol slightly affected while ethanol did not influence activity [12]. The functional group and molecular structure of organic solvents are accountable for designing the artificial environment around the enzyme which affects the enzyme structure and its catalytic properties. Depending upon the type of conformational modifications, phytase activity is affected. It can either stimulate or inhibit the enzyme activity [9]. Limited studies are available on the effect of organic solvents on phytase. Due to the presence of organic solvents solubility of hydrophobic residues is enhanced, it avoids reactions with water molecules, makes enzymes more stable

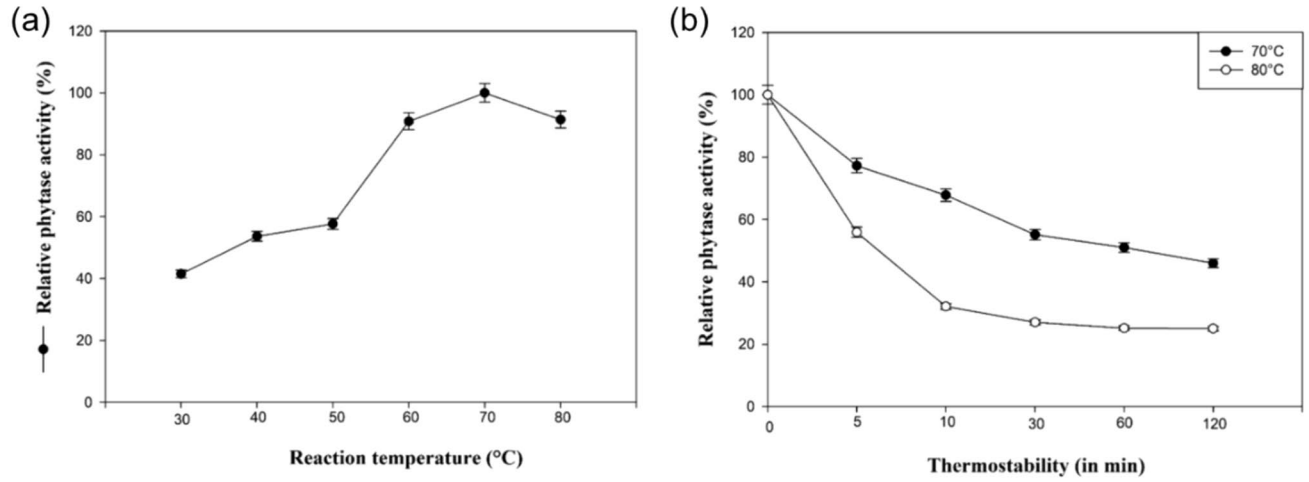


Fig. 1 **a** Effect of temperature range (30–80 °C) on phytase activity. **b** Thermostability studies at 70 and 80 °C

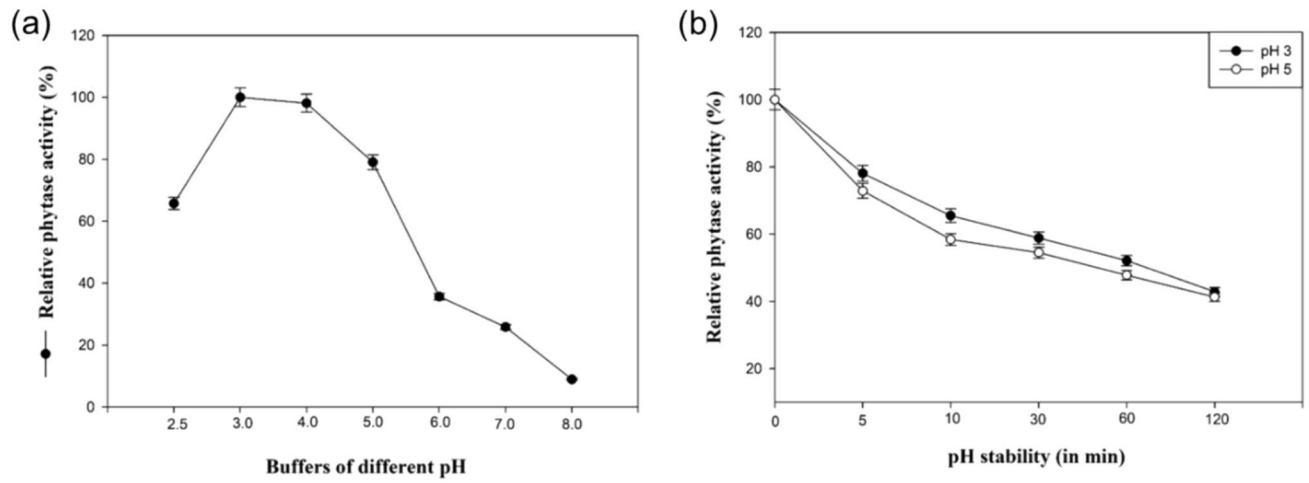


Fig. 2 **a** Effect of pH range (2.5–8.0) on phytase activity. **b** Effect of pH stability studies with pH 3.0 and 5.0

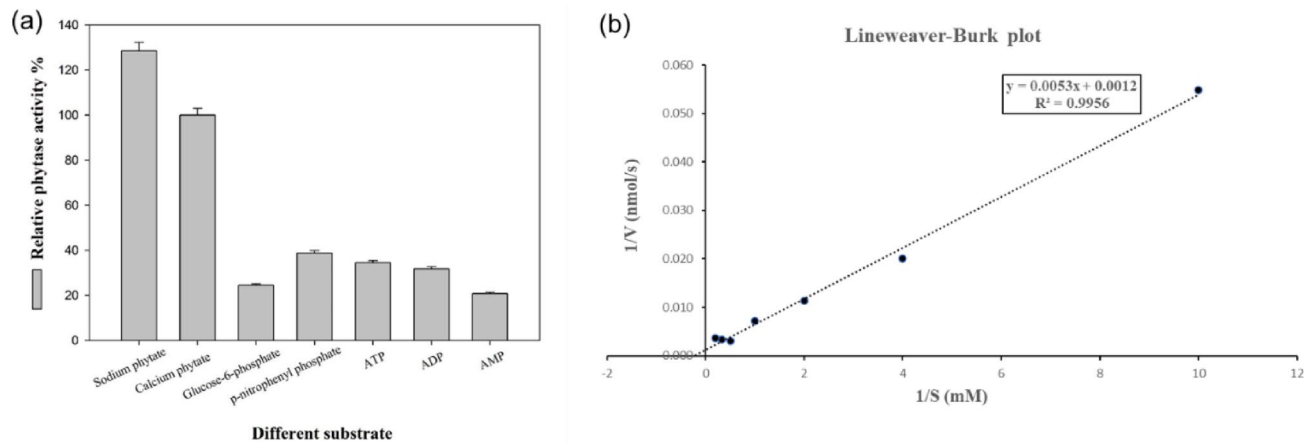


Fig. 3 **a** Effect of different substrate on phytase activity. **b** Lineweaver–Burk plot for K_m and V_{max}

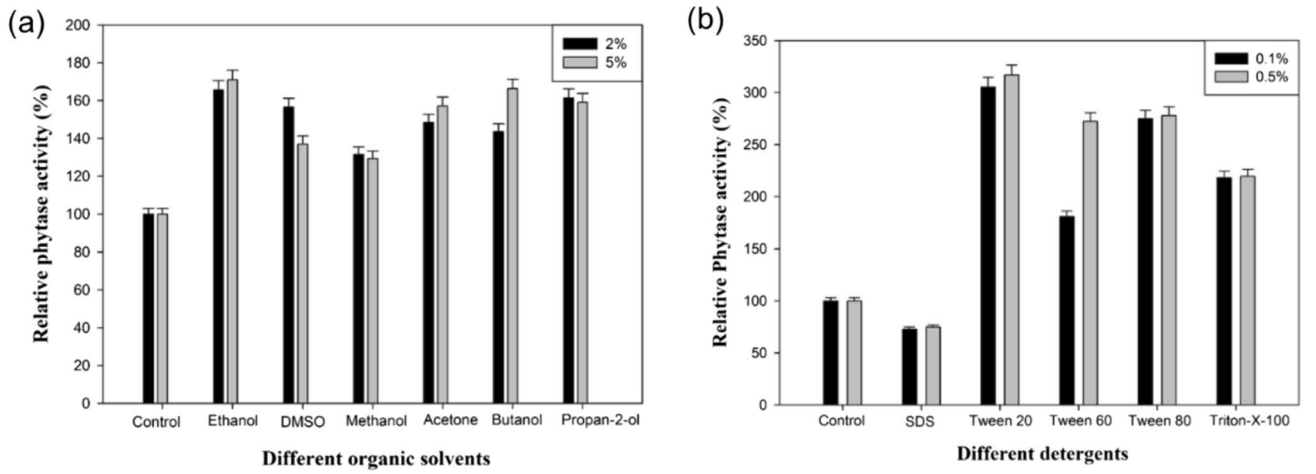


Fig. 4 **a** Effect of different organic solvents on phytase activity. **b** Effect of different detergents on phytase activity

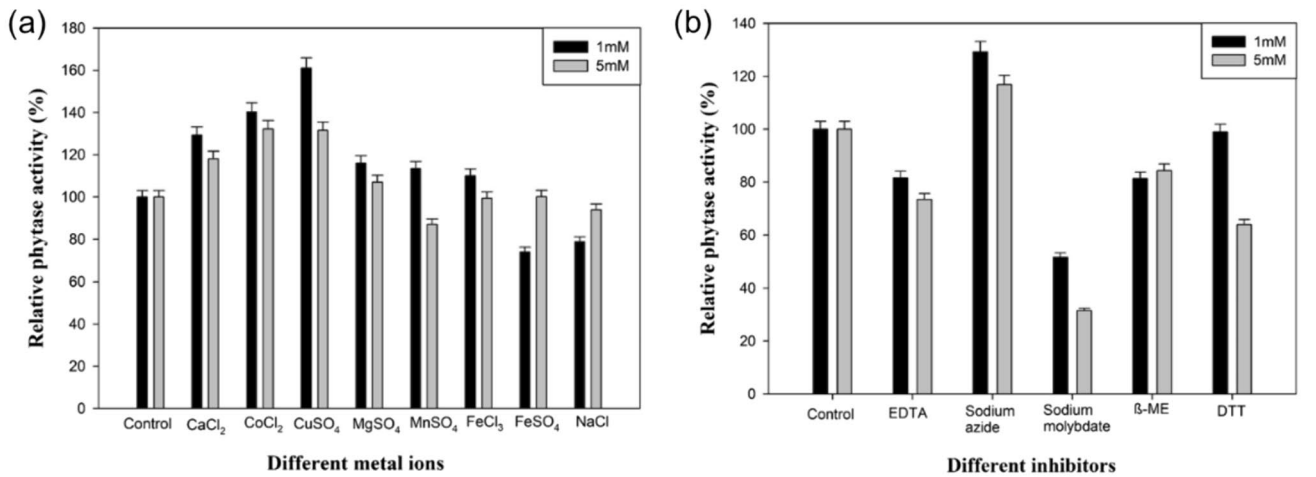


Fig. 5 **a** Effect of different metal ions on phytase activity. **b** Effect of different inhibitors on phytase activity

and reusable, and protects them from microbial contamination [13].

In our result, Tween 20, Tween-60 and Tween-80, Triton-X-100 increased the phytase activity, while SDS decreased the activity (Fig. 4b). SDS has been mentioned as an inhibitor for phytases obtained from *P. oxalicum* [38], *A. oryzae* [9], and *A. fumigatus* [11]. Contrasting to the result, Tween 20 showed the inhibitory action on *A. foetidus* MTCC 11682 phytase [37]. Detergents can also increase or decrease the enzyme activity depending upon the type of interacting molecule whether it is anionic, cationic, or non-ionic. It binds with protein molecules and changes the tertiary structure of protein [9]. SDS is an anionic detergent that is not specific and can attach anywhere to the protein and unfold the protein. It made conformational changes in the active site and thus inactivated the enzyme. Non-ionic detergents (Tweens and Triton X-100) do not bind on the surface of the protein and thus either stimulate or do not influence the activity of phytase by making conformational changes in the enzyme structure.

The involvement of metal ions in the regulation of enzyme catalysis is crucial. The influence of different metal ions was observed on phytase activity (Fig. 5a). Inhibition of phytase activity is possibly due to the association of metal and substrate complex which restricts the binding of substrate with enzyme or the modification of enzyme active site due to binding of ions [12], while enhancement in activity might be due to the requirement of metal ions in the activation of enzymes [44]. In the presence of Ca^{2+}, Co^{2+}, and Cu^{2+}, a significant enhancement is shown in the phytase activity which might be due to an increase in interaction between substrate and enzyme. Similarly, the involvement of Ca^{2+} in stimulating the phytase activity of *A. niger* NT7 [16], *A. fumigatus* [11], *A. oryzae* [9] has been mentioned

in the literature. Co^{2+} was also found to increase the phytase activity in *A. niger* NT7 [16], and *A. oryzae* [9] but in contrast, Cu^{2+} inhibited the phytase activity in *A. niger* NT7 [16], *A. oryzae* [9], and *A. niger* BIONCL8 [15]. Mg^{2+} represented a non-significant positive impact on the activity at both concentrations. A similar stimulatory effect of Mg^{2+} is shown in *A. oryzae* [9], and *A. niger* BIONCL8 [15]. At 1 mM concentration of Mn^{2+}, phytase activity was stimulated moderately but inhibited at high concentrations. Mn^{2+} decrease the activity of phytase obtained from *A. flavus* [28], and *A. niger* BIONCL8 [15]. With both concentrations of Na^{+} ions, activity was low as compared to the control during the study. The inhibitory effect of NaCl was observed with *P. oxalicum* [38] and *A. niger* BIONCL8 [15]. Dissimilar to this, Na^{+} enhances the activity in *R. mucilaginosa* [44]. There is no such effect of metal ions on the activity of *A. tubingensis* TEM 37 phytase [13].

The effect of inhibitors on phytase activity was also analyzed. Except for sodium azide, rest inhibitors (EDTA, ß-ME, DTT, sodium molybdate) were involved in suppressing the phytase activity as shown in Fig. 5b. Maximum inhibition was shown with sodium molybdate. In the literature, EDTA was involved in inhibiting the activity of phytase isolated from *P. oxalicum* [38], and *A. tubingensis* TEM 37 [13] indicating the requirement of metal ions for the activity. Dissimilar to this, EDTA increases the phytase activity in *A. niger* BIONCL8 [15] and has no impact on the phytase activity of *A. oryzae* [9], and *A. aculeatus* APF1 [12]. ß-ME and DTT were also reported to decrease the phytase activity of *A. niger* UFV-1 [45], and *P. oxalicum* [38], depicting the importance of sulfhydryl groups in the catalysis of phytase [45]. Sodium molybdate and sodium azide were involved in lowering the phytase activity of *A. oryzae* [9], and *A. tubingensis* TEM 37 [13] respectively.

Effect of Proteases

Pepsin and trypsin are two important digestive enzymes present in animals and human beings. For the utilization in animal feed, phytase must have the ability to resist the action of both enzymes. As shown in Table 3, there is no significant reduction in phytase activity against trypsin and pepsin treatment indicating the enzyme is protease resistant. Pepsin showed 8 and 12% reductions in phytase activity after incubation of 1 and 2 h respectively while trypsin exhibited 10–20% reduction in phytase activity within 1–2 h. The phytase isolated from *A. niger* UFV-1 exhibited more than 90% activity in the presence of pepsin and trypsin treatment for 1 h [45]. Similarly, the phytase obtained from *A. aculeatus* APF1 and *A. tubingensis* TEM 37 exhibited 74% and 73% and 88 and 98% activity after 120 min of treatment with pepsin and trypsin respectively [12, 13].

Table 3 Effect of proteases on phytase activity

Sample	Phytase activity (%)			
	0 min	30 min	60 min	120 min
Control	100	95 ± 0.98	92 ± 0.99	89 ± 0.32
Pepsin treated	100	99 ± 0.77	92 ± 0.09	88 ± 0.18
Trypsin treated	100	90 ± 1.24	90 ± 1.40	80 ± 1.31

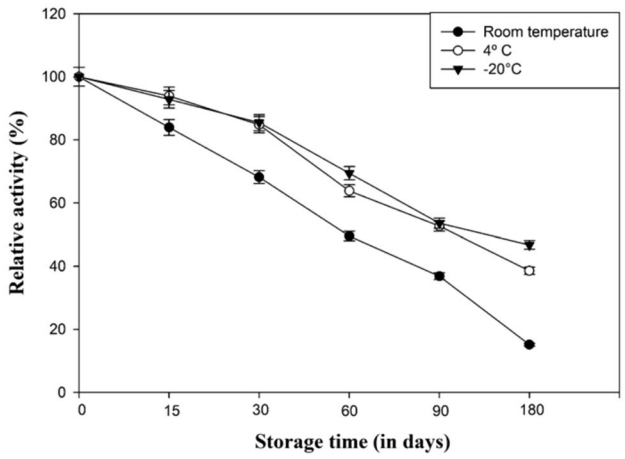


Fig. 6 Effect of storage conditions (RT, 4 °C and −20 °C) on phytase activity

Storage Life

The stability of the enzyme depends upon the molecule and storage conditions [9]. As shown in Fig. 6, 43% of the phytase activity is exhibited at −20 °C and 38% at 4 ℃ after 6 months, while 15% of the activity is left with phytase stored at room temperature. When the enzyme is stored in liquid, frozen, or powder form at different temperatures, it faces stress due to changing microenvironmental conditions and thus conformational modification occurs in the enzyme [9]. A similar study was observed in the shelf life of *A. oryzae* phytase and 4 ℃ was found as the most suitable storage condition retaining 63% phytase activity after 6 months, while 56% activity was exhibited with the sample kept at −20 °C [9]. *A. niger* CFR 335 showed more than 80% phytase activity at room temperature after the incubation of 3 months [36]. *P. adiposa* phytase expressed a reduction in its activity after incubation of 24 days at room temperature, 32 days at −20 °C, and 42 days at 4 °C [30].

Hydrolysis of the Insoluble Phytates

Phytic acid is associated with various metal ions and is available in the complex form of insoluble metal phytate. These insoluble complexes are unable to be digested by monogastric animals and restrict the assimilation of P and other essential metal ions. In the case of plants, phytic acid is adsorbed on clays or found in the form of insoluble salts of Fe and Al (acidic soils) or salts of Ca (alkaline soils). These insoluble salt forms are unable to be utilized hence must be hydrolyse by phytase before their absorption [25]. In the case of humans and animals, these insoluble metal phytate complex leads to malnutrition conditions [46]. As shown in Fig. 7, *P. oxalicum* PBG30 is capable of hydrolyzing insoluble metal phytate complex at both 37 and 50 °C. It releases inorganic phosphorus from the metal ions complex which increases with time. All the metal complexes of Ca^{2+}, Co^{2+}, Mn^{2+}, Cu^{2+}, and Zn^{2+} showed efficient hydrolysis and release of inorganic phosphate except Mg^{2+}, and Fe^{2+}. Similarly, the hydrolysis of insoluble salt phytates by *A. oryzae* was reported [25]. They showed that hydrolysis action occurred at 50 °C was better than 25 °C and liberation of inorganic phosphate was enhanced with time. *A. niger* NT-7 phytase was also capable of releasing metal ions from the insoluble metal phytate complex and showed more liberation of $Zn^{2+} > Ca^{2+} > Co^{2+} > Fe^{2+} > Fe^{3+}$ at 50 °C than 37 °C [46]. Phytate complexes of Ca, Mn, Mg, Zn, and Cd were hydrolysed more efficiently by *A. niger* phytase than the complexes of Al and Fe-phytates [47].

Phytic acid also makes complexes with proteins, amino acids, and enzymes and restricts their utilization. Plant-based proteins are usually found as insoluble

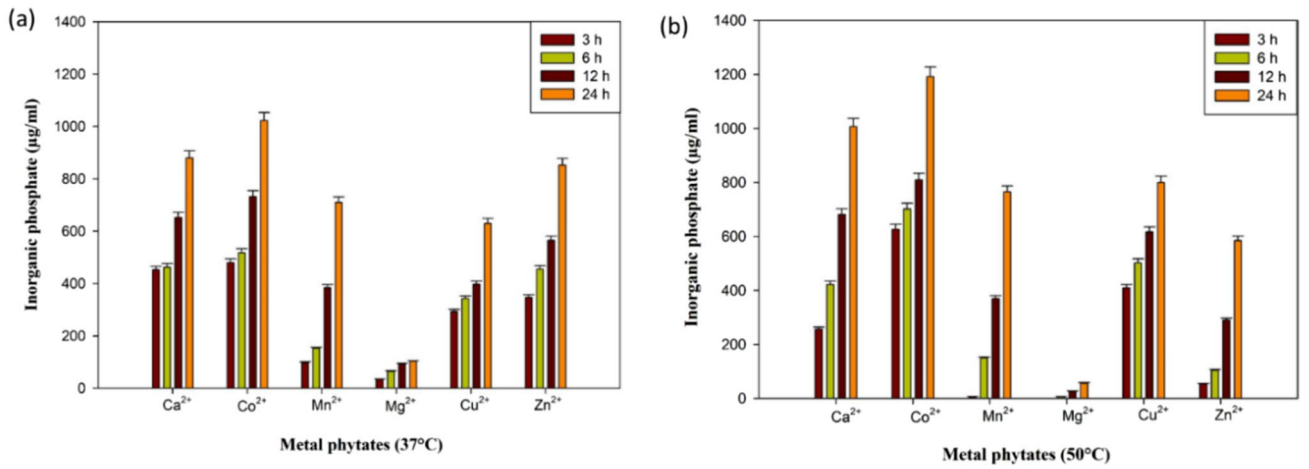


Fig. 7 Hydrolysis of insoluble metal phytates by *P. oxalicum* PBG30 phytase **a** at 37 °C **b** 50 °C

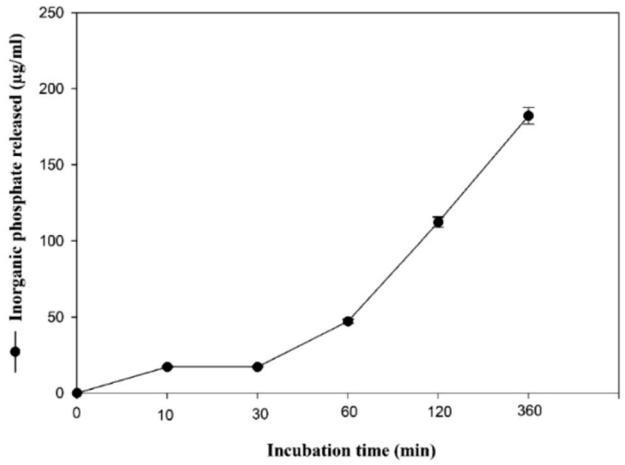


Fig. 8 Hydrolysis of lysozyme-phytate complex by *P. oxalicum* PBG30 phytase at 30 °C

protein-phytate complex forms and are thus beyond the approach of monogastric animals. Due to the phytate-protein complex, turbidity occurred in the solution [46]. In the present study, a turbid solution was observed when sodium phytate was mixed with lysozyme. The release of inorganic phosphate was increased with time as shown in Fig. 8. Similar work was performed on the hydrolysis of phytate association with BSA and lysozyme by using the phytase of *A. oryzae*. Further, they described the time-dependent relation between inorganic phosphorus and the absorbance of the solution. Inorganic phosphorus was increased while absorbance was decreased with time [25]. Kumari and Bansal [46] also studied the relation of the absorbance of turbid solution with the release of inorganic phosphate from the phytate-lysozyme complex in the presence of *A. niger* NT-7.

Phytic Acid Reduction in Fish Feed

In the plant-based feed, phytic acid is present in major amounts. Phytase hydrolyzes the phytic acid and liberates inorganic phosphate along with associated minerals from the feed. The application of phytase in the formulation of fish feed is analyzed to determine the reduction of phytic acid in feed. The phytase-treated sample showed less amount of phytic acid as compared to the untreated sample. As shown in Table 4, the control diet (without the addition of phytase) showed 50.43 mg/g phytic acid which decreases as the dosage of phytase increases in the diet. The sample with 1500 FTU/kg exhibited minimum phytic acid. Similar to this, *S. cerevisiae* MTCC 5421 phytase is mixed with refined wheat flour and pearl millet flour before forming naan and rabadi, and the amount of phytic acid was measured in control food (without phytase) and phytase-treated food. Naan and rabadi prepared with phytase enzyme showed 95% and 100% reduction of phytic acid respectively [48]. In other reports, there was a reduction of phytic acid amount in bread (70%), tandoori (75%), and naan (62.5%) prepared from dough of wheat flour supplemented with recombinant phytase of *Sporotrichum thermophile* [49]. Also, tandoori and naan made from recombinant phytase of *Pichia anomala* showed a 67.5% and 23.2% decrease in phytic acid respectively [50].

Haloperoxidase Activity

Phytase acts like a haloperoxidase when reacted with ammonium metavanadate and enhances the peroxidase activity while decreasing its activity (Fig. 9). The mechanism behind that is vanadium ion gets incorporated in the active site of phytase and makes a semi-synthetic peroxidase. This peroxidase tends to catalyze the sulfoxidation of thioanisole. Structurally, vanadium chloroperoxidase resembles acid phosphatases and acts as a substitute for phosphate and binds to phosphoryl transfer enzyme, thus reaching the intermediate state of reaction and decreasing the phytase activity. Vanadate acts as an anion and binds with positively charged amino acids present in the active site of phytase. By interfering with the phosphoesterase activity, it changes the enzyme functionality and develops a vanadate-dependent haloperoxidase. Only HAP phytase behaves as haloperoxidase due to the similarity in their active site [50, 51]. Likewise, the semi-synthetic peroxidase behaviour of phytase was studied in *S. thermophile* [51], *P. anomala* [50], and *A. oryzae* [20]. The application of vanadate synthesizes peroxidase found in the formation of radiolabelled monoclonal antibodies, quantification of chloride amount in samples, diagnostic purposes, antimicrobial work, and immunological studies [50, 51].

Table 4 Estimation and reduction of phytic acid in the diet after the addition of phytase

Sample	Phytic content (mg/g feed)	Reduction in phytic acid level (%)
Diet I (Control)	50.43 ± 1.15	0
Diet II (Control + 500 FTU/kg phytase)	33.00 ± 0.13	34.6
Diet III (Control + 1000 FTU/kg phytase)	30.37 ± 0.45	39.8
Diet IV (Control + 1500 FTU/kg phytase)	26.91 ± 0.83	46.6

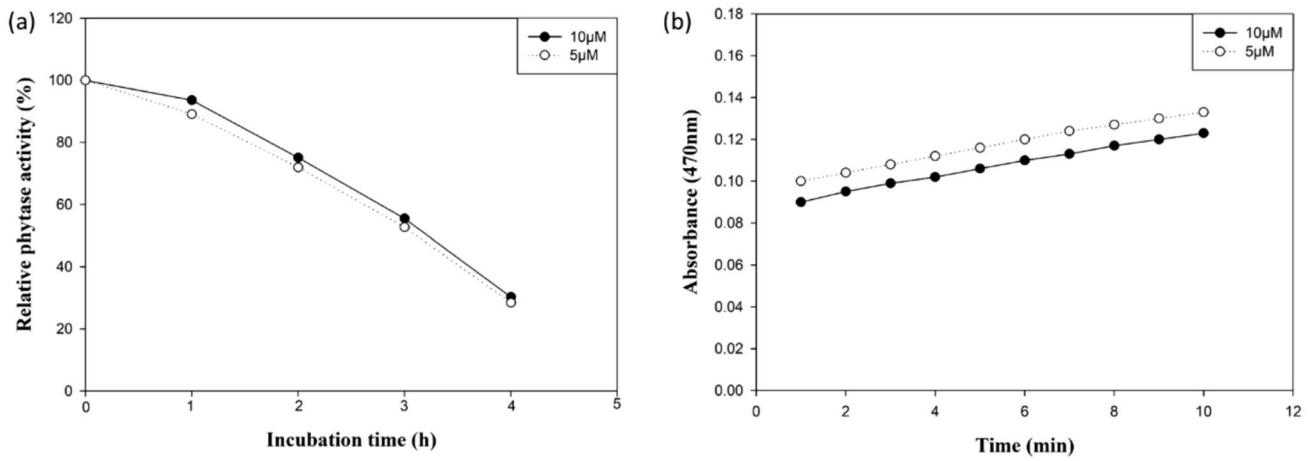


Fig. 9 Reduction of phytase activity **a** and increase in peroxidase activity **b** with time

Conclusion

Wheat bran is used for the production of *Penicillium oxalicum* phytase through the SSF method. After optimizing the conditions using the OVAT approach, phytase production was enhanced by 2.4-fold. The *P. oxalicum* PBG30 phytase exhibits thermostability, acidity tolerance, broad substrate specificity, and protease resistance. Its high thermostability protects the enzyme during the pelleting process in feed production. The acidic nature of phytase helps it to survive and function effectively in the gastrointestinal tract. Its resistance to proteases help in withstanding the action of digestive enzymes such as trypsin and pepsin, allowing it to efficiently degrade phytic acid. Due to these distinctive features, phytase can serve as an ideal additive in the food and feed industry. *P. oxalicum* PBG30 phytase showed the hydrolysis of metal-phytate and protein-phytate complexes and leading to the liberation of inorganic phosphate in a time-dependent manner. This indicates that *P. oxalicum* PBG30 phytase is effective in improving the nutritional content of food and feedstuffs. Additionally, the phytase plays a role in reducing phytate levels in fish feed, which helps decrease P levels in water bodies, thereby mitigating environmental pollution. The presence of haloperoxidase activity in vanadium-modified phytase suggests potential for clinical, immunological, and various biotechnological applications. However, further research is imperative to validate these possibilities.

Acknowledgements The author acknowledges the Department of Biotechnology, Delhi Technological University, Delhi, Department of Microbiology, Maharshi Dayanand University, Rohtak and Aqua Research Lab, Department of Zoology, University of Delhi, Delhi for supporting this research work.

Author's Contributions The original draft was prepared and edited by P and all the experiments were performed by her. Data were analyzed by P, BS, BG and JGS. The draft was critically revised and improved under the supervision of BS, BG, RKG and JGS. All authors read and approved the final manuscript.

Funding The author (Ms. Priya) gratefully acknowledges the financial assistance as Junior/Senior research fellowship from the University Grants Commission (UGC), New Delhi, India.

Declarations

Conflict of interests The authors declare no competing interests.

References

1. Singh B, Pragya, Tiwari SK, Singh D, Kumar S, Malik V (2024) Production of fungal phytases in solid state fermentation and potential biotechnological applications. World J Microbiol Biotechnol 40:22. https://doi.org/10.1007/s11274-023-03783-1
2. Rizwanuddin S, Kumar V, Naik B, Singh P, Mishra S, Rustagi S, Kumar V (2023) Microbial phytase: their sources, production, and role in the enhancement of nutritional aspects of food and feed additives. J Agric Food Res 100559. https://doi.org/10.1016/j.jafr.2023.100559
3. Singh B, Boukhris I, Pragya, Kumar V, Yadav AN, Farhat-Khemakhem A, Kumar A, Singh D, Blibech M, Chouayekh H, Alghamdi OA (2020) Contribution of microbial phytases to the improvement of plant growth and nutrition: a review. Pedosphere 30:295–313. https://doi.org/10.1016/S1002-0160(20)60010-8
4. Priya VI, Pragya, Goswami RK, Singh B, Sharma JG, Giri B (2023) Role of microbial phytases in improving fish health. Rev Aquac 15:1480–1500. https://doi.org/10.1111/raq.12790
5. Rizwanuddin S, Kumar V, Singh P, Naik B, Mishra S, Chauhan M, Saris PEJ, Verma A, Kumar V (2023) Insight into phytase-producing microorganisms for phytate solubilization and soil sustainability. Front microbiol 14:1127249. https://doi.org/10.3389/fmicb.2023.1127249
6. Jain J, Kumar A, Singh D, Singh B (2018) Purification and kinetics of a protease-resistant, neutral, and thermostable phytase from

Bacillus subtilis subsp. *subtilis* JJBS250 ameliorating food nutrition. Prep Biochem Biotechnol 48:718–724. https://doi.org/10.1080/10826068.2018.1487848
7. Menezes LHSD, Oliveira PC, do Espírito Santo EL, Gonçalves MS, Bilal M, Ruiz HA, da Silva EGP, Salay LC, de Oliveira JR, Franco M, (2024) Solid-state fermentation as a green technology for biomass valorization: optimization techniques for bioprocess—an overview. Bioenergy Res 17:42–58. https://doi.org/10.1007/s12155-023-10670-y
8. Jain J, Sapna SB (2016) Characteristics and biotechnological applications of bacterial phytases. Process Biochem 51:159–169. https://doi.org/10.1016/j.procbio.2015.12.004
9. Sapna SB (2017) Purification and characterization of a protease-resistant phytase of *Aspergillus oryzae* SBS50 whose properties make it exceptionally useful as a feed supplement. Int J Biol Macromol 103:458–466. https://doi.org/10.1016/j.ijbiomac.2017.05.077
10. Neira-Vielma AA, Aguilar CN, Ilyina A, Contreras-Esquivel JC, das Graça Carneiro-da-Cunha M, Michelena-Álvarez G, Martínez-Hernández JL, (2018) Purification and biochemical characterization of an *Aspergillus niger* phytase produced by solid-state fermentation using triticale residues as substrate. Biotechnol Rep 17:49–54. https://doi.org/10.1016/j.btre.2017.12.004
11. Sanni DM, Lawal OT, Enujiugha VN (2019) Purification and characterization of phytase from *Aspergillus fumigatus* Isolated from African Giant Snail (*Achatina fulica*). Biocatal Agric Biotechnol 17:225–232. https://doi.org/10.1016/j.bcab.2018.11.017
12. Saxena A, Verma M, Singh B, Sangwan P, Yadav AN, Dhaliwal HS, Kumar V (2020) Characteristics of an acidic phytase from *Aspergillus aculeatus* APF1 for dephytinization of biofortified wheat genotypes. Appl Biochem Biotechnol 191:679–694. https://doi.org/10.1007/s12010-019-03205-9
13. Caliskan-Özdemir S, Önal S, Uzel A (2021) Partial Purification and Characterization of a Thermostable Phytase Produced by Thermotolerant *Aspergillus tubingensis* TEM 37 Isolated from Hot Spring Soil in Gediz Geothermal Field. Turkey Geomicrobiol J 38:895–904. https://doi.org/10.1080/01490451.2021.1977432
14. Ahmed NE, Salem SS, Hashem AH (2022) Statistical optimization, partial purification, and characterization of phytase produced from *Talaromyces purpureogenus* NSA20 using potato peel waste and its application in dyes de-colorization. Biointerface Res Appl Chem 12:4417–4431. https://doi.org/10.33263/briac124.441744431
15. Bhandari Y, Sonwane B, Vamkudoth KR (2023) Isolation and biochemical characterization of acid phytase from *Aspergillus niger* and Its applications in dephytinization of phytic acid in poultry feed ingredients. Microbiology 92:221–229. https://doi.org/10.1134/S0026261722601452
16. Kumari N, Bansal S (2021) Production and characterization of a novel, thermotolerant fungal phytase from agro-industrial byproducts for cattle feed. Biotechnol Lett 43:865–879. https://doi.org/10.1007/s10529-020-03069-8
17. Thakur N, Patel SK, Kumar P, Singh A, Devi N, Sandeep K, Pandey K, Chand D (2022) Bioprocess for hyperactive thermotolerant *Aspergillus fumigatus* phytase and its application in dephytinization of wheat flour. Catal Lett 152:3220–3232. https://doi.org/10.1007/s10562-021-03886-0
18. Priya, Singh B, Sharma JG, Giri B (2023) Optimization of phytase production by *Penicillium oxalicum* in solid-state fermentation for potential as a feed additive. Prep Biochem Biotech, pp 1–11. https://doi.org/10.1080/10826068.2023.2297688
19. Sharma A, Ahluwalia O, Tripathi AD, Singh G, Arya SK (2020) Phytases and their pharmaceutical applications: mini-review. Biocatal Agric Biotechnol 23:101439. https://doi.org/10.1016/j.bcab.2019.101439
20. Pragya, Sharma KK, Singh B (2023) Phytase from *Aspergillus oryzae* SBS50: Biocatalytic reduction of anti-nutritional factor and exhibiting vanadium-dependent haloperoxidase activity. Biocatal Agric Biotechnol 52:102840. https://doi.org/10.1016/j.bcab.2023.102840
21. Sapna SB (2015) Biocatalytic potential of protease-resistant phytase of *Aspergillus oryzae* SBS50 in ameliorating food nutrition. Biocatal and Biotransfor 33:167–174. https://doi.org/10.3109/10242422.2015.1076215
22. Mahmood S, Shahid MG, Nadeem M (2021) Screening of Phytate Degrading fungi and optimization of culture conditions for phytase synthesis using agro-industrial by-products. Pak J Bot 53:763–770. https://doi.org/10.30848/PJB2021-2(12)
23. Fiske CH, Subbarow Y (1925) The colorimetric determination of phosphorus. J Biol Chem 66:375–400. https://doi.org/10.1016/S0021-9258(18)84756-1
24. Pragya, Sharma KK, Kumar S, Manisha, Singh D, Kumar V, Singh B (2023) Enhanced production and immobilization of phytase from *Aspergillus oryzae*: a safe and ideal food supplement for improving nutrition. Lett Appl Microbiol 76. https://doi.org/10.1093/lambio/ovac077
25. Sapna SB (2013) Improved production of protease-resistant phytase by *Aspergillus oryzae* and its applicability in the hydrolysis of insoluble phytates. J Ind Microbiol Biotechnol 40:891–899. https://doi.org/10.1007/s10295-013-1277-3
26. Arabacı N, Karaytuğ T (2023) Alkaline thermo-and oxidant-stable protease from *Bacillus pumilus* Strain TNP93: laundry detergent formulations. Indian J Microbiol 63:575–587. https://doi.org/10.1007/s12088-023-01115-3
27. Gao Y, Shang C, Maroof MS, Biyashev RM, Grabau EA, Kwanyuen P, Burton JW, Buss GR (2007) A modified colorimetric method for phytic acid analysis in soybean. Crop Sci 47:1797–1803
28. Gaind S, Singh S (2015) Production, purification and characterization of neutral phytase from thermotolerant *Aspergillus flavus* ITCC 6720. Int Biodeterior Biodegr 99:15–22. https://doi.org/10.1016/j.ibiod.2014.12.013
29. Arya M, Chauhan G, Fatima T, Verma D, Sharma M (2024) Statistical modelling of thermostable cellulase production conditions of thermophilic *Geobacillus* sp. TP-1 isolated from tapovan hot springs of the Garhwal Himalayan Mountain Ranges. India Indian J Microbiol 64:1132–1143. https://doi.org/10.1007/s12088-024-01258-x
30. Jatuwong K, Kumla J, Suwannarach N, Matsui K, Lumyong S (2020) Bioprocessing of agricultural residues as substrates and optimal conditions for phytase production of chestnut mushroom, *Pholiota adiposa*. Solid State Fermentation J Fungi 6:384. https://doi.org/10.3390/jof6040384
31. Mahendran S, Sankaralingam S, Maheswari P, Dhivya RR, Kathiresan D, Karthikeyan S, Ramya SS, Seethapathy P, Harinathan B, Palpperumal S (2022) Production, characterization, and feed supplement applications of phytase enzyme from *Aspergillus tubingensis* Isolated from Western Ghats Soil. Biomass Convers. Biorefinery 1–11. https://doi.org/10.1007/s13399-022–02894–3
32. Suresh S, Radha KV (2015) Effect of a mixed substrate on phytase production by *Rhizopus oligosporus* MTCC 556 using solid state fermentation and determination of dephytinization activities in food grains. Food Sci Biotechnol 24:551–559. https://doi.org/10.1007/s10068-015-0072-5
33. Chanderman A, Puri AK, Permaul K, Singh S (2016) Production, characteristics and applications of phytase from a rhizosphere isolated Enterobacter sp. ACSS Bioprocess Biosyst Eng 39:1577–1587. https://doi.org/10.1007/s00449-016-1632-7
34. Qasim SS, Shakir KA, Al-Shaibani AB, Walsh MK (2017) Optimization of culture conditions to produce phytase from

Aspergillus tubingensis SKA. Food Nutr Sci 08:733–745. https://doi.org/10.4236/fns.2017.87052
35. Elkhateeb YAM, Fadel M (2022) Bioinformatic studies, experimental validation of phytase production and optimization of fermentation conditions for enhancing phytase enzyme production by different microorganisms under solid-state fermentation. Open Microbiol J 16:1–11. https://doi.org/10.2174/18742858-v16-e2202160
36. Gunashree BS, Venkateswaran G (2015) Extracellular phytase from *Aspergillus niger* CFR 335: purification and characterization. J Food Sci Technol 52:4558–4564. https://doi.org/10.1007/s13197-014-1304-z
37. Ajith S, Ghosh J, Shet D, ShreeVidhya S, Punith BD, Elangovan AV (2019) Partial purification and characterization of phytase from *Aspergillus foetidus* MTCC 11682. AMB Expr 9:1–11. https://doi.org/10.1186/s13568-018-0725-x
38. Kaur R, Saxena A, Sangwan P, Yadav AN, Kumar V, Dhaliwal HS (2017) Production and characterization of a neutral phytase of *Penicillium oxalicum* EUFR-3 isolated from Himalayan region. Nus Biosci 9:68–76. https://doi.org/10.13057/nusbiosci/n090112
39. Suresh S, Radha KV (2016) Statistical optimization and mutagenesis for high level of phytase production by *Rhizopus oligosporus* MTCC 556 under solid state fermentation. J Environ Biol 37:253
40. Gampa M, Nagar S, Kumari K, Tanwar E, Goyal S, Kumar V, Singh B (2023) Cyclic extraction of phosphate from soybean meal using immobilized *Aspergillus oryzae* SBS50 phytase. Bioprocess Biosyst Eng 47:3955. https://doi.org/10.1007/s00449-023-02943-9
41. Singh N, Kumari A, Gakhar SK, Singh B (2015) Enhanced cost-effective phytase production by *Aspergillus niger* and its applicability in dephytinization of food ingredients. Microbiology 84:219–226. https://doi.org/10.1134/S0026261715020149
42. Kim BH, Lee JY, Lee PC (2015) Purification, sequencing and evaluation of a divergent phytase from *Penicillium oxalicum* KCTC6440. J Gen Appl Microbiol 61:117–123. https://doi.org/10.2323/jgam.61.117
43. Lee SH, Cho J, Bok J, Kang S, Choi Y, Lee PCW (2015) Characterization, gene cloning, and sequencing of a fungal phytase, PhyA, from *Penicillium oxalicum* PJ3. Prep Biochem Biotechnol 45:336–347. https://doi.org/10.1080/10826068.2014.923446
44. Yu P, Wang XT, Liu JW (2015) Purification and characterization of a novel cold-adapted phytase from *Rhodotorula mucilaginosa* strain JMUY14 isolated from Antarctic. J Basic Microbiol 55:1029–1039. https://doi.org/10.1002/jobm.201400865
45. Monteiro PS, Guimarães VM, Melo RRD, Rezende STD (2015) Isolation of a thermostable acid phytase from *Aspergillus niger* UFV-1 with strong proteolysis resistance. Braz J Microbiol 46:251–260. https://doi.org/10.1590/S1517-838220120037
46. Kumari N (2022) Production, immobilization and characterization of fungal phytase and its utilization in food and feed industry. Doctoral thesis
47. Sun M, He Z, Jaisi DP (2021) Role of metal complexation on the solubility and enzymatic hydrolysis of phytate. PLoS ONE 16:e0255787. https://doi.org/10.1371/journal.pone.0255787
48. Roopashri AN, Varadaraj MC (2015) Functionality of phytase of *Saccharomyces cerevisiae* MTCC 5421 to lower inherent phytate in selected cereal flours and wheat/pearl millet-based fermented foods with selected probiotic attribute. Food Biotechnol 29:131–155. https://doi.org/10.1080/08905436.2015.1026602
49. Ranjan B, Singh B, Satyanarayana T (2015) Characteristics of recombinant phytase (rSt-Phy) of the thermophilic mold *Sporotrichum thermophile* and its applicability in dephytinizing foods. Appl Biochem Biotechnol 177:1753–1766. https://doi.org/10.1007/s12010-015-1851-4
50. Joshi S, Satyanarayana T (2015) Bioprocess for efficient production of recombinant *Pichia anomala* phytase and its applicability in dephytinizing chick feed and whole wheat flat Indian breads. J Ind Microbiol Biotechnol 42:1389–1400. https://doi.org/10.1007/s10295-015-1670-1
51. Singh B, Sharma KK, Kumari A, Kumar A, Gakhar SK (2018) Molecular modeling and docking of recombinant HAP-phytase of a thermophilic mould *Sporotrichum thermophile* reveals insights into molecular catalysis and biochemical properties. Int J Biol Macromol 115:501–508. https://doi.org/10.1016/j.ijbiomac.2018.04.086

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

JOURNAL OF THERMAL SCIENCE AND ENGINEERING APPLICATIONS

Proposal and evaluation of a novel solar thermal driven cooling-power cogeneration system

Abdul Khaliq
Mechanical Engineering Department, College of Engineering at Yanbu, Taibah University, Yanbu Al-Bahr 41911, Saudi Arabia, Email: khaliqsb@gmail.com

Rajesh Kumar*
Department of Mechanical Engineering, Delhi Technological University (DTU), Bawana Road, Shahbad Daulatpur, Delhi-110042, India, Email: dr.rajeshmits@gmail.com

Tawfiq Al-Mughanam
Department of Mechanical Engineering, College of Engineering, King Faisal University, P.O. Box 380, Al- Ahsa 31982, Saudi Arabia
Email: talmughanam@kfu.edu.sa

Khaled Zeid
Mechanical Engineering Department, College of Engineering at Yanbu, Taibah University, Yanbu Al Bahr 41911, Saudi Arabia, Email: kmzied@gmail.com

Abstract

This study establishes a novel cooling-power cogeneration system which utilizes Parabolic Trough Solar Collectors (PTSC) employing Air, Carbon dioxide (CO_2), and Helium (He) as heat carrying fluids. This heat is utilized by Organic Flash Cycle (OFC) connected to two-phase ejector which simultaneously generates electricity, and cooling for refrigeration and air conditioning. The system proposed provides a promising solution that can help in reducing greenhouse gas emissions and increasing overall efficiency and energy savings. Modelling & simulation through Engineering Equation Solver (EES) is performed to investigate the effect of solar flux, and type of solar heat transfer fluid (SHTF) on exit temperature of SHTF. Promotion of solar flux increases the temperature of SHTF which is found highest for He and lowest for CO_2. A parametric analysis is done to find out the outcome of design parameters. Cogeneration cycle connected to He operated PTSC performs good relative to Air and CO_2 as SHTF. Exergy destruction is found to be 53.87 %, 22.09 %, and 6.15 % in PTSC, OFC, and in two-phase ejector respectively while production of power, exergetic refrigeration, & exergetic air-conditioning as 4.02 %, 10.65 %, 3.22 % respectively.

Keywords: Parabolic trough solar collectors (PTSC), Organic flash cycle (OFC), two-phase ejector, cogeneration, exergy

*Corresponding Author: Email: dr.rajeshmits@gmail.com (Rajesh Kumar)

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

1. Introduction

Increased level of living standard of people and the rapid industrial development led to an exponential increase in demand of energy in the last few decades. Most of this increased energy demand is met by fossil fuel combustion in a single energy generation unit that leads to problematic issues like the fast depletion of fossil fuel reserves, increased level of pollutant emissions, and global warming. To combat these issues, it is required to find out the clean and renewable energy source and development of potential and sustainable combined energy system [1, 2]. Among the renewable sources of energy, solar energy is a universal energy source and has the potential to vanquish all the environmental threats, and has the capability to produce electric power and heat for diverse requirements of domestic and industrial sectors' energy in a sustainable fashion. Kalogirou [3] discussed the different styles of solar panels & collectors and its various applications. Al Mughanam and Khaliq [4] introduced a novel cycle to generate the electricity and cooling, simultaneously, and run by concentrating collectors. They have presented a parametric study of the cycle in view of its exergetic and energetic perspectives.

Concentrating solar collectors (dish collectors, Heliostat field, and PTSCs) are very popular devices worldwide due to their conversion of solar radiations to heat. PTSCs considered as the most promising systems as it supplies energy with enhanced stability and reliability in cost-effective manner in regions & sites receiving high direct normal irradiations (DNI) [5, 6]. However, stand-alone solar operated electricity generation systems are presently not very popular in comparison with traditional fossil fuel driven plants due to its more initial investment, lesser efficiency, discontinuous power supply, etc. So, efficient use of solar energy at peak summer for cooling and power generation is close coincidence of requirement of extreme cooling load when rich solar radiations are available [7, 8]. Much work has been done on the integrated solar heat

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

driven Organic Rankine Cycle (ORC) and Absorption Refrigeration Cycle (ARC) because ORC presents the benefits of operation flexibility and desirable efficiency [9]. Exergy analysis is performed for the solar operated combined system of ARC with the ORC. Result points the components having significant exergy destruction which effecting the cycle function [10]. The investigation of combined ORC (for power), LiBr-H_2O based ARC (for refrigeration), and process heater (for heating), respectively is done and results shows the advantage of using the ARC on the outcome of cycle [11]. The simulation of evacuated tube collector based NH_3-H_2O absorption refrigeration system is done using hourly solar radiation data. The outcome shows the importance of variation of air temperature on the output of system [12]. The exergy analysis of solar tower operated Steam Rankine Cycle (SRC) combined with ARC for electric and refrigeration output. A great improvement in system outcome is noticed with the increase in the DNI [13]. LiBr-H_2O & NH_3-H_2O are very popular working fluid pairs used in ARC for different range of cooling temperatures. When the evaporator temperature requirement is less than 0^oC, LiBr-H_2O based ARC will not work. NH_3- H_2O based ARC is widely used for refrigerating the thermal load of below 0^oC [14]. However, in NH_3-H_2O based ARC, a rectifier (energy consuming device) is required, which degrade the outcome of the system.

In view of the constraints imposed by ARC, a further new alternative heat operated sustainable cooling cycle is required. Ejector Refrigeration Cycle (ERC) now a days seems to be very fascinating cooling choice owing to high energy efficiency, low cost, silent operation, and flexibility in refrigerants use [15, 16]. Therefore, to address the purpose of utilizing the low-grade solar energy source more effectively, ERC & ORC are adopted to utilize waste heat. The exergetic evaluation of a cooling-power system is done wherein an ERC was adopted to augment the thermodynamic performance of ORC [17]. The approach of efficient employment of solar heat by means of ejector in ORC for refrigeration & power is continued by Khaliq [18]. Zhang and Mohamed [19] inspect the response of properties of hydrocarbons refrigerants in solar operated

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

ORC integrated to ERC in view of thermodynamic outcome. Elakhdar et al. [20] hybridized the PTSC, to meet the building energy demand, with a thermal system consists of ORC & ERC which uses Thermal oil (as heat transfer fluid) and the results showed hourly response of the model. Mahdavi et al. [21] investigated a new type of power, cooling, and heating cogeneration system using Therminol VP1 operated PTSC as the primary energy source to drive the dual-pressure ORC combined with ERC. The proposed configuration was analyzed and the system efficiencies found to be 20.26% (energy efficiency) and 15.81% (exergy efficiency), respectively. Yadav et al. [22] showed the exergy analysis of a new configuration combining the ORC with a secondary ejector & an internal heat exchanger and PTSC using Therminol oil as solar fluid. They determined the second law efficiency of the proposed configuration as 18% using isobutane as the working fluid which claimed to be higher than those of the systems employing an ejector between the ORC turbine and condenser.

In the Combined Cooling and Power (CCP) cycles described above, majority of the solar heat was aimed to be utilized for power generation, but most of the works focus on traditional ORC and an ERC to exploit the solar heat to produce electricity and cooling. Power generation with ORC has a disadvantage of temperature mismatching between the thermal energy source stream and the working fluid stream of ORC, resulting in the increase of heat transfer irreversibilities leads to performance reduction. Employment of ERC with the use of single-phase pure leads to lower thermodynamic performance and provide cooling at one level of temperature either for refrigeration or air conditioning. To overcome with such limitations, Organic Flash Cycle (OFC) is considered which uses a separator tank and a flasher and is gaining wide acceptance owing to provide significant reduction in exergy destruction and increased turbine output [23, 24]. Few studies conducted the analysis of OFC from the viewpoints of thermodynamics to utilize the solar heat for power generation. The outcome of OFC & ORC is compared with due focus on thermodynamic behavior of working fluids and heat source temperature is done by Lee et al. [25].

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

They favored OFC because it shows better performance for harnessing the low temperature heat sources. Varma et al. [26] proposed and conducted the thermodynamic investigations of an OFC considering the partial flashing of preheated water and mixing of separated vapor with turbine streams for power enhancement. Their investigation was limited to comparison of the performance of OFC with ORC and did not perform the parametric analysis of the cycle from energy and exergy perspectives which is desired to diagnose the real potential of OFC.

Referring to CCP cycles, very few literatures focus on the integration of OFC with ERC driven by low temperature heat sources. Moreover, employing single-phase ejector driven by exhaust heat of ORC turbine leads to a reduction in turbine power owing to maintain sufficient heat to vaporize the ERC working fluid, further, the cycle has the limitation of providing cooling in single mode either for refrigeration or air conditioning. In this regard, application of two-phase ejector refrigeration cycle which provide cooling in two different temperatures range and can fulfil the requirement of refrigeration and air conditioning, simultaneously, which makes full use of excess power generated by solar energy driven OFC can be considered as an enthralling method in view of improving its energy outcome and conserving the turbine power output. Various investigations are conducted to analyze the outcome of two-phase ejector cooling cycles. A vapor compression refrigeration system is analyzed using different working fluids with two-phase ejector by Kornhauser [27]. Chaiwongsa and Wongwises [28] analyzed the consequences of varying the exit diameter of nozzle & throat for the ERC consist of single evaporator and a refrigerant separator. Pottker et al. [29] analyzed the experimental model for the application of R410A refrigeration in ejector cooling system without considering the ejector dimensions and found a considerable increase in the Coefficient of Performance (COP) of refrigeration system. Lawrence [30] conducted a theoretical & experimental examination of a two-phase ejector refrigeration cycle using the refrigerants working on low pressures. Lawrence and Elbel [31] analyzed the outcome of using the R143a & R1234yf as refrigerants in ERC and an increase of 5% and 6% of COP was

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

observed. For different component efficiencies, the outcome of two-phase ejector was examined by Wang and Yu [32]. They suggest a new method of analysis of ejector by combining its model with reported data.

Above reported investigations clearly indicates that majority of the solar collectors employs to drive the cooling-power cogeneration systems uses water, thermal oils, and molten salts as solar heat transfer fluid. Utilization of these fluids as SHTF brings some challenges such as water is appropriate only for performing the temperature limit up to 100 ℃, and for higher temperature limits, water is pressurized. Thermal oils are being used in many industrial applications up to 400 ℃ and molten salts can be used up to 600 ℃ (solidification at lower temperature occurs) [33 -35]. To mitigate the risks and limitations associated with the use of water, oil, and molten salt, employment of gases as the solar heat transfer medium in the PTSC is being used to operate the systems at high temperature level. Few illustrations can be noticed in the published work where PTSCs using gases as solar heat transfer medium have been successfully utilized for refrigeration, and power-cooling cogeneration [36, 37].

It is observed from the literature review that the investigation of OFC operated cooling-power cycle run by PTSC is very limited and further investigation on this area is required. In this regard, this study introduces an innovative integrated system for combined generation of electricity, and cooling at two different levels of temperature to satisfy the needs of refrigeration and air conditioning. This configuration relies on a unique integration that harnesses high intensity solar radiation of a region and recovers waste heat from the turbine which effectively producing electricity and cooling in this novel setup. A detailed energetic-exergetic evaluation is conducted to diagnose the real potential of proposed cogeneration system. Another contribution of this research is the assessment of the impact of climatic conditions on the rate of cleaner production of cooling and electric power and on the exergetic efficiency of the cogeneration. This detailed

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

analysis and exploration of climatic condition effects on the system's functionality and outcome have not been extensively investigated in past studies. Therefore, this investigation fills a crucial gap of the reported literature by revealing new insights and findings in these areas, thus advancing the knowledge, and understanding of integrated systems.

The main novelties of the present work are as follows:

- To develop a new type of cogeneration system utilizing gaseous working fluids in concentrating solar collectors to operate in high temperature levels for the co-production of electricity and cooling at two different level of temperatures.
- Conduct a comprehensive steady-state energy and exergy analysis of the developed cogeneration system using EES (Engineering Equation Solver) [38].
- Depict the exergy distribution of incident radiations into the various exergetic outcomes.

2. System description

Fig. 1 is the proposed system that shows an innovative combination of OFC using R124 as the working fluid to generate power, along with a bottoming cycle that incorporates a two-phase ejector to simultaneously provide cooling for air-conditioning and refrigeration. PTSCs are chosen as the best solar devices for capturing and utilizing solar radiation. Air, CO_2 and He are used as heat transfer fluids in the PTSC field to transfer heat. The high-temperature heat transfer fluid (state 18) enters the heat recovery vapor generator (HRVG) of the OFC, and the OFC absorbs the thermal energy from the SHTF. High temperature SHTF passes through HRVG causes the increase of fluid temperature of OFC & consequently forms superheated vapor (22). The SHTF exits the HRVG (state point 16) and enters (state point 17) the PTSC field through a solar pump1. One part of pressurized hot liquid (30) from HRVG is given to high pressure flasher (HPF) unit (33) via expansion valve1 (EV1). In this cycle HRVG consists of three zones namely, economizer (ECO), evaporator (EVA) and superheater (SH). In High Pressure Flasher (HPF), this liquid is flashed to

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

saturated vapor (35) which is supplied to the turbine and into liquid (34) connected to Low Pressure Flasher (LPF) via EV2 wherein it flashed and separated into liquid (36) and saturated vapor (37). This arrangement caused two extra streams of fluid into the turbine is, one from HPF and the other from the LPF and hence generate a greater amount of power which is enough to drive the compressor and pumps, and to meet the electricity requirement of a building. The vapors (23) coming from the turbine rejects the heat at regenerator (23−24) and then condensed (25). The condensate is pumped (via pump2) and then passes through regenerator (27) and mixed with liquid (36) is then pumped (via pump3) to HRVG (29) to repeat the cycle. To achieve the desired cooling, the two-phase ejector cycle employs R600a as working fluid. The vapor from the Separator (10) via heat exchanger enters the Compressor (C) (11) and then the condenser2 (Cond2) (1) to reject the heat into ambient. The saturated liquid leaving the Cond2 (2) goes to the air- conditioner (AC) via heat exchanger (HX) and Expansion Valve3 (EV3). This heat exchanger (HX) exchanges the heat between the liquid leaving the Cond2 and the saturated vapor approaching the Compressor. The stream from the air-conditioner (AC) (5) and enters to the Ejector (Eje) as primary fluid of motive nozzle. The high-velocity vapor creates the low-pressure region at the mixing chamber inlet and sucks the vapor from Refrigerator (ref) (9). A normal shock wave occurs as mixed fluid enters the in constant area part and causes a pressure increase and the decelerates in the diffuser as shock takes place and then enters the Separator (Sep) (6). The separator (Sep) separates the fluid into two parts namely, liquid (7) goes to the Refrigerator (ref) that produces refrigeration effect (8) and the saturated vapor (10) is routed to Compressor (C) (11) via heat exchanger (HX). Table 1 depicts some relevant properties of SHTF.

3. Computational Method

Examination of the cogeneration model is conducted in four stages:

- PTSC is working with gaseous SHTF that drives the proposed cogeneration model.

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

- OFC using R124 as the working fluid to generate power
- A two-phase ejector cycle is used to simultaneously provide cooling for air-conditioning and refrigeration
- Energy and exergy examination of the model is conducted.

The modelling is done on EES using three gaseous fluids is made. The properties of the working fluids R124 (working fluid for OFC), and R600a (working fluid for the two-phase ejector) basic were computed by employing the REFPROP 6.01 [39]. Thermal analyses of the proposed model is executed using the EES.

Global Warming Potential, or GWP, is a measure of how destructive a climate pollutant is. Refrigerants today are often thousands of times more polluting than carbon dioxide (CO_2). The GWP of a gas refers to the total contribution to global warming resulting from the emission of one unit of that gas relative to one unit of the reference gas, CO_2, which is assigned a value of 1. GWPs can also be used to define the impact greenhouse gases will have on global warming over different time periods or time horizons. These are usually 20 years, 100 years, and 500 years. The refrigerants considered here in this investigation are R124 and R600a (Hydrocarbon) and their employment in thermodynamic cycles for power generation and cooling offers very less GWP compared to commonly used refrigerants. Therefore, their use provides a global push towards low Global Warming potential alternatives.

Assumptions are made regarding the proposed cogeneration cycle [20, 25, 31, 36]:

- Flow is assumed to be steady, frictionless & 1-dimensional in components of proposed cogeneration model.
- Heat loss from the components, kinetic and potential energy change of stream of proposed cogeneration model assumed negligible.
- Expansion valves follow isenthalpic process.

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

- Pressure drops in components assumed negligible.
- Pumps and expander have fixed Isentropic efficiency.
- Only physical exergy is considered for calculation purpose.
- The ambient Pressure P_0 is 101.325 kPa and Temperature T_0 is 25 ℃.

3.1. Energy and exergy analyses

For the proposed cogeneration model, an exergetic investigation is done, which reports the amount of exergy destruction in the components of the model.

Conservation of mass principle expresses the steady state operation as:

$$\Sigma \dot{m}_i - \Sigma \dot{m}_o = 0 \quad (1)$$

Conservation of energy principle expresses the steady state operation as:

$$\sum \dot{Q} - \dot{W} = \dot{m}\left(h_0 + \frac{C^2}{2} + gz_0\right) - \dot{m}\left(h_i + \frac{C^2}{2} + gz_i\right) \quad (2)$$

where h, c and z are the specific enthalpy, bulk velocity and elevation of flowing stream.

Table 2 shows the Energy balance for the components of the of the model are obtained using Eqs. (1) and (2).

The specific exergy of flowing stream while neglecting the kinetic and potential exergies is formulated as:

$$e_x = (h - h^*) - T_0(s - s^*) \quad (3)$$

where * represents the property at datum state.

For a control volume working on steady state, Exergy destruction equation is given as:

$$\Sigma\left(1 - \frac{T_0}{T}\right)\dot{Q}_j - \Sigma\dot{W} + \Sigma\dot{m}_i e_{x_i} - \Sigma\dot{m}_o e_{x_o} - \dot{E}_{x,D} = 0 \quad (4)$$

Table 3 shows the Exergy destruction expression for the components of the of the model.

3.2. Proposed cogeneration model's overall performance indicators

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

The PTSC system comprises absorber tube which is fixed along the focal length of parabolic concentrator and a reflective cover is employed to decrease the heat loss of absorber tube [20]. The gaseous fluid absorb heat in the collector and it passed via central tube of PTSC. Part of this heat is transferred via HRVG to OFC to generate electricity for the grid and to drive the compressor of ejector cooling cycle. Incident rate of solar energy that PTSC receives ($\dot{Q}_{solar}$) is given by [36]:

$$\dot{Q}_{Solar} = G_b A_p \tag{5}$$

$$\dot{Q}_u = \dot{m}_{SHTF}(h_{18} - h_{17}) = \eta_{en,PTSC}\ G_b A_p \tag{6}$$

Exergy destruction rate of process of solar heat absorption from sun to receiver (s-r) is determined as:

$$\dot{E}_{x,D,s-r} = \eta_{opt}\ \dot{E}_{x,Solar} - \dot{Q}_{abs}\left(1 - \frac{T_0}{T_r}\right) \tag{7}$$

where $\dot{E}_{x,Solar}$ and $\dot{Q}_{abs}$ are defined by the equations below:

$$\dot{E}_{x,Solar} = G_b A_p \left\{1 - \frac{4}{3}\frac{T_0}{T_{solar}} + \frac{1}{3}\left(\frac{T_0}{T_{solar}}\right)^4\right\} \tag{8}$$

$$\dot{Q}_{abs} = \eta_{opt}\ \dot{Q}_{Solar} \tag{9}$$

Exergy destruction rate of the transfer of heat from receiver to fluid (r-f) is expressed as:

$$\dot{E}_{x,D,r-f} = \dot{Q}_u\left(1 - \frac{T_0}{T_r}\right) - \dot{E}_{x,u} \tag{10}$$

$$\dot{E}_{x,u} = \dot{Q}_u - \dot{m}\ c_p.T_0\ \ln\left[\frac{T_{out}}{T_{in}}\right] - T_0\dot{m}\frac{\Delta P}{\rho_{fm}T_{fm}} \tag{11}$$

The term $\dot{Q}_{Solar}$ and $\dot{Q}_u$ are defined above by equations (1) and (2), respectively.

Rate of exergy destroyed in the solar collector is given by:

$$\dot{E}_{x,D,col} = \dot{E}_{x,D,s-r} + \dot{E}_{x,D,r-f} \tag{12}$$

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

Outcome of the proposed model is assessed via thermal and exergy efficiencies, which are determined through the first and second laws of thermodynamics. The thermal efficiency of the proposed model can be expressed as [33]:

$$\eta_{th} = \frac{\dot{W}_{net}+\dot{Q}_{ref}+\dot{Q}_{AC}}{\dot{Q}_{solar}} \quad (13)$$

where, $\dot{W}_{net}$ represents the net power output, $\dot{Q}_{ref}$ denotes the refrigeration capacity and $\dot{Q}_{AC}$ denotes the air-conditioning capacity, $\dot{Q}_{solar}$ is the total solar energy input to the system.

The exergy efficiency of the proposed system can be expressed as follows [17]:

$$\eta_{ex} = \frac{\dot{W}_{net}+\dot{E}_{ref}+\dot{E}_{AC}}{\dot{E}_{solar}} \quad (14)$$

where, $\dot{E}_{ref}$ and $\dot{E}_{AC}$ are exergy accompanied $\dot{Q}_{ref}$ and $\dot{Q}_{AC}$ respectively, and are given by:

$$\dot{E}_{ref} = \dot{Q}_{ref}\left(\frac{T_0}{T_{ref}} - 1\right) \quad (15)$$

$$\dot{E}_{AC} = \dot{Q}_{AC}\left(\frac{T_0}{T_{AC}} - 1\right) \quad (16)$$

where, T_0, T_{ref}, T_{AC} are the ambient temperature, refrigeration temperature, and air-conditioning temperature, respectively.

Mathematical equations formulated for PTSC considering heat transfer details to analyses the SHTF outlet temperature of is provided in the Ref. [34].

Proposed system takes into account the exergy of solar incident collected and divided into exergy output, exergy destructed, and exergy lost. The application of global exergy balance shows:

$$\dot{E}_{x,F,k} = \dot{E}_{x,P,k} + \dot{E}_{x,D,k} + \dot{E}_{x,L,k} \quad (17)$$

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

where, $\dot{E}_{x,F,k}$ = fuel or input exergy; $\dot{E}_{x,P,k}$ = product or output exergy; $\dot{E}_{x,L,k}$ = exergy loss; $\dot{E}_{x,D,k}$= exergy destruction.

For the proposed model, fuel (F) is the incident solar radiation. The distribution of incident exergy into exergy output, exergy destruction, and exergy lost has been presented for better understanding of the proposed model [36].

The input parameters taken for calculation purpose of the model are shown in Table 4.

The value of solar direct beam irradiation, temperature of Sun, and the ambient temperature are taken as 800 W/m^2, 5770 K, and 293K, respectively, because these are the most acceptable values of solar irradiance, Sun temperature, and ambient temperature in the Middle East region. The concentration ratio 26.37, the aperture area 69.6 m^2, the parabola width 5.8 m, and an optical efficiency 80% were taken because these have been considered as the satisfying values for the commercial parabolic trough collector, Eutrotrough ET150. This collector is a representative example of PTSCs as it has been used in many applications worldwide.

4. Results and discussion

4.1 Model validation

In order to ensure the accuracy of the calculations, the model presented in this paper was first verified. Due to the lack of a system of the same form for comparison, the OFC and the ejector sub-systems were verified separately. Table 5 shows a comparison between the present model and the literature results for the OFC [26]. The model for the OFC performance is validated by Verma and Srinivas [26]. The OFC turbine output and its exergy efficiency are determined as 1292 kW and 26.84%, respectively compared to OFC turbine output 1248 kW and exergy efficiency 26.21% obtained in [26]. The performance of the two-phase ejector is verified through the results obtained by Lawrence and Elbel [31]. The cooling outputs (refrigeration and air conditioning capacities) of

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

the present system are determined as 165 kW and 515 kW, respectively, compared to 1558 kW and 504 kW of the results obtained by [31]. The cooling output of the presently developed system is marginally higher than that of the [31] as PTSC using gaseous heat transfer fluids entails a major input to the system. A small deviation is found between the present findings and the results from the literature [26, 31] and in sum, the data obtained from present investigation make a solid accept to conduct future research in the area for the exploitation of solar thermal potential of the hot region to enhance the sustainable production of cooling and electric power. The relative errors of energy efficiency are within 3%, while those of exergy efficiency are within 1%.

Since, proposed system performance depends on the climatic conditions, therefore, the influence of variation of average solar intensity (solar flux), the ambient temperature, the temperature of the evaporators of refrigeration, and air conditioning sub-systems is examined. Influence of solar flux is investigated on the power produced by turbine of OFC and on the cooling outputs.

4.2 Effect of climatic conditions on cogenerations system performance

The performance of the cogeneration system depends on the climatic conditions; the solar flux and the ambient temperature.

Fig. 2 shows the effect of variation of solar flux on the temperature (T_{18}) of SHTF leaving the PTSC where the values of outlet temperature of the collector are different from gas to gas. Promotion of solar flux (G_b) raises the incident rate of solar energy received by PTSC which leads to increase in the useful heat gained by the SHTF as it is indicated from eq. (6). For the fixed mass flow rate of SHTF, increase in useful heat gained by the SHTF results in the rise of outlet temperature of SHTF (T_{18}) leaving the PTSC. This trend is observed for all three SHTFs (Air, CO_2, and He). So, solar flux (G_b) has the positive effect on the temperature of SHTF (T_{18}) leaving the PTSC. The outlet temperature of SHTF (T_{18}) is achieved highest for He and lowest for CO_2 owing to higher thermal conductivity of He and lower for CO_2. For He, it increases from 312.5 ℃

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

to 434.7 ℃, for Air from 300 ℃ to 413.6 ℃, and for CO_2 from 283.9℃ to 387.1℃, when solar flux is increased from 600-1000 W/m^2. The highest outlet temperature in case of He is achieved because of its higher heat transfer coefficient and lowest pressure drop. Higher heat transfer coefficient assists in increasing the heat transfer from the absorber to fluid which reduces the absorber temperature and resulted in lower thermal losses. Based on these findings, Helium can be considered as the most suitable solar heat transfer fluid for the PTSC employed to drive the cooling-power cycles applied in large number of applications up today.

Fig. 3 illustrates the variation in mass flow rate of the working fluid (R124) of OFC by changing solar flux (600-1000 W/m^2) for three different SHTFs namely Air, CO_2, and He. This is explained as the value of solar flux increases, useful heat gain by SHTF increases and consequently, more heat is supplied to the OFC working fluid at Evaporator (EVA) which resulted in high flow rate of OFC working fluid because of the rise of temperature of SHTF which in turn raise the degree of superheat at the superheater (SH) as shown by equations of EVA and SH respectively in Table 2. The mass flow rate of the working fluid of OFC cycle is highest when the SHTF is He followed by CO_2 and Air because the specific heat of He is highest followed by CO_2 and Air. The mass flow rate of OFC cycle for He increases from 37.19 kg/s to 60.9 kg/s, for CO_2 from 36.48 kg/s to 60.16 kg/s, and for Air from 35.76 kg/s to 59.38 kg/s with the change in solar flux from 600-1000 W/m^2.

It is clear from the Fig. 1 that the Turbine power ($\dot{W}_T$) is produced by three streams of OFC: first stream (22) is coming from SH, second stream (35) is coming from High Pressure Flasher (HPF) and third stream (37) is coming from Low Pressure Flasher (LPF). As explained in Fig. 3, with the increase in solar flux, useful heat gain by SHTF increases which results in the increased mass flow rate and increased superheating of OFC working fluid. So, power produced by first stream (22), increases with the increase in solar flux. As the pressure of HPF is fixed, the temperature of flash

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

stream (35) is fixed. So, increase in solar flux, results in the increased mass flow rate of flash vapor (35) of OFC working fluid. So, power produced by second stream (35) increases with the increase in solar flux. Similarly, pressure of LPF is fixed, the temperature of flash stream (37) is fixed. So, increase in solar flux, results in the increased mass flow rate of flash vapor (37) of OFC working fluid. So, power produced by third stream (37) increases with the increase in solar flux. So, total power produced by turbine ($\dot{W}_T$) will be the summation of the power produced by all three streams. The variation of solar flux on the power produced ($\dot{W}_T$) by the turbine of OFC cycle is shown in Fig. 4. It is evident that the turbine output ($\dot{W}_T$) of the OFC cycle increases for all SHTFs (Air, CO_2, and He) with the increase in the solar flux. The increased turbine power will not only increase the electricity but also produce more cooling because of the greater amount of power supplied to compressor of two-phase ejector. This unique system is designed such a way that the excess amount of electricity will be supplied to off grid areas at the time of off-peak cooling load requirements. It is also noticed that the turbine output is highest for He and followed by the CO_2 and Air which is due to the highest mass flow rate of working fluid (R124) of OFC cycle for He followed by the CO_2 and Air. Turbine output ranges from 981.4 kW to 1600 kW for He, 960.2 kW to 1579 kW for CO_2, and 939.7 kW to 1557 kW for Air, with the change in the solar flux from 600-1000 W/m^2.

The turbine output of the OFC cycle is utilized for two purposes. Half of the power produced by turbine is employed for electricity generation and remaining power output is utilized to drive the compressor of ejector cycle producing the refrigeration and air-conditioning outputs. It is explained in Fig.4, increase in solar flux leads to increase in turbine power output ($\dot{W}_T$) of the OFC cycle. This is because increase of solar flux raises the evaporator temperature which correspondingly increase the turbine entry pressure leading to higher pressure ratio of turbine. As half of the turbine power is utilized to run the compressor of ejector cycle, the compressor input

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

power also increases with the increase in solar flux. Increase in compressor power leads to increase the refrigeration output as well air-conditioning output of the two-phase ejector cycle. Fig. 5 depicts the variation of the refrigeration output and air-conditioning output of ejector cycle with solar flux. It is shown that both the outputs of two-phase ejector cycle increase for all SHTFs (Air, CO_2, and He) with the increase in the solar flux. For the He, the refrigeration output ranges from 1245 kW to 2063 kW and air-conditioning output from 376.7 kW to 624.1 kW is highest, followed by CO_2 [refrigeration output from 1261 kW to 2081 kW and air-conditioning output from 381.9 kW to 630 kW] and for Air it is lowest [refrigeration output from 1276 kW to 2097 kW and air-conditioning output from 387 kW to 635.4 kW] with the increase of solar flux from 600-1000 W/m^2.

Fig. 6 depicts the effect of increase of solar flux on cooling exergy outputs of two-phase ejector cycle. According to the figure, both the exergetic outputs of two-phase ejector are improved for all SHTFs (Air, CO_2, and He) with the rise in the solar flux. For the He, the exergy of refrigeration output ranges from 241.5 kW to 394 kW and exergy of air-conditioning output from 10.008 kW to 15.94 kW is highest, followed by CO_2 [exergy of refrigeration output from 236.5 kW to 388.8 kW and exergy of air-conditioning output from 9.506 kW to 15.33 kW] and for Air it is lowest [exergy of refrigeration output from 231.3 kW to 383.3 kW and exergy of air-conditioning output from 9 kW to 14.91 kW] with the increase of solar flux from 600-1000 W/m^2. The reason of increase in both the exergy outputs being the increase in turbine output of OFC cycle due to the increase in mass flow rate of the working fluid of OFC cycle which is utilized to run the compressor of this cycle as mentioned in Fig. 4. The exergetic output of air conditioning sub-system is found to be significantly lower than the exergetic output of refrigeration sub-system. This is because the evaporator temperature of refrigeration system is far away from the ambient temperature and hence provide larger exergy compared to air conditioning which is close to ambient. This clearly shows

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

analysis of exergy is an important tool for the development of strategies to use energy resource for the production of cooling at different levels of temperature.

Thermal efficiency of the system is defined as overall useful energy output (net electrical output, refrigeration output and air-conditioning output) to the total energy input ($\dot{Q}_{solar}$) as defined in eq. (5). Similarly, the exergy efficiency of the system is defined as overall exergy output (net electrical output, exergy of refrigeration output and exergy of air-conditioning output) to the total exergy input ($\dot{E}_{solar}$) as defined in eq. (8). The variation of thermal and exergy efficiencies of the system for different SHTFs (Air, CO_2, and He) with the change in solar flux as shown in Fig. 7. The thermal and exergy efficiency is found to decrease for all the SHTFs with the increase in the solar flux. With the increase in solar flux, the increase in the energy output (i.e., net electric output of turbine, refrigeration output, and air-conditioning output) is less in comparison to increment in the solar energy input which lead to the decrease the thermal efficiency of the system for all SHTFs. For He, the thermal efficiency decreases from 16.72% to 16.62%, for CO_2, it decreases from 16.71% to 16.61%, and for Air 16.70% to 16.60% with the increase in solar flux from 600- 1000 W/m^2. It is also important to note that with the increase in solar flux, the increase in the exergy outputs (summation of all exergetic outputs) is less in comparison to increment in the exergy of solar energy input which lead to decrease the exergy efficiency of the system for all SHTFs. Furthermore, the exergy efficiency for He decreases from 6.096 % to 6.067%, for CO_2 from 6.094% to 6.063%, and for Air from 6.091% to 6.054% with the increase in solar flux from 600-1000 W/m^2. This is because the increase in the solar flux leads to high temperature of the SHTF which increases the thermal losses (high losses of energy and exergy at high temperature) from the PTSC.

The variation of thermal and exergy efficiencies of the system for different SHTFs (Air, CO_2, and He) with the change in ambient temperature as shown in Fig. 8. The thermal and exergy efficiency

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

is found to decrease marginally for all the SHTFs with the increase in the ambient temperature. For He, the thermal efficiency decreases from 16.72% to 16.68%, for CO_2, it decreases from 16.70% to 16.66%, and 16.68% to 16.64% for Air. Since an increase in ambient temperature decreases the thermal losses between the system and environment which provide an increment in the energy input to the cycle and this leads to decrease the thermal efficiency of the system for all SHTFs. Since exergy accompanied by thermal losses is much less than the thermal losses, therefore, an increment in the exergy supplied to cycle is greater than its energy counterpart and hence, the exergy efficiency of the system is less than its thermal efficiency for all three fluids. Further, the rate of decline in exergy efficiency is greater than the thermal efficiency owing to larger decrease in the exergy accompanied by cooling outputs produced by refrigeration and air conditioning sub-systems. The exergy efficiency for He decreases from 6.099 % to 6.064 %, for CO_2 from 6.097% to 6.062%, and for Air from 6.087% to 6.056% with the increase in ambient temperature from 17^0C to 37^0C. This is because the increment in the exergy output is less than the increment in the exergy input.

4.3 Effect of varying the OFC turbine inlet pressure on cogeneration performance

Fig. 9 depicts the impact of variation of turbine inlet pressure on the thermal and exergy efficiency of the system. As the turbine inlet pressure increases, enthalpy of vaporization of the working fluid (R124) of OFC decreases and consequently mass flow rate of working fluid (R124) of OFC increase. So, for same amount of heat input in PTCS, the work output of turbine increases which results in increase in the energy output (i.e., net electric output of turbine, refrigeration output, and air-conditioning output) and consequently thermal efficiency increases. In similar fashion, for same amount of exergy input in PTCS, the work output of turbine increases which results in increase in the exergy output (i.e., net electric output of turbine, exergy of refrigeration output, and exergy of air-conditioning output) and consequently exergy efficiency increases. So, thermal and

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

exergy efficiency of the system increases for all SHTFs with the increase in turbine inlet pressure of the OFC which is due to the increase in mass flow rate of working fluid (R124) of OFC. For He, the thermal efficiency is highest which increases from 14.83% to 19.38%, followed by CO_2 and increases from 14.62% to 19.17%, and lowest for Air, from 14.41% to 18.95% with the increase the turbine inlet pressure from 1800 kPa to 2200 kPa. Furthermore, the exergy efficiency for He is highest which increases from 5.463 % to 7.123%, followed by CO_2 and increases from 5.361% to 7.021%, and lowest for Air, from 5.256% to 6.914% with the increase in turbine inlet pressure from 1800 kPa to 2200 kPa.

4.4 Effect of varying the refrigeration and air conditioning evaporator temperatures on cogeneration system efficiencies

Fig. 10 illustrates the effect of change of refrigeration evaporator temperature (T_{ref}) on the thermal and exergy efficiencies of the overall system. An increase in refrigeration evaporator temperature (from -25^0C to -15^0C), the refrigeration output ($\dot{Q}_{ref}$) and air-conditioning output ($\dot{Q}_{AC}$) of two-phase ejector cycle increases due to which the energetic output of the proposed system increases. This is due to increase in mass flow rate of the refrigerant (R600a) of two-phase ejector cycle and hence its thermal efficiency increases. For He, the thermal efficiency is highest which increases from 15.72% to 18.19%, followed by CO_2 and increases from 15.62% to 18.08%, and lowest for Air, from 15.50% to 17.97%, with the increase the refrigeration evaporator temperature from -25^0C to -15^0C. On the other hand, exergy efficiency of the overall system decreases with the increase of the refrigeration evaporator temperature. The decrement in the exergy efficiency of the cycle is because the exergy of refrigeration output reduces with the rise of refrigeration evaporator temperature. The exergy efficiency for He is highest which decreases from 6.138 % to 6.018%, followed by CO_2 and decrease from 6.133% to 6.015%, and lowest for Air, from 6.128% to 6.009% with the increase in refrigeration evaporator temperature from -25^0C to -15^0C.

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

Fig. 11 shows the effect of change of air-conditioning (AC) evaporator temperature (T_{AC}) on the thermal and exergy efficiencies of the overall system. An increase in AC evaporator temperature (from 15^0C to 25^0C), the refrigeration output ($\dot{Q}_{ref}$) and air-conditioning output ($\dot{Q}_{AC}$) of two-phase ejector cycle increases due to which the energetic output of the proposed system increases. This is due to increase in mass flow rate of the refrigerant (R600a) of two-phase ejector cycle and hence its thermal efficiency increases. For He, the thermal efficiency is highest which increases from 16.34% to 17.06%, followed by CO_2 and increases from 16.32% to 17.03%, and lowest for Air, from 16.29% to 17.00%, with the increase the AC evaporator temperature from 15^0C to 25^0C. Similarly, the exergy efficiency of the overall system increases marginally with the increase in AC evaporator temperature. The reason being exergy of refrigeration output increases and exergy of air-conditioning output decreases resulting in marginal increase in the exergy efficiency with the rise of AC evaporator temperature. The exergy efficiency for He is highest which increases from 6.076 % to 6.087%, followed by CO_2 and increases from 6.074% to 6.085%, and lowest for Air, from 6.068% to 6.079% with the increase in AC evaporator temperature from 15^0C to 25^0C.

4.5 Breaking down of solar exergy input into the exergy produced, exergy destroyed, and the exergy loss

Fig. 12 show the breaking down of solar exergy and its destruction due to generation of entropy in various components of the system when it is operated at baseline conditions. The exergy accompanied by solar rays falling on the PTSC is considered as 100% exergy source driving the proposed cogeneration. The highest exergy destruction rate is observed at PTSC (53.87%) followed by OFC (22.09%), two phase ejector cycle (6.15%), and remaining 17.89% is the exergy output of cogeneration system. This percentage of exergy reveals critical insights into the performance of individual sub-systems, highlighting areas for performance improvement of real-world cogeneration systems driven by solar energy. Computation of exergy destruction in sub-systems assist in diagnosing the real potential of this newly developed cogeneration system.

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

Comparing with the exergy destroyed in the PTSC employing thermal oil as solar heat transfer fluid (58.47%) [21], the exergy destructed in presently used PTSC (53.87%) is considerably lesser. Results shows PTSC has by far the highest exergy destructive sub-system, followed by the OFC. Therefore, it seems that system modifications, of which the aim is to reduce exergy destruction in the PTSC and OFC, have a major potential to increase the power output of OFC for enhancing the electric power output. The exergy destruction in the OFC can be mitigated by employing the zeotropic mixtures instead of pure organic fluids as the working medium. The third highest exergy destructive sub-system is the two-phase ejector where majority of the exergy is destroyed in its condenser and evaporators. The exergy in these components can be significantly mitigated by increasing the evaporator temperature and decreasing the condenser temperature which in turn decrease the exergy destroyed in the two-phase ejector. This will contribute towards an enhancement in the cooling output desired to refrigerate the thermal loads of air conditioning and refrigeration. Classification of the output exergy gives turbine output (4.02%), refrigeration output (10.65%) and air-conditioning output (3.22%). The exergy destruction in PTSC is dominant in the whole system because high quality solar insolation is absorbed by the receiver at higher temperature which leads to greater thermodynamic losses in the collector. The second highest destruction of solar exergy is found in the OFC mainly due to throttling of cycle working fluid from economizer pressure to evaporator pressure which is the saturation pressure corresponding to temperature of high-pressure flasher. Expansion of the mixture of two-streams in the turbine with the significant difference exergy of two streams further contribute to larger destruction of solar exergy by the OFC. The two-phase ejector is determined as the third largest destruction of solar exergy which is majorly due to friction losses associated with the fluid stream passes through the ejector and due to the entropy generated via non-isentropic expansion of flowing stream in the nozzle. Separation of refrigerant R600a stream into two streams; saturated vapor (state 10) and

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

saturated liquid (state 7) and the throttling of saturated liquid in the expansion valve (7- 8) further contributes to destruction of exergy in the cooling cycle.

5. Conclusions

In conclusion, this investigation undertook an in-depth evaluation of an innovative solar-driven cooling-power cogeneration system, seamlessly integrating diverse technologies such as parabolic trough collectors, organic flash cycle, and two-phase ejector cooling cycle. Through rigorous energy and exergy analyses, the quantified system performance yielded key quantitative outcomes affirming its efficacy. Important features of this proposal are that system does not rely on electricity from grid. Part of electricity generated by OFC is supplied to grid, and remaining is expended in driving the compressor of two-phase ejector which is aimed to produce refrigeration and air conditioning. The influence of using gases as solar fluid and change of weather conditions have been examined on the energetic and exergetic performances of the devised cogeneration and major findings of the study can be shown as below:

- Helium (He) is determined to be the gaseous SHTF which absorbed highest solar energy by showing a 28.11% increase in its temperature while passing through the collector and in case of the use of CO_2 it is increased by 26.65% when solar flux is promoted from 600 to 1000 W/m^2.

- By increase of solar flux from 600 to 1000 W/m^2, mass flow rate of the working medium of OFC is significantly enhanced but it is marginally varied with the change in type of SHTF employed in PTC.

- OFC turbine output, and the refrigeration and air conditioning outputs of two-phase ejector shows upward trends when the solar flux increases from 600 to 1000 W/m^2.

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

- Cooling exergy produced by two-phase ejector is found considerably increasing with the rise of solar flux. Rate of increase in exergy of refrigeration is greater than the rate of exergy accompanied by air conditioning.

- The impact of varying the OFC turbine inlet pressure on system efficiencies is considerably large but change of solar flux shows a little impact on them.

- Results obtained for the distribution of solar exergy in the system components reveal that 53.87% of solar exergy is destroyed by PTSC followed by 22.09% in OFC, and 6.15% in two-phase ejector cycle, and remaining solar exergy is converted into power exergy (4.02%), refrigeration exergy (10.65%) and exergy accompanied by air-conditioning (3.22%). The findings clearly reveal PTSC as the maximum exergy destructive component (53.87%) of the system and it needs improvement on priority. The exergy destruction in the PTSC can be obtained by employing use of multiple concentration ratios in the collector, use of internal fins in absorber of PTSC, and by applying coating over absorber tubes which decrease the loss of heat from collector to environment and therefore, by decreasing the transfer of heat between ambient and system boundary the entropy transfer via heat reduces which in turn decrease dissipation of exergy.

This investigation's unique integration of multiple technologies paves a sustainable pathway to address the ever-increasing demands of a growing global population and rising living standards. As results through this research provides essential groundwork, subsequent efforts must translate the immense promise into practical implementations.

The proposed system of cooling and power come with numerous scopes and novelty. The developed model mainly investigates on a few operational parameters, whereas there are many more aspects of employment in real-world scenario such as environmental impact

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

and economic expenses along with the consideration of the variation of heat losses, pressure drop, and component efficiency. All investigations are conducted with the steady-state condition which limits the ability to predict the variation of cooling loads of refrigeration and air conditioning across various times of the day. To facilitate the scopes of improvements of this investigation, an advanced exergy analysis, exergo-economic evaluation, multi-objective optimization can be carried out. Investigation into the scalability and adaptability of the proposed cooling and power system for different geographical locations and CSP technologies will be essential for its broader implementation, ensuring a comprehensive and robust approach.

Acknowledgement

The authors acknowledge the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research at King Faisal University, Saudi Arabia, for financial support under the annual funding track [KFU251111].

Nomenclature

Symbols

A_p Aperture area [m^2]

c Bulk velocity of fluid [m/s]

e_x Specific flow exergy [kJ/kg]

$\dot{E}x$ Exergy rate [kW]

G_b Solar flux [W/m^2]

g Acceleration due to gravity [m/s^2]

h Specific enthalpy [kJ/kg]

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

$\dot{m}$ Mass flow rate [kg/s]

P Pressure [MPa]

$\dot{Q}$ Heat transfer rate [kW]

T Temperature [K]

s Specific entropy [kJ/kg K]

$\dot{W}$ Output power [kW]

z Elevation of flowing stream [m]

Greek symbols

η Efficiency [%]

∑ Summation

Superscript

***** property at datum state

Subscript

abs Absorbed

AC Air-conditioner/ Air-conditioning

0 Ambient

C Compressor

Cond Condenser

D Destruction

Eje Ejector

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

EV Expansion valve

ECO Economizer

EVA Evaporator

en Energy

ex Exergetic

F Fuel

i Inlet

HX Heat exchanger

L loss

o Outlet

PTSC Parabolic Trough Solar Collector

P Product

s-r Sun to receiver

r-f receiver to fluid

opt Optical

r Receiver

ref Refrigerator/refrigeration

Sep Separator

SH Superheater

Solar Sun

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

T	Turbine
th	Thermal
u	Useful
1,2,3..	State points

Abbreviations

ARC	Absorption Refrigeration Cycle
CCP	Combined Cooling and Power
CO_2	Carbon Dioxide
Cond	Condenser
COP	Coefficient of Performance
DNI	Direct Normal Irradiations
EES	Engineering Equation Solver
Eje	Ejector
ERC	Ejector Refrigeration Cycle
ECO	Economizer
EV	Expansion valve
EVA	Evaporator
HRVG	Heat Recovery Vapor Generator
He	Helium

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

HX	Heat exchanger
HPF	High Pressure Flasher
LPF	Low Pressure Flasher
OFC	Organic Flash Cycle
ORC	Organic Rankine Cycle
PTSC	Parabolic Trough Solar Collector
ref	Refrigerator
Sep	Separator
SH	Superheater
SHTF	Solar Heat Transfer Fluid
SRC	Steam Rankine Cycle

References

[1] Çakir, U., Çomakli, K., Yüksel, F., 2012, "The role of cogeneration systems in sustainability of energy," Energy Convers. Manage. 63 pp. 196-202.

[2] Bouguetaia, N., Bellel, N., Lekbir, A., 2023, "Absorption chiller system driven by the solar hybrid system: Case study in the Algeria weather condition," ASME J. Thermal Sci. Eng. Appl. 15 (6) pp. 061009.

[3] Kalogirou, S.A., 2004, "Solar thermal collectors and applications," Progress in Energy and Combustion Science 30 (3) pp. 231-295.

[4] Al Mughanam, T., Khaliq, A., 2023 "Thermodynamic investigation of a novel synergetic integration of a solar based Kalina cycle and ejector refrigeration cycle," Int J Exergy 41 (4) pp. 403-430.

[5] Abdullatif, Y.M., Okonkwo, E.C., Al-Ansari, T., 2021, "Thermal performance optimization of a parabolic trough collector operating with various working fluids using copper nanoparticles," ASME J. Thermal Sci. Eng. Appl. 13 (5) pp. 051011.

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

[6] Siddiqui, M. A., 2024, "Thermodynamic Analysis and Performance Assessment of a Novel Solar-Based Multigeneration System for Electricity, Cooling, Heating, and Freshwater Production," ASME J. Sol. Energy Eng., 146(2) pp. 021007.

[7] Hasan, A. A., Goswami, D.Y., 2003, "Exergy analysis of a combined power and refrigeration thermodynamic cycle driven by a solar heat source," ASME J. Sol. Energy Eng. 125(1) pp. 55-60.

[8] Khaliq, A., Kumar R., Mokheimer, E.M.A., 2018, "Investigation on a solar thermal power and ejector-absorption refrigeration system based on first and second law analyses," Energy, 164 pp. 1030-1043.

[9] Bao, J., Zhao, Li., 2013, "A review of working fluid and expander selections for organic Rankine cycle for solar thermal power generation," Renew. Sustain. Energy Rev. 24 pp. 325-342.

[10] Grosu, L., Marin, A., Dobrovicescu, A., Oueiros-code, D., 2016, "Exergy Analysis of a Solar Combined Cycle: organic Rankine cycle and absorption cooling system," Int. J. of Energy and Environmental Engineering 7 pp. 449–459.

[11] Eisavi, B., Khalilarya, S., Chitsaz, A., Rosen, M.A., 2018, "Thermodynamic analysis of a novel combined cooling, heating and power system driven by solar energy," Applied Thermal Engineering 129 pp. 1219-1229.

[12] Ozgoren, M., Bilgili, M., Babayigit, O., 2012, "Hourly performance prediction of ammonia-water solar absorption refrigeration," Applied Thermal Engineering 40 pp. 80-90.

[13] Khaliq, A., Mathkar, M.A., Alqaed, S., Mokheimer, E.M.A., Kumar, R., 2020, "Analysis and assessment of tower solar collector driven trigeneration system," ASME Trans. Journal of Solar Energy Engineering," 142 pp. 051003-1.

[14] Yu, Z., Han, J., Liu, H., Zhao, H., 2014, "Theoretical Study on a Novel Ammonia-Water Cogeneration System with Adjustable Cooling to Power Ratios," Appl Energ 122 pp. 53-61.

[15] Abbady, K., Al-Mutawa, N., Almutairi, A., 2023, "Energy and exergy analysis of an automobile hybrid ejector refrigeration system utilizing its exhaust waste heat," ASME J. Thermal Sci. Eng. Appl. 15 (9): pp. 0191003.

[16] Siddiqui, M. A., Khaliq, A., and Kumar, R., 2022, "Thermodynamic and Comparative Analysis of Ejector Refrigeration Cycle and Absorption Refrigeration Cycle Integrated

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

Wet Ethanol-Fueled HCCI Engine for Cogeneration of Power and Cooling," ASME J. Thermal Sci. Eng. Appl., 14(4) pp. 041003.

[17] Dai, Y., Wang, J., Gao, L., 2009, "Exergy Analysis, Parametric Analysis and Optimization for a Novel Combined Power and Ejector Refrigeration Cycle," Appl. Therm. Eng. 29 pp. 1983-1990.

[18] Khaliq, A., 2017, "Energetic and exergetic performance investigation of a solar based integrated system for cogeneration of power and cooling," App. Therm. Eng. 112 pp. 1305-1316.

[19] Zhang, T., Mohamed, S., 2015, "Conceptual Design and Analysis of Hydrocarbon-Based Solar Thermal Power and Ejector Cooling Systems in Hot Climates," ASME J. Sol. Energ. 137 pp. 021001.

[20] Elakhdar, M., Landoulsi, H., Tashtoush, B., Nehdi, E., Kairouani, L., 2019, "A combined thermal system of ejector refrigeration and organic Rankine cycle for power generation using a solar parabolic trough," Energy Convers. Manage. 199 pp.111947.

[21] Mahdavi, N., Ghaebi, H., Minaei, A., 2022, "Proposal and multi-aspect assessment of a novel solar-based trigeneration system; investigation of zeotropic mixture's utilization," Appl. Therm. Eng., 206 pp. 118110.

[22] Yadav, V.K., Sarkar, J., Ghosh, P., 2024, "Performance optimization and multi-objective analysis of an innovative solar-driven combined power and cooling system," Energy and Buildings, 307 pp. 113943.

[23] Dagdas, A., "Performance analysis and optimization of double-flash geothermal power plants," ASME J Energy Resour Technol 129 pp. 125-133.

[24] Shokati, N., Ranjbar, F., Yari, M., 2015, "Comparative and parametric of double flash and single flash/ORC combined cycles based on exergoeconomic criteria," Appl Therm Eng 91 pp. 479-495.

[25] Lee, H.Y., Park, S.H., Kim, K.H., 2016, "Comparative analysis of thermodynamic performance and optimization of organic flash cycle (OFC) and organic Rankine cycle (ORC)," Appl Therm Eng 100 pp. 680-690.

[26] Varma, G.V.P., Srinivas, T., "Power generation from low temperature heat recovery," Renewable and Sustainable Energy Reviews 75 pp. 402-414.

[27] Kornhauser, A.A., 1990, "The use of an ejector as a refrigerant expander," International Refrigeration and Air Conditioning Conference. Paper 82.

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

[28] Chaiwongsa, P., Wongwises, S., 2007, "Effect of throat diameters of the ejector on the performance of the refrigeration cycle using a two-phase ejector as an expansion device," Int. J. Refrigeration 30 pp. 601-608.

[29] Pottker, G., Guo, B., Hrnjak, P.S., 2010, "Experimental investigation of an R410A vapor compression system working with an ejector," International Refrigeration and Air Conditioning Conference, Paper 1135.

[30] Lawrence, N.D., 2012, "Analytical and experimental investigation of two-phase ejector cycles using low-pressure refrigerants (MS Thesis)," University of Illinois, Urbana, IL, USA. http://hdl.handle.net/2142/42337

[31] Lawrence, N., and Elbel, S., 2013, "Theoretical and practical comparison of two-phase ejector refrigeration cycles including first and second law analysis," Int. J. Refrigeration 36 pp. 1220-1232.

[32] Wang, X., and Yu, J., 2016, "An investigation on the component efficiencies of a small two-phase ejector," Int. J. Refrigeration 71 pp. 26-38.

[33] Khaliq, A., Refaey, H.A., Alharthi, M.A., 2021, "Development and analysis of a novel CSP source driven cogeneration cycle for the production of electric power and low temperature refrigeration," Int. J. Refrigeration 130 pp. 330–346.

[34] Bellos, E., Tzivanidis, C., Antonopoulos, K.A., 2017, "A detailed working fluid investigation for solar parabolic trough collectors," Appl. Therm. Eng. 114 pp. 374-386.

[35] Chao Xu, Zhifeng Weng, Xin Li, Feihu Sun, 2011, "Energy and exergy analyses of solar power tower plants," Applied Thermal Engineering 31 pp. 3904-3913.

[36] Almatrafi, E., Khaliq, A., Alquthami, T., 2022, "Thermodynamic investigation of a novel cooling-power cogeneration system driven by solar energy," Int. J. Refrigeration 138 pp. 244-258.

[37] Almatrafi, E., Khaliq, A., Kumar, R., Bamasag, A., Ehtisham, M., 2023, "Thermodynamic analysis of a solar refrigeration system based on combined supercritical CO2 power and cascaded refrigeration cycle," Int. J. Exergy 41 (2) pp. 182-199.

[38] Klein, S. A., 2012, "Engineering Equation Solver (EES) for Microsoft Windows operating systems: Academic professional version," F-chart software, Medison, W.I. (Available online http://www.fchart.com)

[39] REFPROP., 2013., "NIST Reference Thermodynamic and Transport Properties," version 9.1.

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

FIGURE CAPTIONS:

Fig. 1. Schematic diagram of proposed cogeneration system

Fig. 2. Variation of outlet temperature of different SHTF for with the change in solar flux

Fig. 3. Variation of mass flow rate of OFC working fluid for different SHTF with the change in solar flux

Fig. 4. Variation of turbine output of OFC for different SHTF with the change in solar flux

Fig. 5. Variation of refrigeration output and air-conditioning output of ejector refrigeration cycle for different SHTF with the change in solar flux

Fig. 6. Variation of exergy outputs of ejector refrigeration cycle for different SHTF with the change in solar flux

Fig. 7. Variation of thermal and exergy efficiencies of the cogeneration system for different SHTF with the change in solar flux

Fig. 8. Variation of thermal and exergy efficiencies of the cogeneration system for different SHTF with the change in Ambient temperature

Fig. 9. Variation of thermal and exergy efficiencies of the cogeneration system for different SHTF with the change in turbine inlet pressure of OFC

Fig. 10. Variation of thermal and exergy efficiencies of the cogeneration system for different SHTF with the change in refrigeration evaporator temperature of ERC

Fig. 11. Variation of thermal and exergy efficiencies of the cogeneration system for different SHTF with the change in AC evaporator temperature of ERC

Fig. 12. Percentage distribution of exergy input into exergy destruction and exergy output of the proposed cogeneration system

TABLE CAPTIONS:

Table 1. Thermal properties of SHTF at mean temperature (377 ℃)[34]

Table 2. Energy balance equations for the components of the proposed cogeneration system

Table 3. Exergy destruction equations for the components of the proposed cogeneration system

Table 4. Input data for modeling of the Proposed cogeneration system

Table 5: Comparison of the computed results with the reported results of literature [26, 31]

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

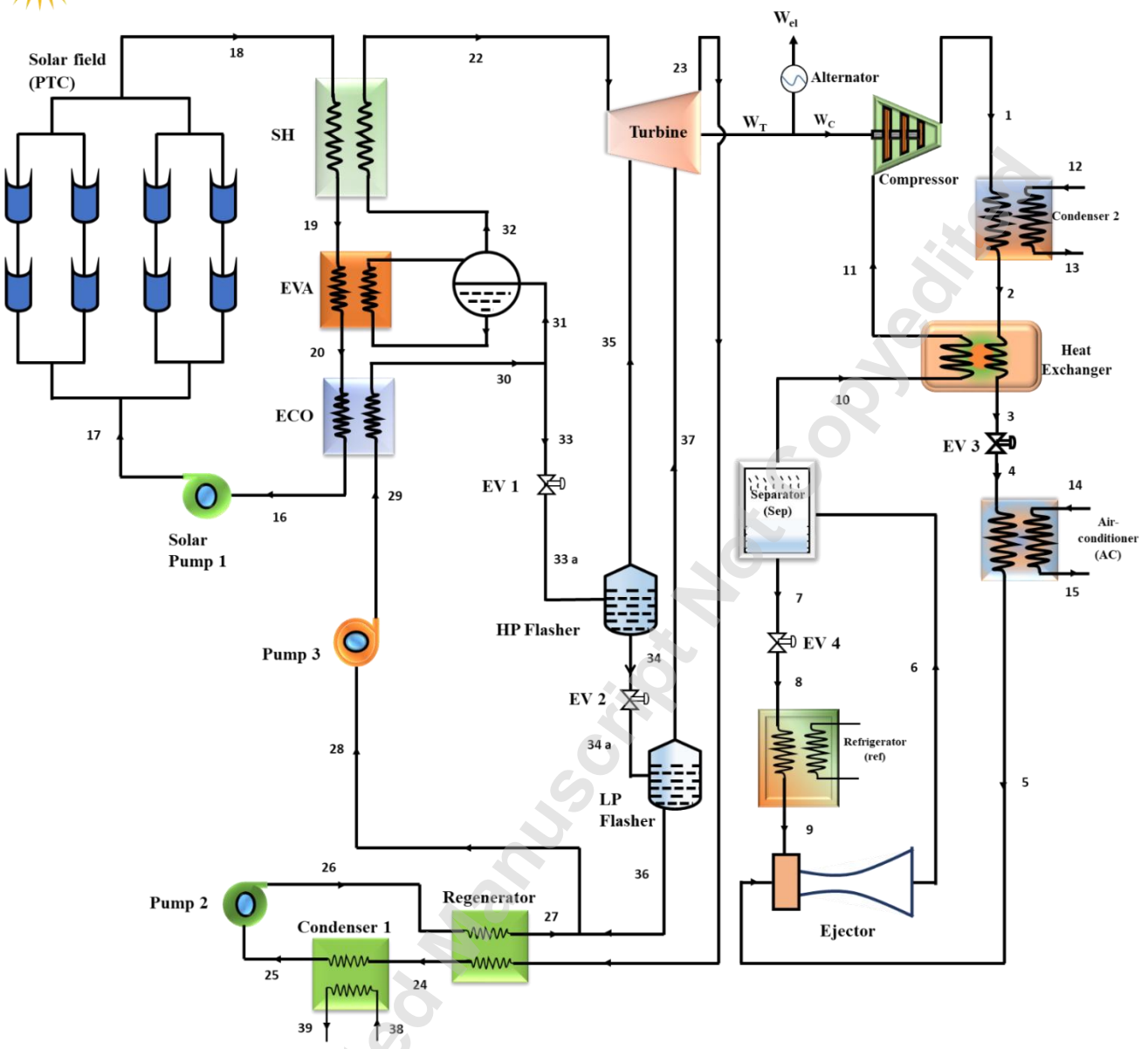


Fig. 1. Schematic diagram of proposed cogeneration system

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

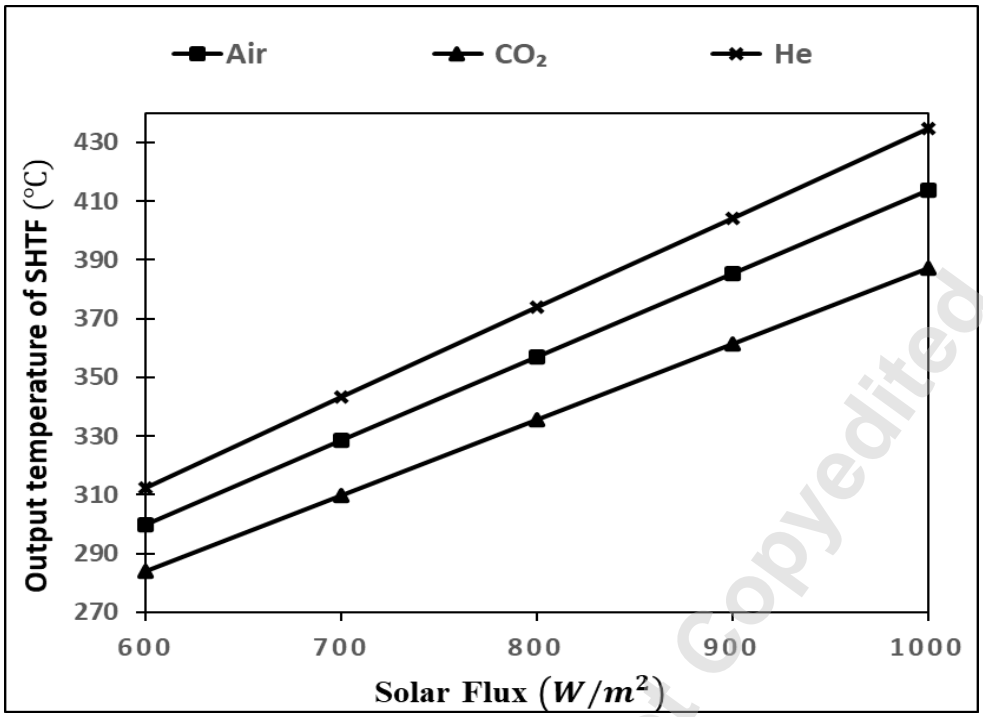
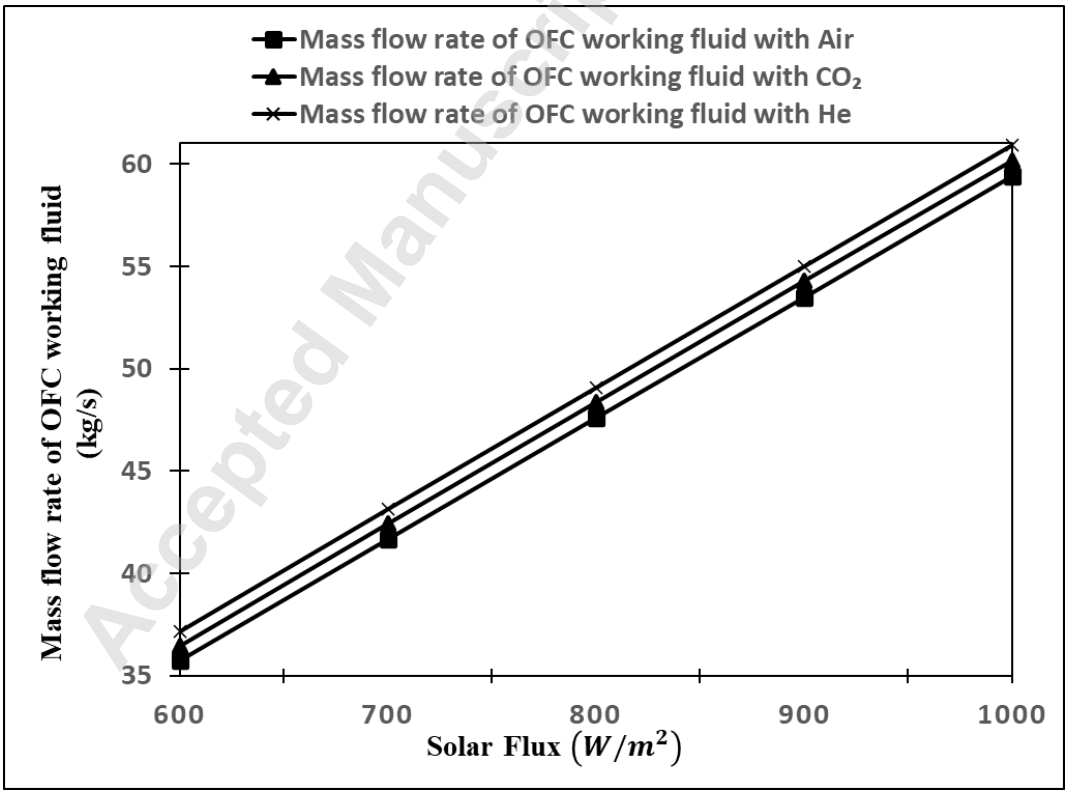


Fig. 2. Variation of outlet temperature of different SHTF for with the change in solar flux



Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

Fig. 3. Variation of mass flow rate of OFC working fluid for different SHTF with the change in solar flux

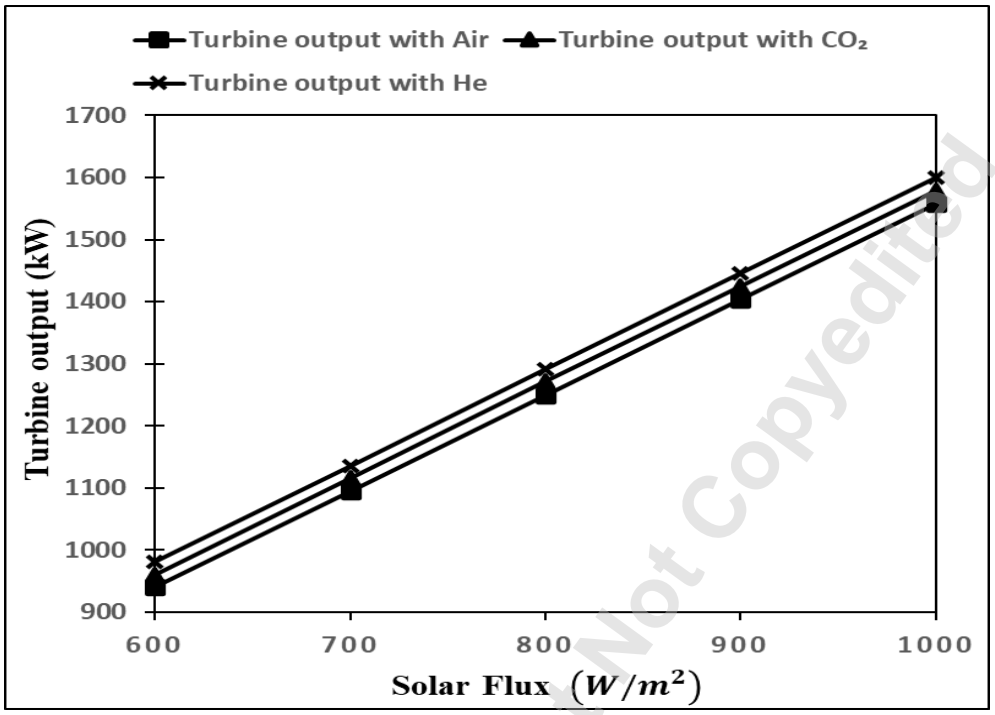
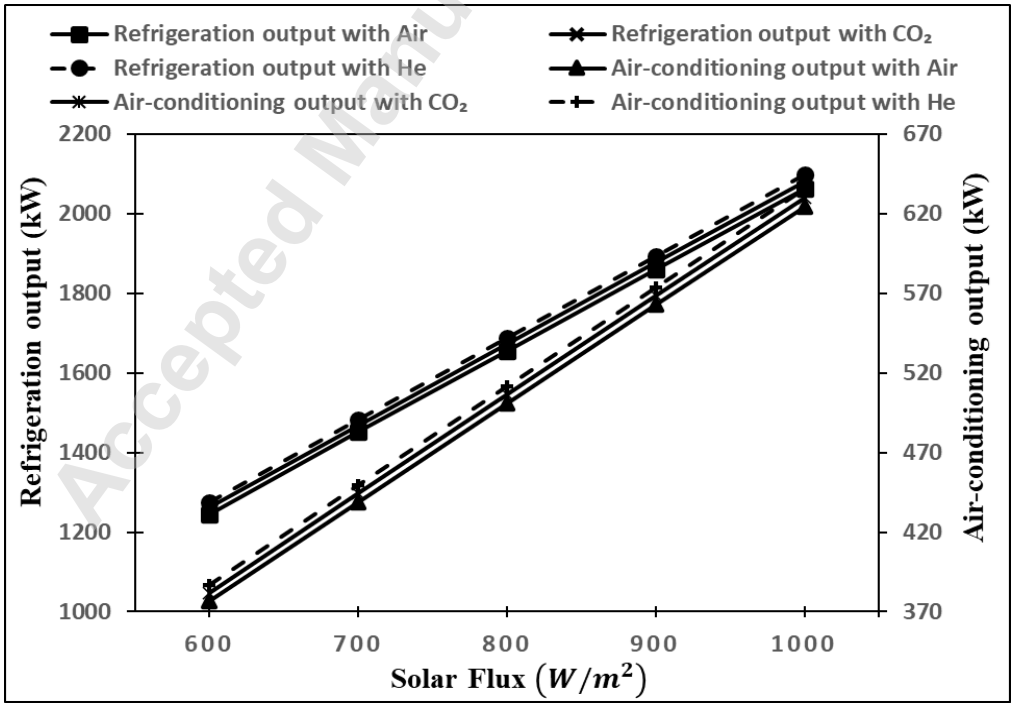


Fig. 4. Variation of turbine output of OFC for different SHTF with the change in solar flux



Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

Fig. 5. Variation of refrigeration output and air-conditioning output of ejector refrigeration cycle for different SHTF with the change in solar flux

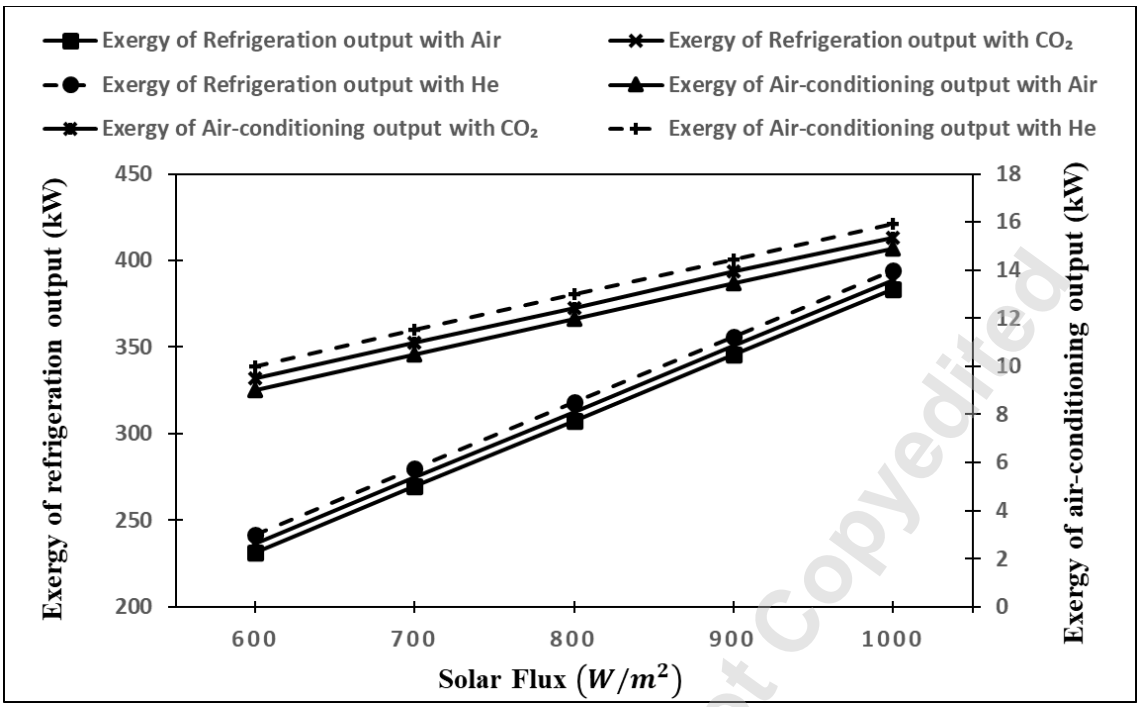


Fig. 6. Variation of exergy outputs of ejector refrigeration cycle for different SHTF with the change in solar flux

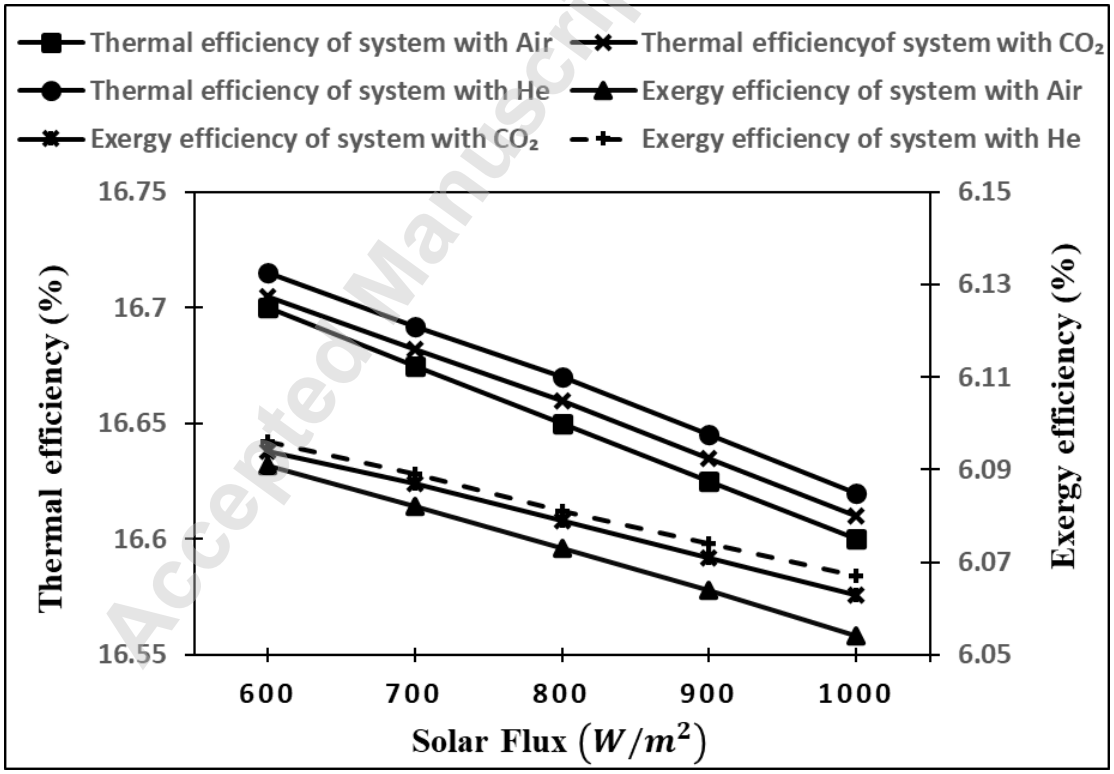


Fig. 7. Variation of thermal and exergy efficiencies of the cogeneration system for different SHTF with the change in solar flux

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

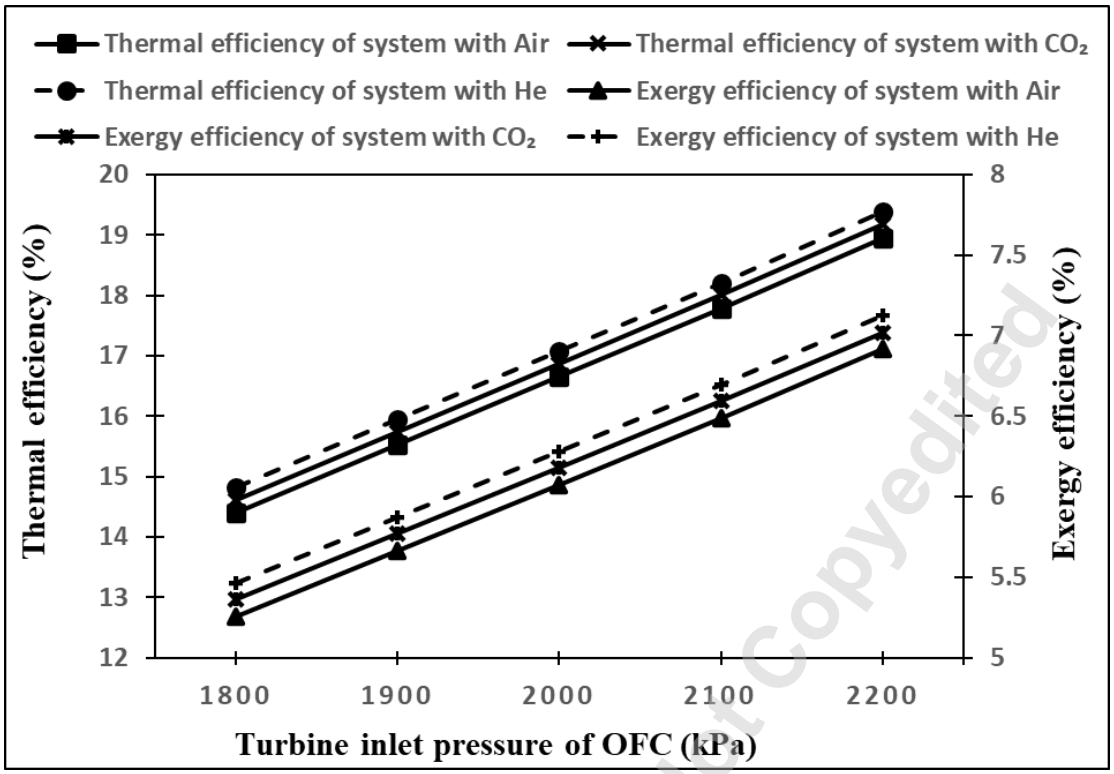
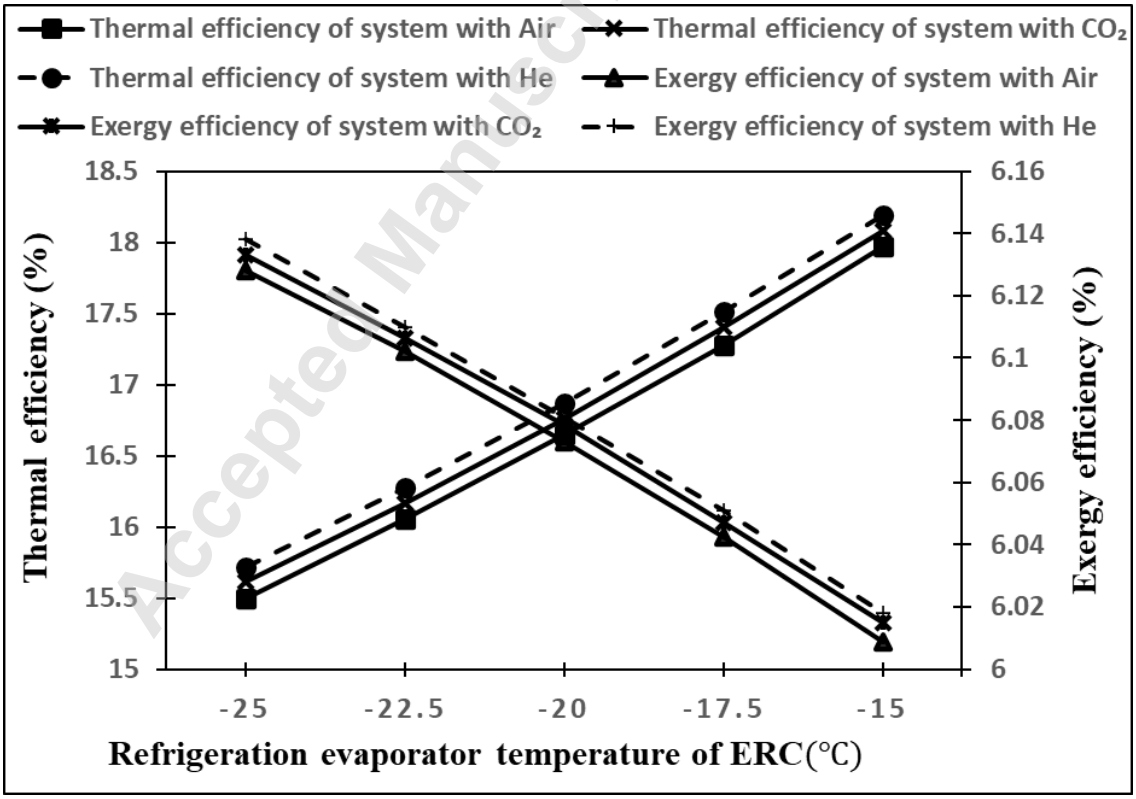


Fig. 8. Variation of thermal and exergy efficiencies of the cogeneration system for different SHTF with the change in turbine inlet pressure of OFC



Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

Fig. 9. Variation of thermal and exergy efficiencies of the cogeneration system for different SHTF with the change in refrigeration evaporator temperature of ERC

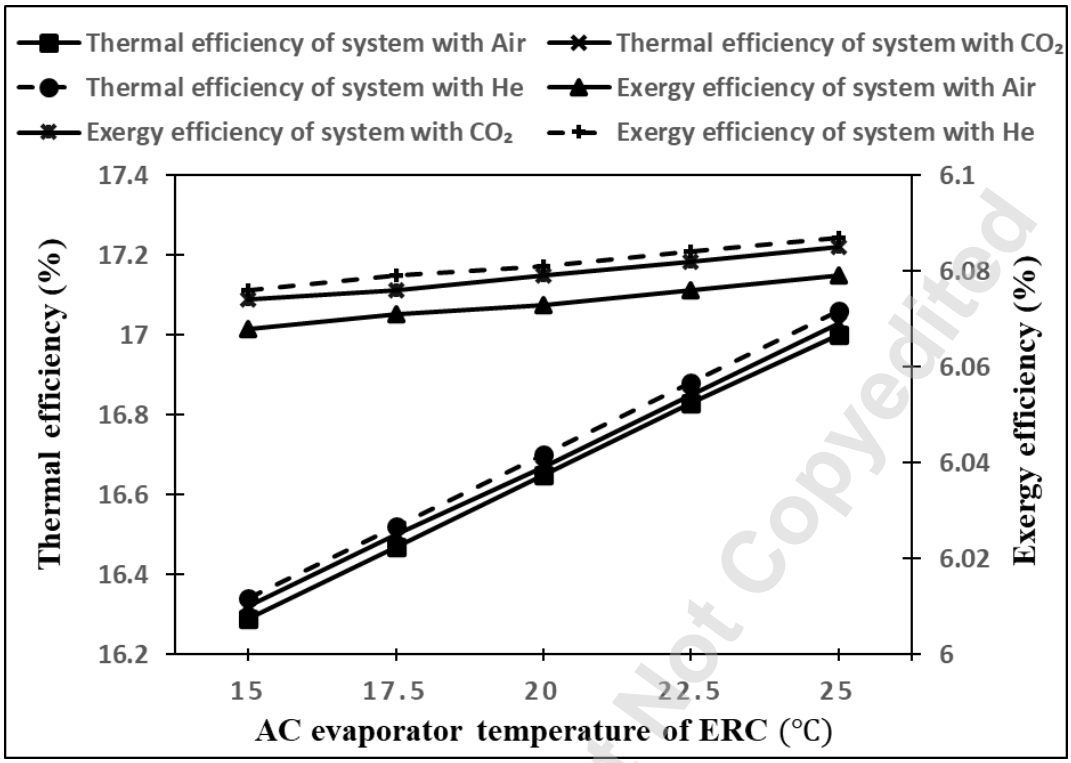
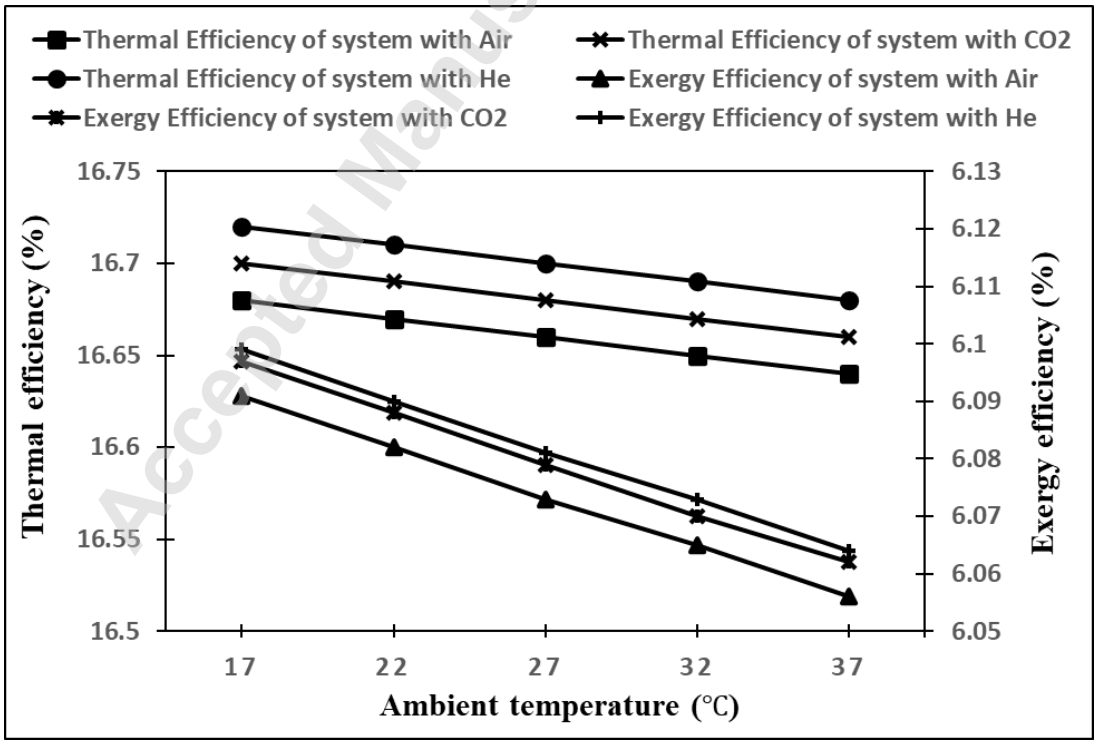


Fig. 10. Variation of thermal and exergy efficiencies of the cogeneration system for different SHTF with the change in AC evaporator temperature of ERC



Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

Fig. 11. Variation of thermal and exergy efficiencies of the cogeneration system for different SHTF with the change in Ambient temperature

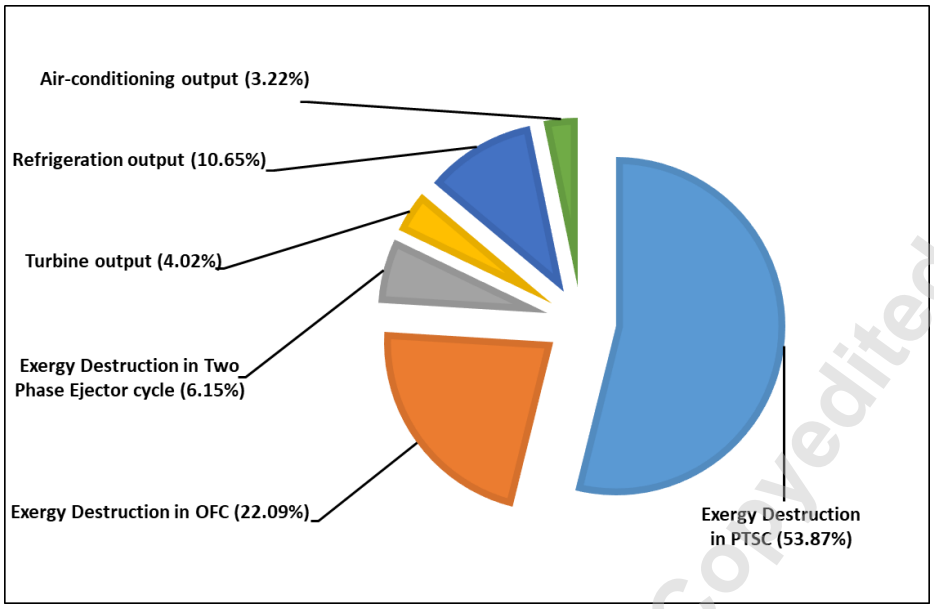


Fig. 12. Percentage distribution of exergy input into exergy destruction and exergy output of the proposed cogeneration system

Table 1. Thermal properties of SHTF at mean temperature (377 ℃)[34]

Solar heat transfer fluid (SHTF)	**Dynamic viscosity μ (Pa s)**	**Specific heat, C_p (J kg^{-1} K^{-1})**	**Density, ρ (kg m^{-3})**	**Thermal conductivity, k (W m^{-1} K^{-1})**
Air	$3.5x10^{-5}$	1086	0.467	0.054
CO_2	$3.3x10^{-5}$	1145	0.713	0.051
He	$3.3x10^{-5}$	5193	0.076	0.251

Table 2. Energy balance equations for the components of the proposed cogeneration system

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

Component	Balance Equations
Superheater (SH)	$\dot{m}_{18}(h_{18} - h_{19}) = \dot{m}_{22}(h_{22} - h_{32})$
Evaporator (EVA)	$\dot{m}_{19}(h_{19} - h_{20}) = \dot{m}_{32}(h_{32} - h_{31})$
Economizer (ECO)	$\dot{m}_{20}(h_{20} - h_{16}) = \dot{m}_{29}(h_{30} - h_{29})$
Solar pump1(P1)	$\dot{W}_{P1} = \dot{m}_{16}(h_{17} - h_{16})$
Turbine(T)	$\dot{W}_T = \dot{m}_{22}\dot{h_{22}} + m_{35}h_{35} + \dot{m}_{37}h_{37} - \dot{m}_{23}h_{23}$
High Pressure Flasher (HPF)	$\dot{m}_{33a}\dot{h_{33a}} = m_{35}h_{35} + \dot{m}_{34}h_{34}$
Low Pressure Flasher (LPF)	$\dot{m}_{34a}\dot{h_{34a}} = m_{36}h_{36} + \dot{m}_{37}h_{37}$
Expansion valve1(EV1)	$h_{33} = h_{33a}$
Expansion valve2 (EV2)	$h_{34} = h_{34a}$
Regenerator	$\dot{m}_{23}(h_{23} - h_{24}) = \dot{m}_{26}(h_{27} - h_{26})$
Condenser1(Cond1)	$\dot{Q}_{Cond1} = \dot{m}_{24}(h_{24} - h_{25})$
Pump2 (P2)	$\dot{W}_{P2} = \dot{m}_{25}(h_{26} - h_{25})$
Pump3 (P3)	$\dot{W}_{P3} = \dot{m}_{28}(h_{29} - h_{28})$
Compressor (C)	$\dot{W}_C = \dot{m}_1(h_1 - h_{11})$
Condenser2 (Cond2)	$\dot{Q}_{Cond2} = \dot{m}_1(h_1 - h_2)$
Heat exchanger (HX)	$\dot{Q}_{HX} = \dot{m}_2(h_2 - h_3) = \dot{m}_{11}(h_{11} - h_{10})$
Expansion valve3 (EV3)	$h_3 = h_4$
Air-conditioner (AC)	$\dot{Q}_{AC} = \dot{m}_5(h_5 - h_4)$
Ejector (Eje)	$\dot{m}_5h_5 + \dot{m}_9h_9 = \dot{m_6}h_6$
Separator (Sep)	$\dot{m}_7h_7 + \dot{m}_{10}h_{10} = \dot{m_6}h_6$
Expansion valve 4 (EV4)	$h_7 = h_8$
Refrigerator (ref)	$\dot{Q}_{ref} = \dot{m}_9(h_9 - h_8)$

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

Table 3. Exergy destruction equations for the components of the proposed cogeneration system

Component	**Balance Equations**
Super Heater (SH)	$\dot{E}_{x,D,SH} = \dot{m}_{18}(e_{x,18} - e_{x,19}) + \dot{m}_{22}(e_{x,32} - e_{x,22})$
Evaporator (EVA)	$\dot{E}_{x,D,EVA} = \dot{m}_{19}(e_{x,19} - e_{x,20}) + \dot{m}_{32}(e_{x,31} - e_{x,32})$
Economizer (ECO)	$\dot{E}_{x,D,ECO} = \dot{m}_{20}(e_{x,20} - e_{x,16}) + \dot{m}_{29}(e_{x,29} - e_{x,30})$
Solar pump (P1)	$\dot{E}_{x,D,P1} = \dot{m}_{16}(e_{x,16} - e_{x,17}) + \dot{W}_{P1}$
Turbine(T)	$\dot{E}_{x,D,T} = \dot{m}_{22}e_{x,22} + \dot{m}_{35}e_{x,35} + \dot{m}_{37}e_{x,37} - \dot{m}_{23}e_{x,23} - \dot{W}_{T}$
High Pressure Flasher (HPF)	$\dot{E}_{x,D,HPF} = \dot{m}_{33a}e_{x,33a} - \dot{m}_{35}e_{x,35} - \dot{m}_{34}e_{x,34}$
Low Pressure Flasher (LPF)	$\dot{E}_{x,D,LPF} = \dot{m}_{34a}e_{x,34a} - \dot{m}_{36}e_{x,36} - \dot{m}_{37}e_{x,37}$
Expansion valve1(EV1)	$\dot{E}_{x,D,EV1} = \dot{m}_{33}(e_{x,33} - e_{x,33a})$
Expansion valve2(EV2)	$\dot{E}_{x,D,EV2} = \dot{m}_{34}(e_{x,34} - e_{x,34a})$
Regenerator	$\dot{E}_{x,D,Reg} = \dot{m}_{23}(e_{x,23} - e_{x,24}) + \dot{m}_{26}(e_{x,26} - e_{x,27})$
Condenser1(Cond1)	$\dot{E}_{x,D,Cond1} = \dot{m}_{24}(e_{x,24} - e_{x,25}) - \dot{Q}_{Cond}\left(1 - \frac{T_0}{T_{Cond1}}\right)$
Pump2(P2)	$\dot{E}_{x,D,P2} = \dot{m}_{25}(e_{x,25} - e_{x,26}) + \dot{W}_{P2}$
Pump3(P3)	$\dot{E}_{x,D,P2} = \dot{m}_{28}(e_{x,28} - e_{x,29}) + \dot{W}_{P3}$
Compressor (C)	$\dot{E}_{x,D,C} = \dot{m}_{1}(e_{x,20} - e_{x,10}) + \dot{W}_{C2}$
Condenser2 (Cond2)	$\dot{E}_{x,D,Cond2} = \dot{m}_{1}(e_{x,1} - e_{x,2}) - \dot{Q}_{Cond}\left(1 - \frac{T_0}{T_{Cond2}}\right)$
Heat exchanger (HX)	$\dot{E}_{x,D,HX} = \dot{m}_{2}(e_{x,2} - e_{x,3}) + \dot{m}_{10}(e_{x,10} - e_{x,11})$
Expansion valve3 (EV3)	$\dot{E}_{x,D,EV3} = \dot{m}_{3}(e_{x,3} - e_{x,4})$
Air-conditioner (AC)	$\dot{E}_{x,D,AC} = \dot{m}_{4}(e_{x,4} - e_{x,5}) - \dot{Q}_{AC}\left(\frac{T_0}{T_{AC}} - 1\right)$
Ejector (Eje)	$\dot{E}_{x,D,Eje} = \dot{m}_{5}e_{x,5} + \dot{m}_{9}e_{x,9} - \dot{m}_{6}e_{x,6}$
Separator (Sep)	$\dot{E}_{x,D,Sep} = \dot{m}_{6}e_{x,6} - \dot{m}_{7}e_{x,7} - \dot{m}_{10}e_{x,10}$

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

Expansion valve 4 (EV4)	$\dot{E}_{x,D,EV4} = \dot{m}_7(e_{x,7} - e_{x,8})$
Refrigerator (ref)	$\dot{E}_{x,D,ref} = \dot{m}_8(e_{x,8} - e_{x,9}) - \dot{Q}_{ref}\left(\frac{T_0}{T_{ref}} - 1\right)$

Table 4. Input data for modeling of the proposed cogeneration system

Solar heat transfer fluids	He, Air, and CO_2
Working fluid of OFC cycle	R124
Working fluid of the two-phase ejector cooling cycle	R600a
Ambient temperature (^{0}C)	25
Ambient pressure (kPa)	101.32
Effectiveness of heat exchanger	0.8
Inlet temperature of SHTF (^{0}C) [at point 17]	377
Solar direct beam irradiation (G_b) (W/m^2)	800
Concentration ratio of collector	26.37
Aperture area (m^2)	69.6
Sun temperature (K)	5770
Width (m)	5.8
Diameter of inner receiver (mm)	66
Diameter of outer receiver (mm)	70
Diameter of inner cover (mm)	120
Diameter of outer cover (mm)	125
Optical efficiency (%)	80
Tube length (m)	12
Nozzle efficiency (%)	85-95
Refrigeration evaporator Temperature(℃)	(-15) - (-25)
Condenser Temperature (℃)	30-40
Mixing chamber efficiency (%)	90
Diffuser efficiency (%)	85
AC evaporator Temperature (℃)	15-25

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Journal of Thermal Science and Engineering Applications. Received December 15, 2024;
Accepted manuscript posted April 1, 2025. doi:10.1115/1.4068339
Copyright © 2025 by ASME

Table 5: Comparison of the computed results with the reported results of literature [26, 31]

	Study/value	Literature/value
OFC turbine output kW	Present	Varma and Srinivas [26]
	1292	1248
Exergy efficiency of OFC	Present	Varma and Srinivas [26]
	26.84%	26.21%
Cooling outputs of two-phase ejector (Refrigeration output kW)	Present	Lawrence and Elbel [31]
	1565	1558
Cooling outputs of two-phase ejector (Air conditioning output kW)	Present	Lawrence and Elbel [31]
	515	504

Accepted Manuscript Not Copyedited

Downloaded from http://asmedigitalcollection.asme.org/thermalscienceapplication/article-pdf/doi/10.1115/1.4068339/7456566/tsea-24-1619.pdf by Delhi Technological University user on 08 April 2025

Check for updates

Concurrency and Computation: Practice and Experience

WILEY

RESEARCH ARTICLE

Ranking Influential Nodes in Social Networks Based on the Game of Thieves Algorithm

Minni Jain | Aditya Gupta | Aaryan Arora | Aayush Patel

Delhi Technological University, Delhi, India

Correspondence: Minni Jain (minnijain@dtu.ac.in)

Received: 4 September 2024 | **Revised:** 11 February 2025 | **Accepted:** 24 February 2025

Keywords: centrality | complex networks | game of thieves | node removal procedure (NRP) | online social networks | susceptible-infected-recovered (SIR) model

ABSTRACT

With advancements in technology, the number of people using various social networks to share information daily has increased exponentially. To spread information to maximum users, there is a need to effectively identify highly influential nodes in social networks. Still, existing centrality measures and methods have drawbacks relating to computational complexity and accuracy. The proposed method considers the Game of Thieves algorithm to rank influential nodes in social networks effectively with lower computational time that outperforms existing closeness centrality, eigenvector, Game Theory, and VoteRank^{++} algorithms. To evaluate the performance of the proposed work, experiments were conducted on synthetic and real-world network datasets. The results were evaluated with the help of the Susceptible-Infected-Recovered (SIR) model and Node Removal Procedure (NRP) evaluation metrics. These results showed that GOT outperforms the existing measures by 12% on average while giving more accurate results than the standard algorithms for large-scale networks.

1 | Introduction

In recent decades, social networking has become a crucial medium for connecting with people and disseminating information with the help of technology. It plays an important role in modern society, allowing people to share information. In recent years, Social Networking platforms like X, LinkedIn, Facebook, Instagram, and so on have grown exponentially. Currently, there are 4.95 billion users[1] globally on over 130 social media platforms, which are expected to reach 5.85 billion by 2027. A social network can be described as a group of nodes (people, groups, organizations, etc.) that is connected by links/edges showing relations between them. It is widely applicable in areas such as organizational studies, psychology, economics, information science, and criminology [1] leading to additional challenges for the application of social network analysis (SNA) methods.

SNA focuses on the exchange of resources between individuals, groups, or organizations. Many complex systems, like social networks, utilize SNA to identify key or influential individuals. These individuals can be thought of as "hubs" within the network, and their actions can significantly influence the behavior of the entire network. Thus, it is possible to track the spread of information by determining the key players in the dissemination of messages. This can be useful in understanding how rumors spread, allowing for the potential to intervene and control the spread of misinformation [2].

The existing approaches used to find these influential nodes include centrality measures, like degree, closeness, and betweenness centrality measures [3]. These measures assess the importance of a node based on its connections to other nodes in the network. Degree centrality measures the number of neighbors of

© 2025 John Wiley & Sons Ltd.

a node but fails to accurately measure the centrality of a node. A node having a higher number of connections but lying in the periphery of the network would have a higher degree as compared to a more central node with a lesser number of connections [4]. Closeness centrality is based on the distance between each pair of nodes in the network, while betweenness centrality is based on the frequency with which a node acts as a bridge between other nodes. Betweenness and closeness centralities give better results than degree centrality, but high computational complexity limits their application in large-scale social networks [3].

K-shell algorithms [5] strip away the outer layers of a network, with the core nodes considered more influential. However, in a significant drawback, they may assign the same influence value to many nodes in some special networks.

The above-discussed methods assume that a node's influence is contingent only on its own importance. Information entropy [6] considers the impact of neighbors' importance on a node's influence. Popular random walk models like PageRank [5] and LeaderRank [7] give better results for finding influential nodes in directed networks but fail to do so for undirected networks. Recent approaches using game theory [8] model social interactions as cooperative games to detect dominant influencers in the networks.

While there have been many efforts to identify influential nodes in social networks, the scope for improvement exists in terms of accuracy and efficiency. Further research is needed to address the limitations of existing methods and to develop new approaches to better determine the influence of nodes in real-world networks. This study will explore the use of the Game of Thieves (GOT) algorithm for identifying influential nodes in a network and evaluate its performance through experiments on real-world and synthetic datasets.

The GOT algorithm is a new approach for analyzing complex networks, including identifying the influential nodes in networks. This algorithm relies on the notion of a "diffusion process," in which a certain behavior or characteristic spreads through the network from one individual to another. The GOT algorithm simulates this diffusion process and uses it to determine which individuals are most likely to start or stop the spread of the behavior. It considers the strategic behavior of the nodes during the influence spread, and it can identify influential nodes more realistically [3].

The main challenge with earlier proposed methods was the greater amount of computation cost associated with these methods. GOT can compute the importance of both vertices and edges in a network in polylogarithmic time as opposed to the quadratic time taken by classical methods. The worst-case implementation of GOT is limited by $O(\log^3 |V|)$ where V is the number of nodes in a network. The lower computational complexity of GOT makes it much faster and more scalable (both in the number of nodes and edges/connections) than already existing methods [3].

The evaluation results suggest that GOT outperforms other centrality algorithms based on computational complexity by an average of 12% and provides more accurate results for large-scale networks.

The key contributions of this research work are as follows:

1. This paper introduces a comprehensive evaluation of the GOT algorithm for finding influential nodes in social networks.
2. The algorithm identifies influential nodes that have a significant impact on the network structure and properties when removed from the network. These nodes can represent individuals who have many connections within the network or bridge different subgroups within the network. Once influential nodes have been identified, they can be targeted for interventions to leverage their influence within the network.
3. In terms of computational complexity, it outperforms earlier proposed methods to find nodes that drive the dissipation of information, opinions, or behaviors within the network.
4. This study makes several significant contributions to the field of network analysis and provides an effective algorithm for identifying the influential nodes in a network.

The remainder of this paper is organized as follows. Section 2 presents various studies conducted in the field of finding influential nodes in a complex network. Section 3 describes the preliminary concepts that need to be understood for a better understanding of the proposed work. Section 4 details the proposed methodology and datasets used. The algorithmic explanation and flowchart along with its illustration for a network are described in Section 4. In Section 5, the experimental setup used is explained to compute the results. The results of the analysis are discussed in Section 6. Finally, concluding remarks and the scope of future work are discussed in Section 7.

2 | Related Work

The identification of influential nodes in complex networks is a critical area of research in SNA. A wide range of centrality measures has been proposed to quantify node importance, reflecting their position and connections within the network. These measures aim to determine the nodes that have the most significant influence, either directly or indirectly, on the network's structure and dynamics.

Traditional centrality measures: Common centrality measures, such as degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality [2], are widely used to identify influential nodes. Degree centrality quantifies the importance of nodes based on the number of direct connections, while betweenness centrality highlights nodes that serve as bridges in the network. Closeness centrality, on the other hand, focuses on nodes that can quickly access other nodes, and eigenvector centrality considers the importance of neighbors when evaluating a node's influence. These traditional centrality measures are foundational but have limitations, particularly in dynamic networks where the influence of nodes may change over time.

Cluster-based and heuristic approaches: The clustered local degree centrality measure improves on traditional centrality measures by considering local clustering around a node and the

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70044 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

degrees of its neighbors [9]. This approach identifies influential nodes based on the local network structure rather than global network properties. Meanwhile, heuristic clustering methods group nodes into clusters based on similarity, with the most influential node in each cluster selected as a potential seed node. Although these methods are computationally efficient, they often exhibit low accuracy, as they are prone to getting trapped in locally optimal solutions [10].

Limitations of traditional methods and diffusion models: Traditional centrality measures often fail to capture the complexity of diffusion processes within networks. To address this, more sophisticated methods, such as the PageRank algorithm, are used for influence maximization problems. The PageRank algorithm assesses the importance of nodes by considering both the quantity and quality of their connections, reflecting the spreading of influence through the network. However, these methods still fall short in many scenarios, prompting the development of alternative approaches.

Optimization and simulation-based approaches: One notable advancement is the use of optimization techniques, such as the gray wolf optimization algorithm, proposed by Zareie [11]. This method formulates the influence maximization problem as an optimization challenge, taking into account the influence of nodes and the distances between them. Compared to earlier heuristic approaches, the gray wolf optimization algorithm showed improved performance and reduced computational time.

Simulation-based methods also provide valuable insights into node influence. The Linear Threshold Model (LTM) [12], for instance, simulates the spread of information across a network to identify influential nodes by analyzing their potential to initiate or stop the spread. While these models are effective, they are sensitive to parameter choices and may not always accurately represent real-world dynamics.

VoteRank^{++} and game theory approaches: Recent advancements in influence maximization have introduced algorithms like VoteRank^{++} [12], which iteratively select influential nodes based on varying voting abilities. In this method, nodes with different degrees have different voting powers, allowing for more accurate identification of important nodes in the network.

Game theory has also been increasingly applied to influence maximization problems. The explosion-trust (ET) Game Model [13] incorporates two critical factors—the degree of rumor explosion and the degree of trust in the source node—to develop a model for rumor spreading. This approach, along with the EDGly algorithm and SV-based games for community-based networks [14], has shown promise in finding influential nodes within specific communities.

Hybrid and deep learning approaches: While individual centrality measures are effective in specific contexts, they often lack the capacity for comprehensive evaluation. Therefore, combining multiple centrality measures has become a popular approach. Random walk algorithms [7] integrate these measures through preference analysis, ranking influential nodes more robustly.

Furthermore, hybrid approaches that combine classical methods with modern machine learning techniques have gained attention. For example, deep learning methods that utilize network embedding and deep reinforcement learning have been proposed to address the influence maximization problem, offering more efficient and accurate results. Graph embedding techniques help overcome the limitations of traditional algorithms by providing a more nuanced representation of nodes and their relationships.

Several modified approaches have also been proposed, such as the modified spider monkey optimization-based method, which incorporates graph embedding and the k-means algorithm to improve the detection of influential nodes in graph networks [15].

Community-aware and multicomponent centrality approaches: Recent work has increasingly focused on community-aware centrality approaches and multicomponent centrality metrics [16–26]. These methods recognize that nodes within communities may have different roles and influences depending on their membership in specific subgroups. By integrating community structure into the centrality measures, these approaches offer a more granular and context-aware assessment of node influence in complex networks.

Within the research community, one can see that considerable work has already been done to find the influential nodes in complex networks. Despite the advances, there are still some drawbacks to the current study. For example, centrality measures can be unreliable in capturing the true influence of a node in a network because they depend on limited information and can be influenced by the network structure. In addition, game-theoretic models can be computationally complex and may not accurately reflect the dynamics of real-world networks.

The GOT algorithm is a recent addition to the literature that addresses these limitations. This research builds upon the existing literature by applying the GOT algorithm to identify influential nodes in a network and comparing its performance to traditional centrality measures and simulation-based methods using real-world and synthetic datasets. This paper presents the fundamental ideas and principles of the GOT approach and how it identifies the influential node in an optimal manner for a social network.

3 | Preliminaries

This section includes background knowledge of the concepts required for the better understanding of the proposed work.

3.1 | GOT

The algorithm analyzes the interaction between a set of "virtual diamonds" (vdiamonds) and a group of wandering "thieves" (moving across the network in search for vdiamonds) in a graphical network.

At the start, each node is endowed with wandering thieves and vdiamonds. A thief has two states: "loaded" (it carries one vdiamond) and "empty" (it does not carry any vdiamond). A thief

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70044 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

moves from one node to a randomly chosen neighboring node to look for vdiamonds; when he finds a vdiamond, the thief brings it back to his home node by retracing the previous path and goes in search of another vdiamond. This vdiamond becomes immediately available for the other wandering thieves to steal.

Initially, all node hosts have the same given number of thieves and vdiamonds [3]. The algorithm moves in epochs. At each epoch, all thieves move from their current node to the next one and change state when they steal a new vdiamond from other nodes.

Consider an undirected graph, $G = (V, E)$, where (V) is the set of vertices/nodes, and (E) is the set of edges between nodes.

At the first epoch; initial state:

$$\upsilon_{0n} : \text{number of vdiamonds at node } n, n \in V$$

As the game proceeds, at epoch T, the state of the graph can be described as:

υ_{Tn}: number of vdiamonds at node $n, n \in V$.

α_{Tn} : number of loaded thieves crossing link, $e \in E$.

So, the average number of vdiamonds, υ_n, present at node n, $n \in V$ and average number of loaded thieves, α_e, traversing link e, $e \in E$ after the game runs for T epochs can be calculated as:

$$\upsilon_n = \frac{1}{T}\sum_{i=0}^{T}\upsilon_{in} \tag{1}$$

$$\alpha_e = \frac{1}{T}\sum_{i=0}^{T}\alpha_{ie} \tag{2}$$

At the end of all the game runs, the node with lower value of υ_n is considered more influential than a node with higher value of υ_n.

3.2 | Information Diffusion Models

The information diffusion models are mathematical frameworks widely used for evaluating how information and behaviors spread among individuals in SNA. With the help of these models, the dynamics of information flow and the factors that influence the rate and extent of diffusion in social networks can be studied. To validate the efficacy of the suggested method, the node removal procedure (NRP) and the susceptible infected recovered (SIR) models have been used to assess the rationality of the node ranking results.

3.2.1 | SIR

The SIR model is the most popular diffusion model used in SNA to study the propagation of infectious diseases, rumors, and information in a network. It has been applied to various real-world scenarios, including the spread of HIV, COVID-19, and SARS [27].

This model takes the entire disease-spread population to estimate the number of individuals infected in a period and predicts the growth of the disease.

The model divides the individuals/nodes in a network into one of three discrete states: susceptible (S), infected (I), and recovered (R).

The state variables of each state can be defined as:

$S(t)$ = number of individuals who are not infected yet at time t

$I(t)$ = number of infected individuals who transit disease at time t

$R(t) = n$umber of individuals recovered at time t

The set of differential equations that describe the dynamics of disease transmission and recovery in a network can be expressed as follows:

$$\frac{dSt}{dt} = -\beta StIt \tag{3}$$

$$\frac{dIt}{dt} = \beta StIt - \gamma It \tag{4}$$

$$\frac{dRt}{dt} = \gamma It \tag{5}$$

In the model, at time $t = 0$, few nodes are in I state and act as disease spreaders in the network. The rest nodes are in state, S. The rate at which nodes get infected can be defined as transmission coefficient, β, and γ denotes the recovery rate of infected nodes to become immune to the disease. There has also been an assumption for recovered nodes that they will never be reinfected again. The variations in the susceptible (S), infected (I), and recovered (R) population can be studied over the course of the infection. Diffusion dynamics play a significant role in determining the spread of influence in networks, and the SIR model provides a relevant framework for studying this phenomenon.

3.2.2 | NRP

NRP is a widely accepted technique for determining the accuracy of identified influential nodes in a network. It is a non-compartmental diffusion model [28] used in SNA to simulate the spread of information and opinions in a network. The model includes a node removal process, which removes nodes from the network based on certain criteria, such as their centrality.

This plays a pivotal role in the experimental process. Based on the scores assigned by the metric, each node in the graph is arranged in sorted order. The sorting process begins with the most critical/influential node and proceeds down to the least significant node. Nodes are gradually removed from the graph, one by one, following the order established during the sorting step [29]. As

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70044 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

each component is removed, two crucial aspects of the remaining graph are carefully measured: the size of the Giant Component (GC) and the number of connected components (CC).

The primary objective in NRP is to identify a metric that demonstrates desirable characteristics. A metric is considered superior if it produces a greater number of connected components while simultaneously minimizing the size of the giant component. By maintaining this delicate balance, the metric remains effective throughout the NRP process.

4 | Proposed Methodology

4.1 | Brief Description of GOT

As discussed in Section 3.1, GOT identifies influential nodes in a complex network by employing the idea of a diffusion process. The drivers of this algorithm are "thieves," with one thief present at each node. Each node of the network is considered to have a predetermined quantity of virtual diamonds or "vdiamonds" present initially.

The strategic behavior of the thieves to collect the vdiamonds is used to find the influential nodes in the network. Each thief has an associated state which can be "empty," if the thief is not carrying a vdiamond or "loaded," if the thief is carrying a vdiamond. A thief in "empty" state picks a vdiamond from a node and deposits it at its home node. The influence of each node is measured by the average of the number of vdiamonds at that node. The nodes with lower averages are considered more influential in comparison to nodes with higher average numbers of vdiamonds. A lower average signifies that a greater number of thieves have frequented a particular node, and therefore, this node can exert its influence on the home nodes of those thieves.

The GOT algorithm can also be used for finding the influence of edges by keeping a track of the average number of thieves in a loaded state passing through the edge. The edge with a higher average is considered more influential than the edge with a lower average. For the purpose of this study, the influence of edges is not considered in the network. During initialization, random initialization is chosen to simulate a variety of starting conditions or to ensure that results are not biased by a fixed starting configuration.

4.2 | Example Network

UlFigure 1 illustrates the functioning of GOT on an example network of eight nodes, labeled A–H. The label for each figure in the illustration represents the influence of the nodes at the mentioned epoch. The illustration shows the behavior of thieves in the network at certain intervals between epochs 1 and 1000. The first epoch represents the initial state of the network. Each node has two thieves and five vdiamonds at epoch 1 in the illustration above.

At each epoch, the network follows the GOT algorithm once, and the number of vdiamonds gets updated. The mean of the diamonds for each node after each epoch is calculated. The lowest mean value represents the most influential node. In order to determine the rank, the nodes are arranged in ascending order using the mean value of the vdiamonds. The labels alongside the nodes represent the ranks after that number of epochs; that is, the

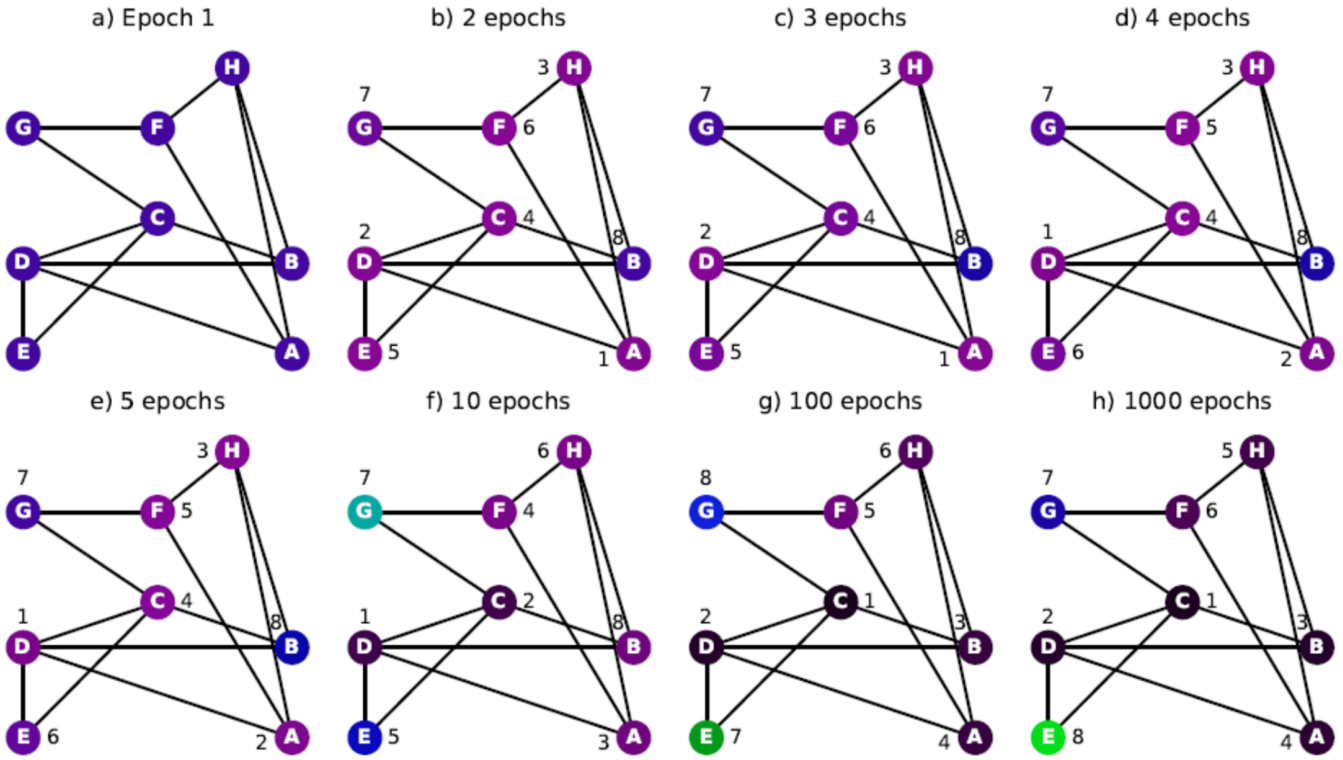


FIGURE 1 | Running GOT on a sample network.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70044 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

rank 1 node is the most influential node in the network, and the rank 8 node is the least influential.

4.3 | Algorithm Description

Based on the technical framework defined in Section 4.1 and the example covered in Section 4.2, this section formulates the algorithm for the proposed approach. For a proper understanding of the method, some notations are defined as follows:

- n_u^t is the number of vdiamonds at vertex $u \in N(G)$ at epoch t
- h_x is the home node of thief x
- c_x is the current location of thief x
- s_x is the current state of thief x representing either "empty" or "loaded" state
- $\prod_u$ is the set of all nodes directly connected to node u by an edge
- Δ_x is the path followed by thief x to reach the current node from its home node
- ϕ_u is the average number of vdiamonds at each node u
- Ω_u is the total number of vdiamonds at each node u
- T_f is the number of epochs for which the algorithm is run on the network G (Algorithm 1).

ALGORITHM 1

Input: Target Network, $G(N, E)$
Output: Average number of vdiamonds, ϕ_u, at each node $u \in N(G)$

```
1 for i in range (1, T_f):
2   for each thief j in G:
3     if s_j is empty then:
4       set c_j = random(u) where u ∈ j
5       if c_j ∈ j then:
6         update Δ_j to include path till first occurrence of c_j
7       else
8         append c_j to Δ_j
9       if n^i_{c_j} >  0 then:
10        set s_j = loaded and update n^i_{c_j} := n^i_{c_j} − 1
11    else
12      if c_j= h_j then:
13        set s_j= empty and update n^i_{c_j} := n^i_{c_j} + 1
14      else
15        set c_j to the last but one location in Δ_j and remove
          the last node from Δ_j
16  end for
17 for each node v in G:
18   Ω_v := Ω_u + n^i_v
19 end for
20 end for
21 for each node v in G
22   set φ_v = Ω_v/T_f
23 end for
24 Return u for each node u ∈ N(G)
```

A brief description of the algorithm is provided below:

Step 1: This step covers the lines 1–2 of the algorithm. In this step, all the thieves in the graph are iterated over for a total of T_f epochs.

Step 2: This step describes the lines 3–8 of the algorithm. In this step, the next location is decided for a thief j if the current state of the thief, s_j is "empty." A node, $u \in \prod_u$, is chosen as the next location for thief j based on the weighted probability of the edges given by:

$$p\left(e_v^u\right) = \frac{n_u^t}{\sum_{u \in \prod_u} w_v^x} \tag{6}$$

where e_v^u is an edge between nodes v and u, w_v^x is the weight of edge between nodes v and u and $p\left(e_v^u\right)$ is the probability of selecting e_v^u.

The dynamic list Δ_j contains the nodes previously visited by the thief. If the thief has already visited the next location c_j previously then, Δ_j is modified such that it contains the nodes up to the first occurrence of cj. If the thief has not visited c_j previously then, c_j is appended to Δ_j.

Step 3: In this step, the new location of the thief in "empty" state is checked for vdiamonds. If c_j has vdiamonds, the thief takes one vdiamond and its state changes to 'loaded' and the number of vdiamonds at this node is updated. This step concludes the actions taken by a thief in an initially "empty" state for an epoch. This step covers lines 9 and 10 of the algorithm.

Step 4: This step covers lines 11–13 of the algorithm. In this step, c_j of a thief j in the "loaded" state initially is considered. If the thief is at its home node, it drops its held vdiamond at this node and its state is set to "empty" with the number of vdiamonds at the home node of the thief increasing by 1.

Step 5: This step covers lines 14–16 of the algorithm. If the thief carrying vdiamond is not at its home node, its next location is the last but one node in Δ_j. Then c_j is updated and the last node is deleted from Δ_j. This step concludes the series of actions taken by a thief in initially "loaded" state for an epoch.

Step 6: This step illustrates lines 17–19 of the algorithm. All the nodes in the network are iterated over to calculate the number of the vdiamonds at each node. For each node, n_v^i is added to Ω_v.

Step 7: This step covers lines 21–23 of the algorithm. In this step, ϕ_v is calculated by dividing Ω_v by the number of epochs T_f.

4.4 | Flowchart

The flowchart in Figure 2 provides a step-by-step overview of the approach followed. The target network G is taken as the input for the algorithm. In the pre-processing phase, the network is

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70044 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

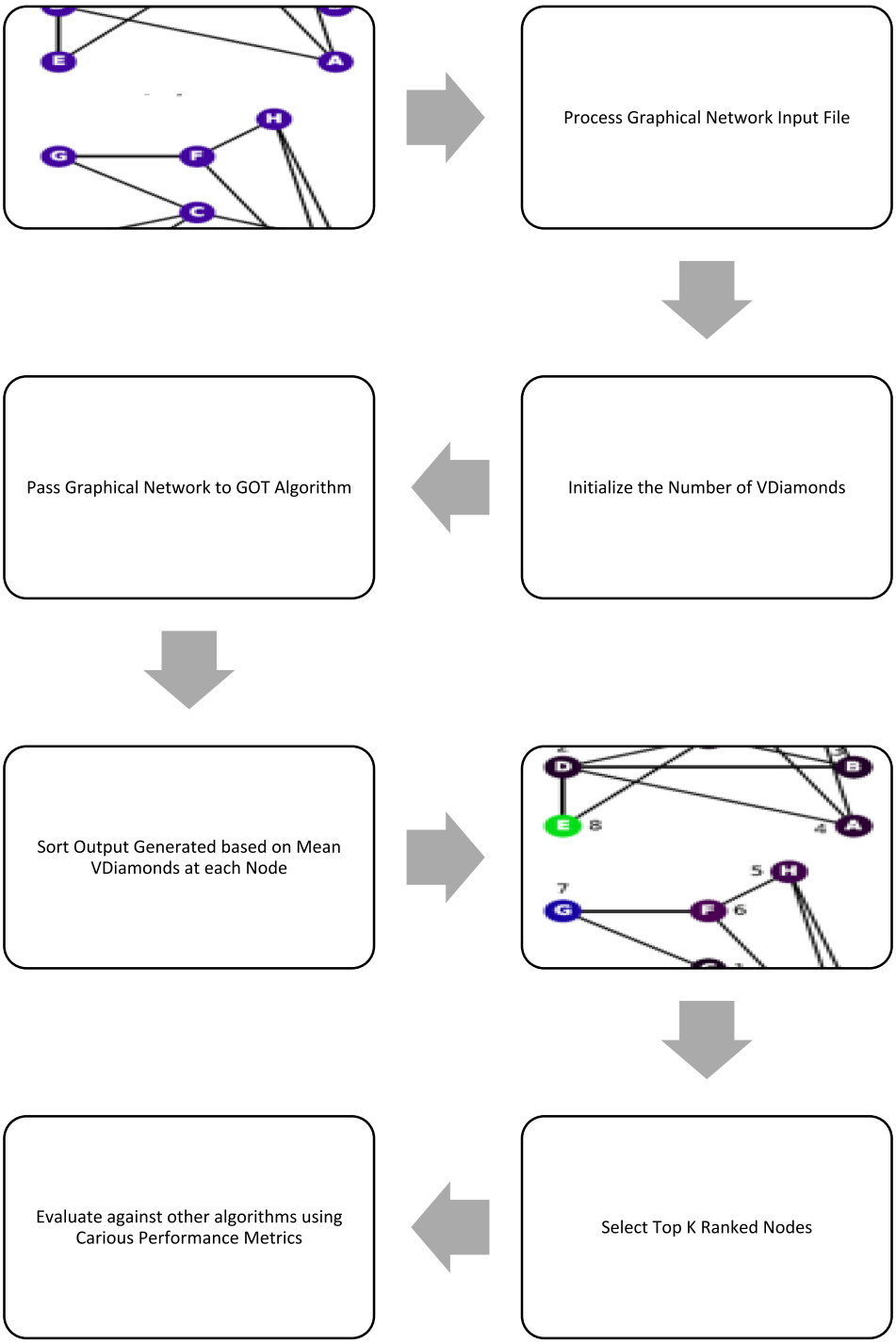


FIGURE 2 | Overview of the proposed work.

prepared for the algorithm by allocating an equal number of vdiamonds to each node initially. The GOT algorithm is applied on the network, and the output obtained is then sorted according to decreasing influence, that is, in increasing order of the average number of vdiamonds at nodes.

The next phase involves comparing the output of GOT with other approaches of finding influential nodes (such as centrality measures, etc.) and evaluating the performances using the different diffusion models as discussed in previous section (Section 3.2.1), respectively. The most influential k nodes are selected for evaluation, where k is a pre-decided number. For this research, k is taken as the number of nodes in the entire network. The results obtained from the diffusion models are represented as graphs and tables, which are further discussed in the results section.

5 | Experimental Setup

This section presents the experimental setup for the proposed work. GOT and the other measures, including Centrality, were

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70044 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

used to find the order of influence of nodes in a social network. The social networks used consisted of both synthetic and real-world datasets, as mentioned in Section 5.1. The datasets are further classified into weighted and non-weighted networks. Section 5.2 gives the details about the evaluation metrics used to compare the algorithms. The evaluation metrics considered in the study are: computational time analysis (Section 5.2.1), SIR (Section 5.2.2), and NRP (Section 5.2.3).

The implementation was done in Python and mainly using the NetworkX library for simulating the social networks. GOT was implemented by taking inspiration and ideas from other centrality-based algorithms that were used directly from the inbuilt implementations. VoteRank^{++} and game theory were taken from their respective sources [14] and re-implemented based on the requirements.

For the implementation, computation, and evaluation of results, a typical desktop PC was used (Intel i5 9600k with six cores and six threads with 16GB RAM on Windows 10 Operating System). The code was implemented on PyCharm IDE and run on Python 3.

5.1 | Datasets

This section describes the various online real-world social network datasets as well as synthetic datasets of varying contexts, sizes, and complexities used for the study.

The description of the real-world networks' datasets used is as follows:

1. Email-eu [30]: This dataset has information about all incoming and outgoing emails between members of the research institution. There is an edge (a, b) in the network if person a has sent person b at least one email.
2. Newman [31]: This dataset has a two-mode network due to two types of "nodes"—authors and papers—with connections existing between different node types.
3. Facebook Combined [32]: Facebook data has been made anonymous by adding a new value to each user's Facebook internal ID.
4. Congress-Bills [33]: In this dataset, US Congresspeople are the nodes, and the network represents the sponsors and co-sponsors of legislative bills.
5. US Power-Grid [34]: Transformers, substations, and generators serve as the nodes of this network of a high-voltage power grid, and high-voltage transmission lines serve as the ties.
6. Facebook Like [35]: The Facebook-like online social network represents an interesting two-mode network where weight can be assigned based on the number of messages that a user posts to a topic.
7. Wiki-Votes [36]: Administrators make up a very small portion of Wikipedia contributors. Users cast votes to advance a current user to the administrator position.
8. p2p Nutella network [37]: A series of images showing the topology of the Gnutella peer-to-peer file-sharing network, with nodes standing in for hosts and edges standing for connections between the hosts.
9. CA-Hep [37]: Citation-Author (CA) high energy physics (HEP) collaboration network is from the e-print that covers scientific collaborations between authors' papers submitted to this category.
10. Epinions [38]: This online social network is a who-trust-whom part of the general consumer review website Epinions.com. The statistical details of real-world networks used are shown in Table 1.

TABLE 1 | Statistical details of used real-world networks.

Real-world non-weighted networks				Real-world weighted networks			
Datasets	**Type**	**Nodes**	**Edges**	**Datasets**	**Type**	**Nodes**	**Edges**
Facebook Combined	Undirected	4039	88,234	Congress-Bills	Undirected	1718	424,932
Wiki-Votes	Directed	7115	103,689	Facebook Like	Directed	899	7046
p2p Nutella	Directed	10,876	39,994	Newman	Undirected	16,264	47,594
CA-Hep	Undirected	12,008	237,010	US PowerGrid	Directed	4941 5	6594
Email-eu	Directed	1005	16,706				

TABLE 2 | Statistical details of synthetic networks used.

Synthetic non-weighted networks				Synthetic weighted networks			
Datasets	**Type**	**Nodes**	**Edges**	**Datasets**	**Type**	**Nodes**	**Edges**
Edros Renyi	Undirected	10,000	499,950	Edros Renyi	Undirected	10,000	499,950
Watts Strogatz	Undirected	10,000	500,000	Watts Strogatz	Undirected	10,000	500,000
Barabasi Albert	Undirected	10,000	990,000	Barabasi Albert	Undirected	10,000	990,000

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70044 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The results were calculated for both non-weighted and weighted synthetic networks. The weights of the edges in weighted networks were randomly assigned, ranging from 1 to 10.

The statistical details of the synthetic networks used are shown in Table 2 and the description of the three synthetic networks used for the experiments is as follows:

1. Erdos Renyi [39]: A random graph $G_{n,p}$, generated by randomly connecting nodes (n) with edges (M), where each edge is included with a fixed probability (p), independent of other edges. The other parameter *seed* is set as a seed for random number generation. The number of edges can be calculated: $E = p \times {}^{n}_{2}C$.

   The graph can be represented as:

   $$G_{n,p} = erdos_renyi_graph(N = n, p = 0.15, seed = 7)$$

2. Barabasi Albert [40]: This models a random synthetic network $G_{n,m}$, where new nodes are added incrementally and connected to the nodes having a higher degree. The new node is connected to the node (i) with probability (p_i):

   $$p_i = \frac{k_i}{\sum_j k_j}, \text{ where } k_i \text{ is the degree of node } i.$$

   The total edges E can be found as follows:

   $$E = n \times m; \text{where } m = int(0.01 \times n) \text{ denotes the number of nodes in the initial graph.}$$

   The graph can be represented as:

   $$G_{n,m} = barabasi_albert_graph(N = n, m, seed = 7)$$

3. Watts Strogatz [41]: This model, $G_{n,k,p}$, interpolates between these two extremes by randomly rewiring some edges, k of a regular lattice. The number of edges in Watts–Strogatz network is given by:

   $$E = n \times k/2, \text{ where } k \text{ is mean degree.}$$

   The graph can be represented as:

   $$G_{n,k,p} = watts_strogatz_graph(N = n, k = int(0.01 * N), p = 0.4, seed = 7)$$

TABLE 3 | Computational complexity.

Algorithm	Computational complexity
Degree centrality [3]	$O(V)$
Page Rank [5]	$O(k * V)$
Betweenness centrality [3]	$O(V \|\| E)$
Closeness centrality [3]	$O(V^3)$
Game theory [11]	$O(k * V^2)$
VoteRank^{++} [12]	$O(k * V^2 * \log \|V)$
Eigen Vector [42]	$O(V^3)$
Game of Thieves (*GOT*) [1]	$O(log^2 \|V\|) < O(GOT) < O(log^3 \|V\|)$

5.2 | Evaluation Metrics

5.2.1 | Computational Time (C_t)

In comparing influence estimation algorithms, C_t is an important factor, as most applications require faster and more efficient algorithms. Computationally efficient algorithms can have practical applications in real-world problems involving large-scale, complex networks. Hence, a comparison was done based on the time required for different algorithms to identify influential nodes in each dataset. The algorithms considered for this purpose are degree centrality [3], Page Rank [5], betweenness centrality [3], closeness centrality [3], game theory [11], Eigen Vector [42] VoteRank^{++} [12], and GOT [1].

Table 3 shows the time complexity of the above-mentioned algorithms in terms of vertices, V, and edges, E. GOT's computational complexity is limited by $O(\log^2|V|) < O(\text{GOT}) < O(\log^3|V|)$. This is primarily because GOT has a fully distributed [5] design for the coverage of nodes, which means each node in the network can run parallelly without interfering with other node's computations. This distributed architecture allows it to analyze networks more quickly when traditional methods are inefficient.

TABLE 4 | Algorithm runtime (in seconds) for real networks.

Datasets	GOT	Betweenness	Closeness	Degree	Eigen vector	Page rank	VoteRank^{++}	Game theory
CA-Hep	333.246	834.299	976.332	0.006	4.457	4.837	182.235	544.078
Congress Bills	531.683	135.412	362.515	0.001	20.028	18.089	3.484	9614.25
Facebook-Combined	261.482	101.862	211.234	0.002	7.567	3.737	19.433	320.222
Newman	204.916	988.696	712.983	0.008	7.322	2.496	331.368	820.95
P2p Nutella	138.712	633.305	481.165	0.005	5.261	1.895	146.241	528.019
Wiki-votes	253.535	365.132	494.108	0.004	3.808	4.399	63.112	511.027
Email-eu	17.957	6.222	9.967	0.001	0.537	0.654	1.208	44.901
Facebook-Like	9.511	3.448	4.159	0.0	0.278	0.327	0.984	9.063
US Powergrid	26.817	74.324	63.602	0.002	0.947	0.604	30.195	54.728

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70044 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 5 | Algorithm runtime (in seconds) for synthetic networks.

Datasets	GOT	Betweenness	Closeness	Degree	Eigen vector	Page rank	VoteRank^{++}	Game theory
Weighted networks								
Erdos Renyi	1348.137	2082.576	3368.035	0.005	8.124	14.354	120.813	1896.849
Watts Strogatz	1301.513	1273.205	2632.612	0.004	61.372	13.673	126.020	1869.934
Barabasi Albert	3110.793	2807.24	6934.177	0.008	15.381	32.348	123.416	1883.478
Non-weighted networks								
Erdos Renyi	1355.243	2162.618	3425.229	0.006	10.369	16.783	142.659	1825.327
Watts Strogatz	1278.478	1230.734	2540.018	0.005	60.455	13.572	122.548	1747.52
Barabasi Albert	2807.457	2291.112	6690.13	0.008	15.246	32.331	126.293	1790.58

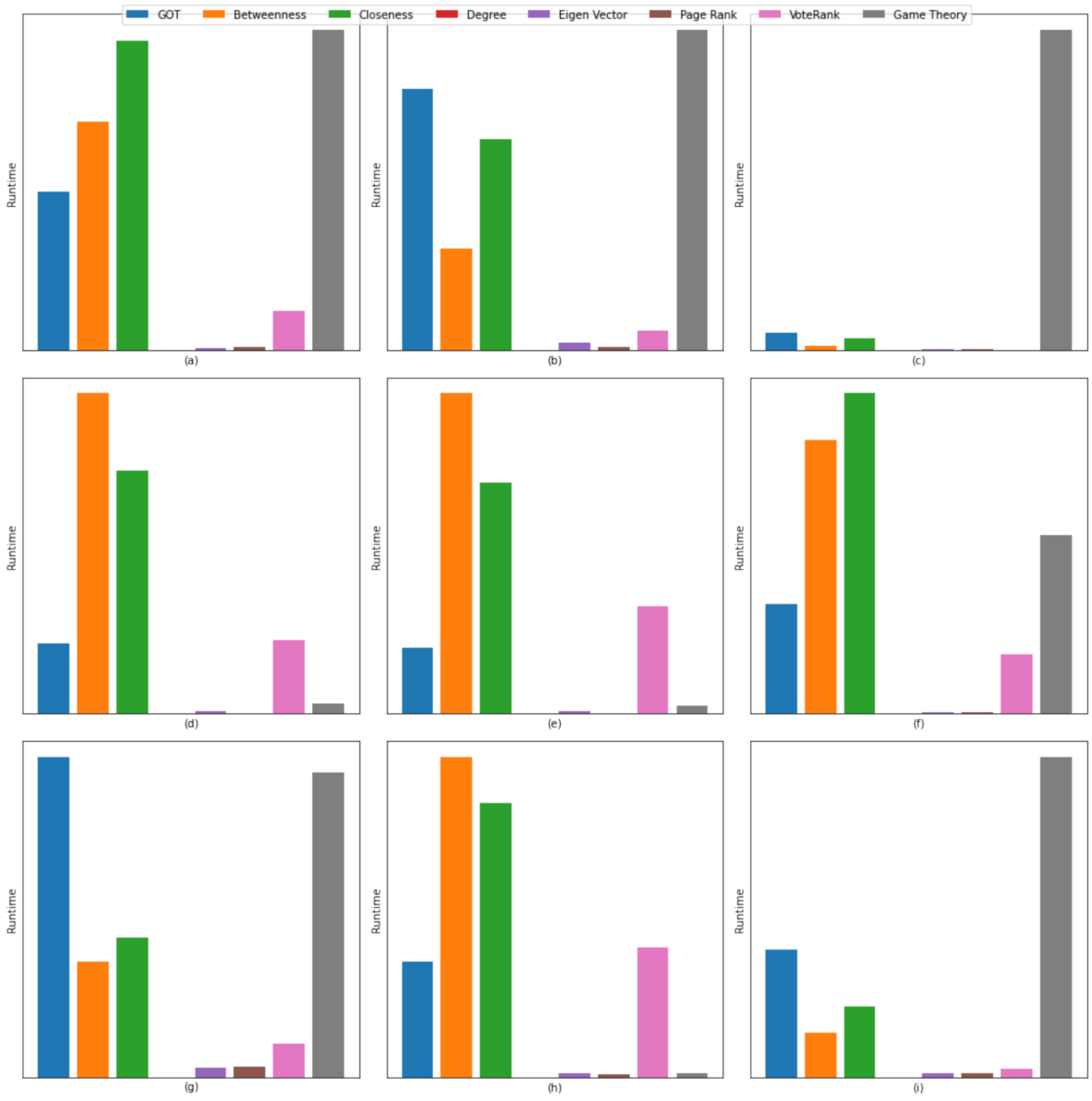


FIGURE 3 | Comparison of runtime of different algorithms on real networks: (a) Wiki Votes, (b) Facebook Combined, (c) Congress Bills, (d) p2p Nutella, (e) Newman, (f) CA-hep, (g) Facebook Like, (h) US Power Grid, and (i) Email-eu.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70044 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

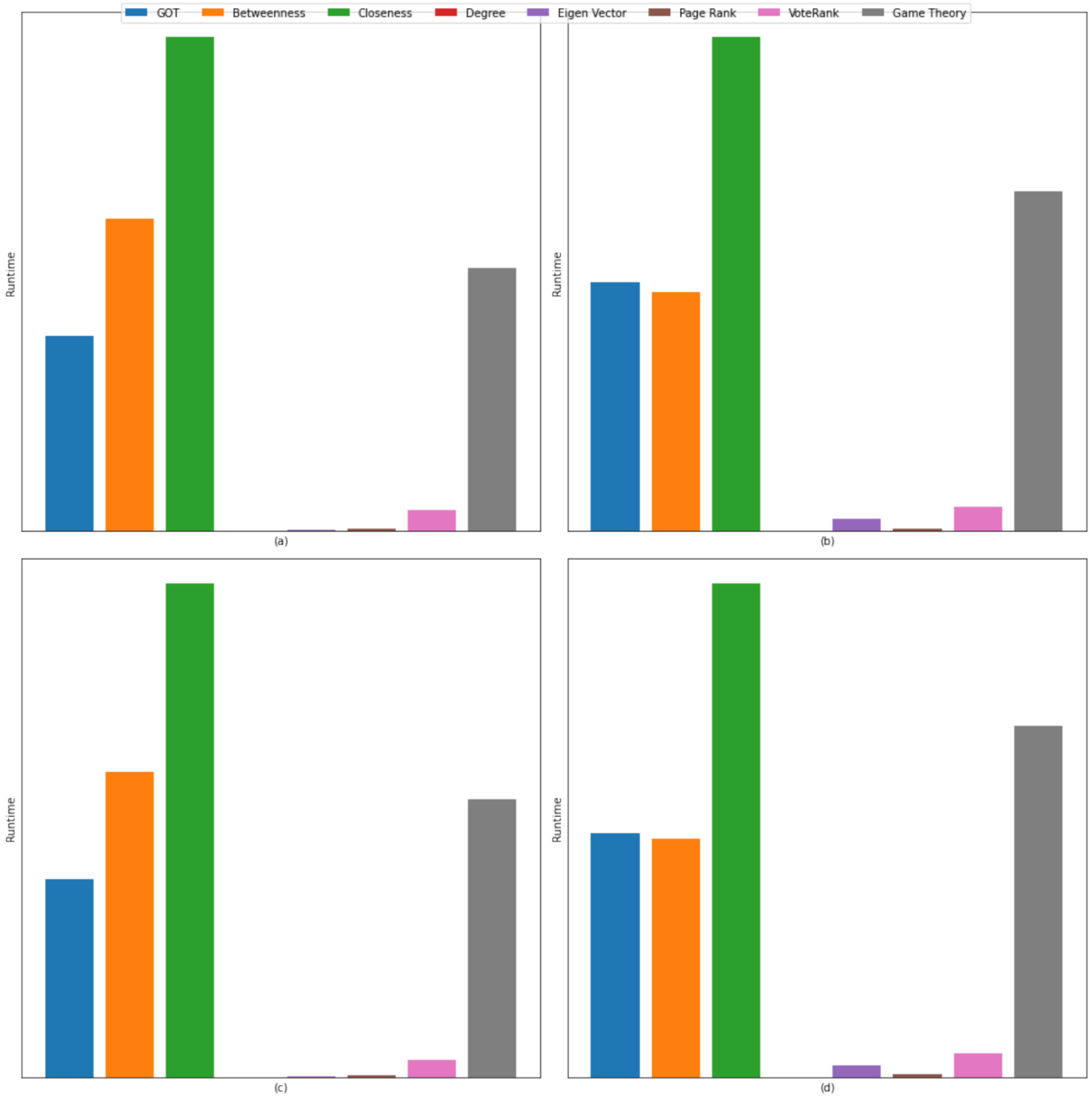


FIGURE 4 | Comparison of runtime of different algorithms on Synthetic Networks: (a) Erdos Renyi (non-weighted) (b) Watts Strogatz (non-weighted), (c) Erdos Renyi (Weighted), and (d) Watts Strogatz (Weighted).

5.2.2 | SIR

SIR model was used to simulate how well were the top $K\%$ influential nodes, identified by each algorithm, able to spread influence in the network. Initially, top 1% nodes were infected with the rest of the nodes considered as susceptible. The infection factor, β, was set to 0.15 and the recovery factor, γ, was set to 0.1. An average of 100 runs was considered to determine the results of the model. The model works [3] on the idea that at every stage, an infected node can further infect its neighbors with a probability β and an infected node can recover with the probability γ. The term $R_0 = \frac{\beta}{\gamma}$ is known as the reproduction rate [12] and is extremely important for explaining the spreading potential. If $R_0 > 1$, then rumor is likely to spread widely. Otherwise, it dies out.

5.2.3 | NRP

In social networks, node removal can help to predict how the removal of individuals affects the information spreading nature [29] in a network. This can be useful for identifying the most influential nodes. For example, in the case of social media platforms, it can help us to determine who the top influencers or nodes in the network are. This essentially determines the

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70044 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

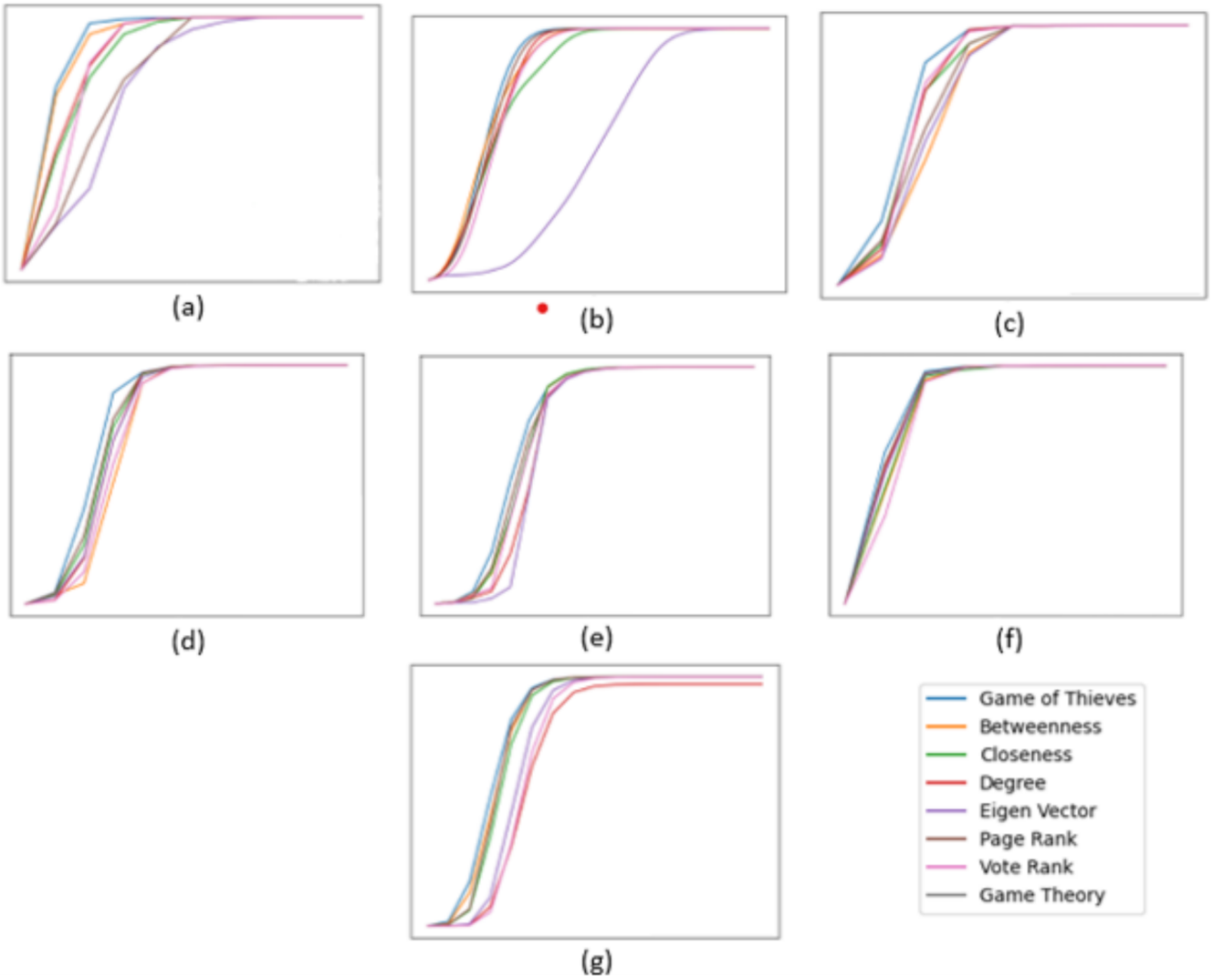


FIGURE 5 | Results for SIR for datasets: (a) Facebook combined; (b) US Power Grid; (c) Facebook Like; (d) p2p Nutella; (e) Newman; (f) Email-eu; and (g) CA-Hep.

influence spreading capacity and helps to evaluate the various algorithms.

Each algorithm is run to assign influence scores to the nodes, and they are sorted based on these scores. The order is from the most influential node to the least influential node. Further, the nodes from the sorted list are removed one by one from the network, and the size of the Giant Component and the number of connected components are computed after the removal. A node influence estimation metric is better if the number of connected components is as big as possible. In addition, the size of the giant component should be as small as possible [27].

6 | Results Discussions

This section lists the results of the experiments conducted as part of the study: computational time analysis (Section 6.1), SIR results (Section 6.2), and NRP results (Section 6.3). The experiments were conducted under the setup detailed in Section 5.

6.1 | Runtime Comparison of Proposed GOT With Existing Popular Algorithms

Tables 4 and 5 show the time (in seconds) each algorithm takes on the various real-world benchmark datasets and synthetic datasets, respectively. This data is also represented in Figures 3 and 4 for real-world and synthetic datasets, respectively.

GOT outperforms closeness and betweenness centrality measures substantially in Wiki Votes, US Power Grid, p2p Nutella, Newman, and CA-hep datasets, but underperforms in Facebook Combined, Facebook Like, Email-eu, and Congress Bills datasets. A clear distinction is observed in the relative performances of GOT and the two centrality measures based on the dataset size. Betweenness and closeness centralities perform better on datasets with a lesser number of nodes. Meanwhile, GOT performs better on datasets with a larger number of nodes. This difference can be attributed to the computational cost of overheads associated with GOT for smaller networks.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70044 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

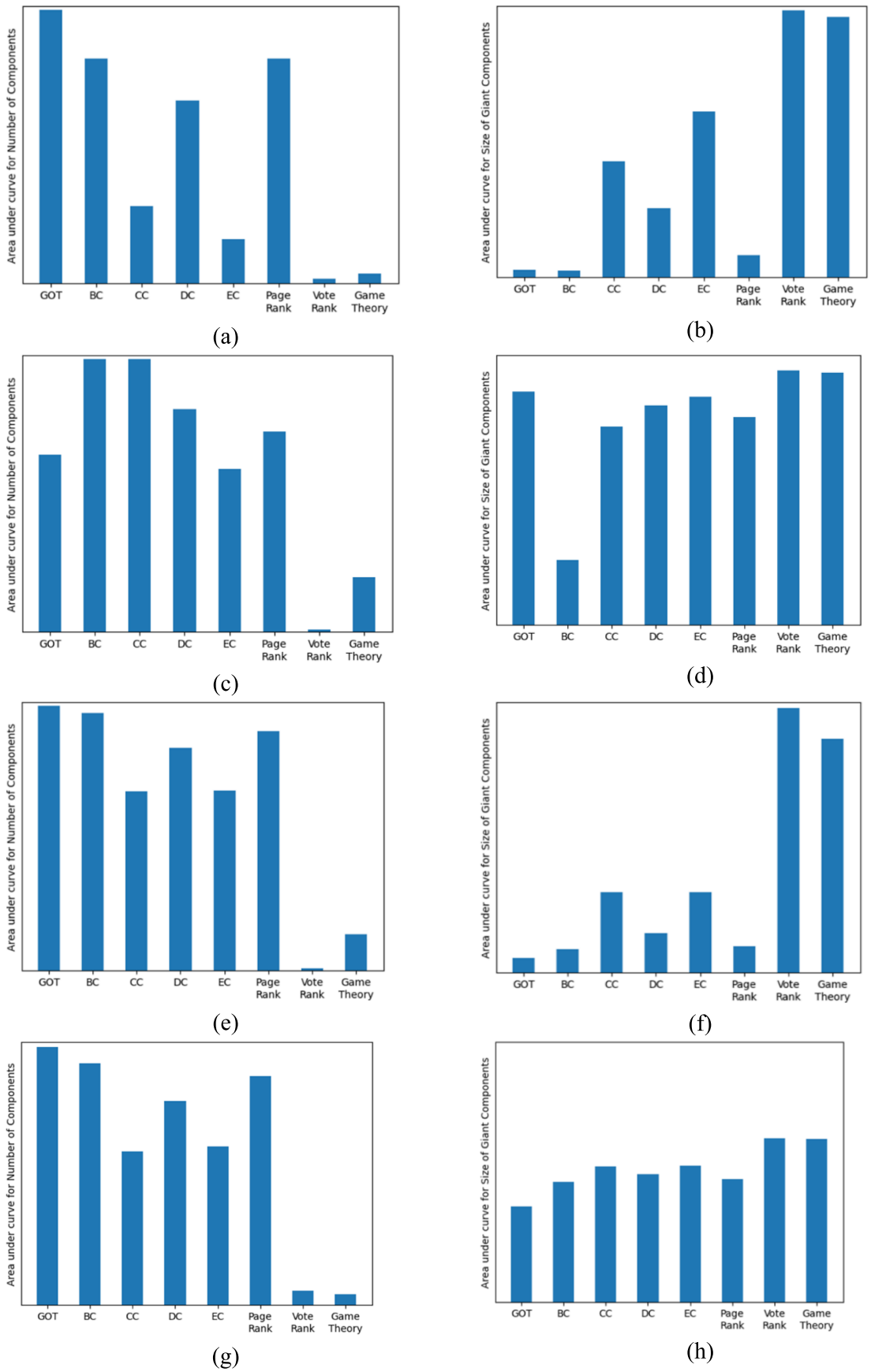


FIGURE 6 | Node removal procedure results for CA-Hep Dataset (a) and (b); Congress Bills Dataset (c) and (d); Facebook Combined Dataset (e) and (f); Erdos-Renyi Dataset (g) and (h). BC: betweenness centrality; CC: closeness centrality; DC: degree centrality; EC: eigenvector centrality; Vote Rank: VoteRank^{++}.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70044 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 6 | AUC for size of giant components versus number of removed nodes in NRP.

Dataset	GOT	Betweenness	Closeness	Degree	Eigen vector	Page rank	VoteRank^{++}	Game theory
CA-Hep	19,918,788	19,842,783	35,790,181	28,875,608	43,045,921	22,099,443	57,779,021	56,853,584
Congress Bills	1,453,480	1,286,674	1,418,811	1,440,244	1,448,964	1,428,680	1,474,895	1,472,720
Facebook-Like	204,661	210,776	250,640	222,086	250,371	212,829	378,590	357,047
P2p Nutella	22,046,917	26,070,724	32,192,810	27,110,379	33,974,124	24,642,085	51,594,509	53,677,769
Wiki-votes	17,045,375	6,912,178	8,138,794	7,713,975	8,281,848	7,326,541	21,893,234	22,878,320
Email-eu	28,875,608	19,842,783	35,790,181	28,875,608	43,045,921	22,099,443	57,779,021	56,853,584
Facebook-Combined	3,742,153	1,698,703	3,290,721	4,819,473	6,234,006	4,164,579	7,279,401	6,275,227
US Powergrid	1,870,186	1,601,481	4,816,469	1,564,373	7,173,860	1,468,591	4,860,886	9,666,008
Newman	75,057,808	69,845,477	78,614,610	78,406,168	83,008,676	75,033,251	85,657,249	66,544,329
Erdos Renyi	49,946,986	49,959,731	49,967,846	49,963,612	49,968,228	49,961,063	49,982,401	49,982,161
Watts Strogatz	49,932,455	49,925,151	49,934,994	49,931,449	49,964,609	49,927,634	49,937,176	49,995,000
Barabasi Albert	49,954,626	49,942,248	49,925,403	49,934,132	49,929,106	49,933,810	49,988,582	49,994,680

TABLE 7 | Calculated number of connected components in NRP for each algorithm.

Dataset	GOT	Betweenness	Closeness	Degree	Eigen vector	Page rank	VoteRank^{++}	Game theory
CA-Hep	36,151,685	31,287,540	16,627,016	27,154,405	13,301,409	31,305,501	9,379,921	9,888,904
Congress Bills	9003	12,979	12,979	10,891	8417	9962	1725	3900
Facebook-Like	187,817	183,219	134,915	161,553	135,086	172,010	25,490	46,733
P2p Nutella	36,134,279	31,700,059	25,053,178	30,305,414	23,399,678	32,882,917	7,095,328	5,198,417
Wiki-votes	8,015,647	17,987,577	16,503,065	16,867,053	16,315,970	17,416,805	3,357,706	2,208,027
Email-eu	36,151,685	31,287,540	16,627,016	27,154,405	13,301,409	31,305,501	9,379,921	9,888,904
Facebook-Combined	1,504,904	1,265,739	328,432	959,192	215,345	1,216,511	283,715	86,233
US Powergrid	7,555,172	5,429,316	1,295,510	6,641,287	1,304,638	7,239,404	3,243,628	383,293
Newman	27,165,958	24,736,612	24,925,553	25,103,894	23,408,061	26,765,759	25,760,354	25,181,964
Erdos Renyi	40,990	39,539	31,809	36,256	32,240	38,424	19,584	19,313
Watts Strogatz	29,432	31,704	30,225	32,409	28,307	33,003	33,432	9999
Barabasi Albert	41,123	48,321	59,566	57,430	59,620	55,804	15,665	7332

Degree centrality, eigenvector centrality, and Page Rank are almost instantly computed. A possible explanation for this result can be based on their inbuilt implementations and possible precomputations at the time of the network building phase.

For large datasets (having more than 50,000 nodes) like Twitter and Epinions, GOT was able to compute the influential nodes in under 2 h with the proposed experimental setup. Other algorithms could not complete the computation even after a reasonable period (over 36 h). Their runtime could not be ascertained due to system constraints. This is explained by the cubic and quadratic worst-case time complexities for these methods in comparison to the polylogarithmic time complexity for GOT, as explained in Section 5.1.

6.2 | SIR

Figure 5 shows the results obtained with SIR simulations on the real-world datasets. The graphs shown are plotted with the x-axis representing time and the y-axis representing the sum of infected and recovered nodes. The plotted graphs show the spread of information in the network over time. A steeper rise in the graphs denotes a faster information spread, indicating a more accurate identification of influential nodes.

As shown in Figure 5, GOT slightly outperforms all algorithms for every dataset. It dissipates the information at a faster rate than the other algorithms, indicated by its steep rise. It reaches the peak faster than the rest of the algorithms in the Facebook Combined and US power grid datasets and are on par with other algorithms for the rest of the datasets. This shows improved accuracy by GOT in identifying the most influential nodes in the network.

6.3 | NRP

The graphs in Figure 6 show the results obtained after performing the NRP on the datasets, as discussed in Section 5.2.3. The area under the curve (AUC) for the size of the giant component

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70044 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and the number of connected components is plotted to perform an overall comparison between the algorithms. The greater value of the number of connected components and the lower value of the size of giant components are the preferred characteristics. Table 6 shows the AUC for the size of giant components versus the number of removed nodes in the NRP, and Table 7 shows the calculated number of connected components in NRP for each algorithm.

GOT outperforms all algorithms in high energy physics (Figure 6a,b), Facebook combined (Figure 6d,e) and Erdos-Renyi (Figure 6g,h) datasets using both AUC—size of giant component and AUC—number of connected components. It has the highest area under the graph for the number of components and the lowest for the size of giant components among all the algorithms in each of these datasets. This result highlights the superior ability of GOT in identifying the influential nodes in a network.

Figure 6c,d show the Congress Bills Dataset in which GOT lags behind the other algorithms. Though it can also be observed that in this dataset Closeness and Betweenness perform the best in the number of connected components, while betweenness outperforms all the algorithms substantially in terms of the size of giant component.

This highlights the limitation of GOT, particularly in smaller networks like Congress Bills, which are not complex enough for the algorithm to understand the influence spreading behavior.

As was observed for runtime results in Section 6.1, GOT underperforms relative to other algorithms in smaller and less complex datasets.

7 | Conclusion

SNA has become increasingly important in the era of online social networks having billions of daily active users. This has led to a massive increase in the sizes of social networks along with their complexity. In this work, the evaluation of the GOT algorithm to rank influential nodes in social networks effectively with lower computational time is done. GOT was evaluated against multiple influential node identification algorithms in real-world and synthetic datasets using run time comparison, SIR, and NRP evaluation metrics. The results illustrate that GOT gives a faster and more accurate results in comparison to the other algorithms, particularly for large-scale networks. The limitations of GOT are highlighted in small networks. It is observed that GOT performs worse than most of the other algorithms, even the established centrality metrics, in small-sized networks. The algorithm takes more time to run due to the overheads and implementation inefficiencies in the smaller datasets. Additionally, the use of randomness in the algorithm causes poor performance in smaller networks due to less network complexity and reduced epochs. The random placement of thieves can sometimes lead to suboptimal starting positions, affecting the efficiency and convergence of the algorithm. While this randomness promotes exploration, it can also limit the algorithm's ability to consistently find the optimal solution, especially in complex or noisy networks. In the future, more potential strategies, such as enhanced initialization techniques or hybridizing GOT with other optimization approaches, may be applied to mitigate these limitations. Further study in the field of influential nodes is required to find a single algorithm that offers good results in all kinds of networks with reasonable computational time.

Endnotes

[1] Globally, there were 5.3 billion internet users as of October 2023. In total, 4.95 billion people use social media platforms, representing 61.4% of the world's population.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

References

1. A. Ficara, G. Fiumara, P. De Meo, and A. Liotta, "Correlations Among Game of Thieves and Other Centrality Measures in Complex Networks," in *Data Science and Internet of Things*, ed. G. Fortino, A. Liotta, R. Gravina, and A. Longheu (Springer, 2021), https://doi.org/10.1007/978-3-030-67197-6_3.

2. B. A. Bhuiyan, "An Overview of Game Theory and Some Applications," *Philosophical Progress* 59, no. 1–2 (2018): 111–128.

3. D. C. Mocanu, G. Exarchakos, and A. Liotta, "Decentralized Dynamic Understanding of Hidden Relations in Complex Networks," *Scientific Reports* 8 (2018): 1571, https://doi.org/10.1038/s41598-018-19356-4.

4. J. Golbeck, *Chapter 21—Analyzing Networks: Introduction to Social Media Investigation* (Syngress, 2015), 221–235, https://doi.org/10.1016/B978-0-12-801656-5.00021-4.

5. J. Sheng, J. Zhu, Y. Wang, B. Wang, and H. Zheng'ang, *Identifying Influential Nodes of Complex Networks Based on Trust-Value* (School of Computer Science and Engineering, Central South University, 2020).

6. S. Gao, J. Ma, Z. Chen, G. Wang, and C. Xing, "Ranking the Spreading Ability of Nodes in Complex Networks Based on Local Structure," *Physica A* 403 (2014): 130–147.

7. C. Mao and W. Xiao, "A Comprehensive Algorithm for Evaluating Node Influences in Social Networks Based on Preference Analysis and Random Walk," *Complexity* 2018 (2018): 1–16, https://doi.org/10.1155/2018/1528341.

8. R. Narayanam and Y. Narahari, "A Shapley Value-Based Approach to Discover Influential Nodes in Social Networks," *IEEE Transactions on Automation Science and Engineering* 8, no. 1 (2011): 130–147, https://doi.org/10.1109/TASE.2010.2052042.

9. M. Li, R. Zhang, R. Hu, F. Yang, Y. Yao, and Y. Yuan, "Identifying and Ranking Influential Spreaders In Complex Networks By Combining A Local-Degree Sum And The Clustering Coefficient," https://doi.org/10.1142/S0217979218501187.

10. Y. Chen, X. Cao, and K. J. Liu, "Community Detection in Networks: A Game-Theoretic Framework," *EURASIP Journal on Advances in Signal Processing* 2019 (2019): 60, https://doi.org/10.1186/s13634-019-0655-z.

11. A. Zareie, A. Sheikhahmadi, and M. Jalili, "Identification of Influential Users in Social Network Using Gray Wolf Optimization Algorithm," *Expert Systems with Applications* 142 (2020): 112971, https://doi.org/10.1016/j.eswa.2019.112971.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70044 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

12. P. Liu, L. Li, S. Fang, and Y. Yao, "Identifying Influential Nodes in Social Networks: A Voting Approach," *Chaos, Solitons & Fractals* 152 (2021): 111309, https://doi.org/10.1016/j.chaos.2021.111309.

13. C. Guo, X. Chen, D. Chen, H. Gao, L. Yang, and J. Ma, "Influential Nodes Identification in Complex Networks via Information Entropy," *Entropy* 22, no. 2 (2020): 242.

14. M. Jain, A. Jaswani, A. Mehra, and L. Mudgal, "EDGly: Detection of Influential Nodes Using Game Theory," *Multimedia Tools and Applications* 81 (2022): 1625–1647, https://doi.org/10.1007/s11042-021-11440-8.

15. S. Kumar, A. Kumar, A. Mallik, and S. Dhall, "Opinion Leader Detection in Asian Social Networks Using Modified Spider Monkey Optimization," *ACM Transactions on Asian and Low-Resource Language Information Processing* 2022 22, no. 5 (2023): 1–26, https://doi.org/10.1145/3555311.

16. H. Cherifi, G. Palla, B. K. Szymanski, et al., "On Community Structure in Complex Networks: Challenges and Opportunities," *Applied Network Science* 4 (2019): 117, https://doi.org/10.1007/s41109-019-0238-9.

17. Z. Ghalmane, M. El Hassouni, C. Cherifi, et al., "Centrality in Modular Networks," *EPJ Data Science* 8 (2019): 15, https://doi.org/10.1140/epjds/s13688-019-0195-7.

18. S. Rajeh, M. Savonnet, E. Leclercq, and H. Cherifi, "Characterizing the Interactions Between Classical and Community-Aware Centrality Measures in Complex Networks," *Scientific Reports* 11 (2021): 10088, https://doi.org/10.1038/s41598-021-89549-x.

19. S. Rajeh, M. Savonnet, E. Leclercq, and H. Cherifi, "Comparative Evaluation of Community-Aware Centrality Measures," *Quality & Quantity: International Journal of Methodology* 57, no. 2 (2023): 1273–1302, https://doi.org/10.1007/s11135-023-01493-6.

20. S. Rajeh and H. Cherifi, "Ranking Influential Nodes in Complex Networks With Community Structure," *PLoS One* 17, no. 8 (2022): e0273610, https://doi.org/10.1371/journal.pone.0273610.

21. C. Blöcker, J. C. Nieves, and M. Rosvall, "Map Equation Centrality: Community-Aware Centrality Based on the Map Equation," *Applied Network Science* 7 (2022): 56, https://doi.org/10.1007/s41109-022-00477-9.

22. T. Magelinski, M. Bartulovic, and K. M. Carley, "Measuring Node Contribution to Community Structure With Modularity Vitality," *IEEE Transactions on Network Science and Engineering* 8, no. 1 (2021): 707–723, https://doi.org/10.1109/TNSE.2020.3049068.

23. S. Rajeh and H. Cherifi, "A Community-Aware Centrality Framework Based on Overlapping Modularity," *Social Network Analysis and Mining* 13 (2023): 37, https://doi.org/10.1007/s13278-023-01040-9.

24. R. Saxena, S. Kaur, and V. Bhatnagar, "Social Centrality Using Network Hierarchy and Community Structure," *Data Mining and Knowledge Discovery* 32 (2018): 1421–1443, https://doi.org/10.1007/s10618-018-0582-x.

25. C. Sciarra, G. Chiarotti, F. Laio, and L. Ridolfi, "A Change of Perspective in Network Centrality," *Scientific Reports* 8 (2018): 15269, https://doi.org/10.1038/s41598-018-33336-8.

26. S. Rajeh and H. Cherifi, "On the Role of Diffusion Dynamics on Community-Aware Centrality Measures," *PLoS One* 19, no. 7 (2024): e0306561, https://doi.org/10.1371/journal.pone.0306561.

27. C. Castellano and R. Pastor-Satorras, "Thresholds for Epidemic Spreading in Networks," *Physical Review Letters* 105, no. 21 (2010): 218701, https://doi.org/10.1103/PhysRevLett.105.218701.

28. M. Bellingeri, D. Bevacqua, F. Scotognella, et al., "Link and Node Removal in Real Social Networks: A Review," *Frontiers in Physics* 8 (2020): 228, https://doi.org/10.3389/fphy.2020.00228.

29. A.-M. Kermarrec, E. Le Merrer, B. Sericola, and G. Trédan, "Second Order Centrality: Distributed Assessment of Nodes Criticity in Complex Networks," *Computer Communications* 34, no. 5 (2011): 619–628, https://doi.org/10.1016/j.comcom.2010.06.007.

30. H. Yin, A. R. Benson, J. Leskovec, and D. F. Gleich, "Local Higher-Order Graph Clustering," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2017), 555–564.

31. M. E. J. Newman, "The Structure of Scientific Collaboration Networks," *PNAS* 98 (2001): 404–409, http://opsahl.co.uk/tnet/datasets/Newman-Cond_mat_95-99-co_occurrence.txt.

32. J. McAuley and J. Leskovec, *Learning to Discover Social Circles in Ego Networks* (NIPS, 2012).

33. A. R. Benson, R. Abebe, M. T. Schaub, A. Jadbabaie, and J. Kleinberg, "Simplicial Closure and Higher-Order Link Prediction," *Proceedings of the National Academy of Sciences of the United States of America* 115, no. 48 (2018): E11221–E11230, https://doi.org/10.1073/pnas.1800683115.

34. D. J. Watts and S. H. Strogatz, "Collective Dynamics of "Small-World" Networks," *Nature* 393 (1998): 440–442.

35. T. Opsahl, "Triadic Closure in Two-Mode Networks: Redefining the Global and Local Clustering Coefficients," *Social Networks* 35, no. 2 (2013): 159–167, https://doi.org/10.1016/j.socnet.2011.07.001.

36. J. Leskovec, D. Huttenlocher, and J. Kleinberg, *Predicting Positive and Negative Links in Online Social Networks* (WWW, 2010).

37. J. K. Leskovec and C. Faloutsos, "Graph Evolution: Densification and Shrinking Diameters," *ACM Transactions on Knowledge Discovery From Data (ACM TKDD)* 1, no. 1 (2007): 2.

38. M. Richardson, R. Agrawal, and P. Domingos, *Trust Management for the Semantic Web* (ISWC, 2003).

39. P. Erdös and A. Rényi, "On Random Graphs I," *Publicationes Mathematicae Debrecen* 6 (1959): 290–297.

40. A. L. Barabási and R. Albert, "Emergence of Scaling in Random Networks," *Science* 286 (1999): 509–512.

41. M. E. J. Newman and D. J. Watts, "Renormalization Group Analysis of the Small-World Network Model," *Physics Letters A* 263, no. 4–6 (1999): 341–346, https://doi.org/10.1016/S0375-9601(99)00757-4.

42. A. Bihari and M. Pandia, "Eigenvector Centrality and its Application in Research Professionals' Relationship Network," in *2015 1st International Conference on Futuristic Trends in Computational Analysis and Knowledge Management, ABLAZE 2015* (IEEE, 2015), 510–514, https://doi.org/10.1109/ABLAZE.2015.7154915.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70044 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

International Journal of Future Engineering Innovations

Renewable Energy Systems: Engineering Solutions for a Sustainable World

Dr. Rajesh Gupta
Department of Mechanical Engineering, Delhi Technological University (DTU), New Delhi, India

* Corresponding Author: **Dr. Rajesh Gupta**

Article Info

ISSN (online): 3049-1215
Volume: 02
Issue: 02
March-April 2025
Received: 16-01-2025
Accepted: 19-02-2025
Page No: 12-13

Abstract
The increasing demand for energy and concerns about environmental degradation have accelerated the transition to renewable energy systems. This research paper explores the engineering solutions that enable sustainable energy production, including solar, wind, hydro, biomass, and geothermal energy. Advances in energy storage, grid integration, and efficiency enhancement technologies are discussed to highlight the role of engineering in achieving a sustainable energy future. Additionally, the paper examines the challenges in implementing renewable energy on a large scale and proposes innovative solutions for overcoming these barriers.

Keywords: Renewable Energy, Sustainable Engineering, Solar Power, Wind Energy, Energy Storage, Grid Integration, Smart Grids, Energy Efficiency

Introduction
The global energy crisis and climate change concerns necessitate a shift from fossil fuels to renewable energy sources. Engineering innovations play a crucial role in harnessing renewable energy efficiently and integrating it into existing infrastructure. This paper examines various renewable energy systems, their technological advancements, and engineering solutions for improving sustainability.

Solar Energy Systems
1. Photovoltaic (PV) Technology
Solar photovoltaic systems convert sunlight into electricity through semiconductor materials. Advances in PV technology, including perovskite solar cells and bifacial panels, have improved efficiency and reduced costs. Research is also focused on tandem solar cells and multi-junction solar cells, which can achieve higher conversion efficiencies.

2. Solar Thermal Systems
Concentrated solar power (CSP) plants use mirrors to focus sunlight and generate heat, which drives turbines for electricity production. Engineering solutions such as thermal energy storage, hybrid CSP-PV systems, and advanced heat transfer fluids enhance efficiency. The integration of AI-driven tracking systems further optimizes solar panel orientation for maximum energy capture.

3. Building-Integrated Photovoltaics (BIPV)
BIPV systems integrate solar panels into buildings, replacing traditional construction materials while generating electricity. Advances in transparent solar cells and flexible thin-film solar panels are expanding their applications.

Wind Energy Solutions

1. Aerodynamic Blade Design

Innovations in wind turbine blade materials and designs, such as carbon fiber composites and smart rotor technology, enhance energy capture and durability. The development of vertical-axis wind turbines (VAWT) has provided solutions for urban energy generation, where traditional horizontal-axis wind turbines (HAWT) are not feasible.

2. Offshore Wind Farms

Offshore wind farms provide higher energy yields due to stronger and more consistent wind patterns. Floating turbine technology, improved anchoring systems, and predictive maintenance using IoT sensors enhance their feasibility and resilience. Hybrid offshore platforms combining wind, solar, and wave energy are being explored.

Hydropower and Biomass Energy

1. Small-Scale Hydropower

Micro and pico-hydropower systems provide decentralized energy solutions, especially in remote areas. Advances in turbine efficiency, modular designs, and fish-friendly hydropower solutions reduce environmental impact while improving energy generation.

2. Biomass Energy Conversion

Engineering innovations in biomass gasification, anaerobic digestion, and biofuel production optimize energy output while minimizing environmental impact. Waste-to-energy technologies, including municipal solid waste gasification and algae-based biofuels, contribute to circular economy models and carbon neutrality.

Geothermal Energy and Storage Technologies

1. Enhanced Geothermal Systems (EGS)

Engineering advancements in deep drilling techniques, artificial reservoir stimulation, and high-temperature-resistant materials enhance the viability of geothermal energy in various locations. The development of closed-loop geothermal systems has reduced the environmental concerns associated with traditional geothermal extraction.

2. Energy Storage Solutions

Efficient storage systems, including lithium-ion batteries, flow batteries, hydrogen-based storage, and supercapacitors, address intermittency challenges in renewable energy sources. Innovations in solid-state batteries and gravity-based storage systems provide promising alternatives for long-duration energy storage.

Grid Integration and Smart Energy Systems

1. Smart Grids and Demand Response

Smart grids optimize electricity distribution through real-time monitoring, automated control, and demand-side management. AI-driven predictive analytics enhance grid stability by forecasting energy demand and supply fluctuations. Vehicle-to-grid (V2G) technology allows electric vehicles to act as distributed energy storage units, stabilizing the grid during peak demand periods.

2. Hybrid Renewable Energy Systems

Integrating multiple renewable sources with storage and smart grid technologies improves energy reliability and sustainability. Hybrid microgrids incorporating solar, wind, and battery storage provide resilient power solutions for remote communities and disaster-prone areas.

3. Power-to-X Technologies

Power-to-X (P2X) solutions convert excess renewable energy into synthetic fuels, hydrogen, and chemicals for use in industrial applications. Green hydrogen production through electrolysis is gaining traction as a clean energy carrier for sectors that are difficult to electrify, such as heavy industry and transportation.

Challenges and Future Prospects

Despite rapid advancements, challenges such as high initial costs, resource intermittency, energy storage limitations, and policy barriers remain. Future research aims to develop cost-effective, scalable, and efficient renewable energy systems through engineering innovations. International collaboration on standardization, subsidies for clean energy technologies, and investment in next-generation energy infrastructure will accelerate the transition to a fully renewable energy-powered world.

Conclusion

Renewable energy systems are essential for a sustainable future, and engineering solutions play a critical role in their optimization. Advancements in solar, wind, hydro, biomass, geothermal energy, and storage technologies will drive the transition toward a cleaner, more resilient energy landscape. Smart grids, AI-driven energy management, and hybrid renewable energy systems will further enhance efficiency and integration. Through continued innovation and policy support, the global energy sector can achieve carbon neutrality and sustainable development.

References

1. Boyle G. Renewable energy: power for a sustainable future. Oxford University Press. 2012.
2. Twidell J, Weir T. Renewable energy resources. Taylor & Francis. 2021.
3. Hansen K, Breyer C, Lund H. Status and perspectives of 100% renewable energy systems. Energy. 2019;175:471-480.
4. Panwar NL, Kaushik SC, Kothari S. Role of renewable energy sources in environmental protection. Renewable and Sustainable Energy Reviews. 2011;15(3):1513-1524.
5. Jacobson MZ, Delucchi MA. A roadmap for transitioning the world to 100% renewable energy. Energy & Environmental Science. 2011;8(12):2093-2117.
6. Ferrer JV, Iglesias RB. Power-to-X pathways for sustainable energy transition. International Journal of Hydrogen Energy. 2021;46(5):2555-2570.
7. Müller A, Hoffmann C, Becker S. Advances in energy storage technologies for renewable integration. Energy Storage Materials. 2020;30:15-29.

Check for updates

Concurrency and Computation: Practice and Experience

WILEY

RESEARCH ARTICLE

Revocable and Privacy-Preserving CP-ABE Scheme for Secure mHealth Data Access in Blockchain

Anita Thakur | Virender Ranga | Ritu Agarwal

Delhi Technological University, Delhi, India

Correspondence: Virender Ranga (virenderranga@dtu.ac.in)

Received: 9 October 2024 | **Revised:** 6 March 2025 | **Accepted:** 14 March 2025

Keywords: attribute-based encryption | blockchain | scyther tool | security

ABSTRACT

Innovations in technology are revolutionizing healthcare, driving a shift toward patient-centric smart healthcare systems. Mobile health (mHealth) leverages innovations in wearable sensors, telecommunications, and IoT to establish a novel healthcare model that prioritizes the patient, enabling real-time monitoring, personalized interventions, and improved access to care, ultimately fostering a proactive approach to health management and enhancing overall patient outcomes. However, safeguarding patient data transparency, security, and privacy within mHealth systems presents significant challenges, particularly concerning personal health records (PHR). Ciphertext-Policy Attribute-Based Encryption (CP-ABE) offers a competent answer to facilitating one-to-many data sharing in healthcare environments. Nevertheless, several issues must be addressed before CP-ABE can be widely deployed. These include the need for timely and effective attribute revocation when user attributes change, resistance to collusion attacks, and ensuring data integrity. This paper proposes a revocable and secure fine-grained access scheme using blockchain and CP-ABE. We compare four prominent state-of-the-art schemes through comprehensive experimentation with our proposed approach. Our results demonstrate the relative performance of our scheme, showing a significant reduction in computational costs. Specifically, the key generation cost is reduced by ≈35% to 67%, and the encryption cost is reduced by ≈26% to 39%. A detailed analysis of communication, computational, and storage overhead reveals that our suggested solution offers a distinct advantage in terms of efficiency. The Scyther tool is employed to verify the security measures and assess the accuracy of proposed methodologies, subsequently conducting experiments to showcase its efficacy.

1 | Introduction

The world of healthcare has witnessed a noteworthy evolution with the emergence of mHealth, which builds upon the established foundation of electronic health (eHealth). As a subdomain of eHealth, mHealth capitalizes on the ubiquitous nature of mobile devices to deliver three core functionalities. Firstly, it facilitates the acquisition of health-related data directly from patients. Secondly, the acquired data is securely stored on servers managed by healthcare providers, ensuring its integrity and accessibility. Finally, mHealth empowers disseminating healthcare information to authorized users [1]. By implementing secure, attribute-based, fine-grained access controls, this data distribution can be tailored to specific user roles, such as medical practitioners, researchers, and patients [2]. The healthcare sector generates substantial critical data, presenting valuable research opportunities. However, a primary challenge within the mHealth ecosystem lies in safeguarding the inherent sensitivity of this health information. The potential threat of unauthorized access or attacks necessitates robust security measures to mitigate the risk of data breaches caused by internal and external actors [3]. Furthermore, the online dissemination of sensitive patient

© 2025 John Wiley & Sons Ltd.

information to facilitate research activities presents considerable privacy and security challenges for researchers [4]. Integrating mHealth devices with Electronic Health Records (EHRs) presents opportunities and challenges for patient privacy. EHRs, a comprehensive digital record of an individual's medical history, contain highly sensitive information such as diagnoses, allergies, medications, and immunization dates. This data is directly linked to a patient's real identity, making it crucial to ensure appropriate access controls. However, patients may only sometimes desire to share their mHealth data history with every healthcare provider. Historical health data may be sensitive or irrelevant to the current medical condition. Therefore, a key concern within the mHealth ecosystem is the ability of patients to exert granular control over their healthz information [5].

To mitigate this challenge, a pioneering approach termed ABE [6] has garnered significant attention. ABE furnishes a scalable framework for enforcing detailed access control on encrypted data. ABE offers a granular approach to access control for encrypted data [7]. Employing the attributes and access policies built into secret keys and ciphertexts is possible. In CP-ABE, users' secret keys are correlated with a specific set of attributes. Meanwhile, a sender generates ciphertexts accompanied by access policies dictating the required attributes for decryption [8] as illustrated in Figure 1.

1.1 | Our Contributions

Over recent years, the widespread adoption of modern medical sharing systems has facilitated convenient access to healthcare services. Due to its inherent benefits, ABE has emerged as an effective strategy for securing these systems. However, the practical deployment of ABE-based medical sharing raises significant concerns regarding user privacy, data confidentiality, and overall system security. Data users (visitors) seek anonymity regarding their access patterns, even from issuing authorities.

This paper presents a data access control scheme that is both expressive and secure. We propose a new Blockchain-based CPABE (BloCPABE) scheme to achieve revocation and privacy-preserving capabilities. The proposed BloCPABE scheme introduces anonymous key generation, enhancing user anonymity and mitigating the potential for collusion attacks. We design a fine-grained data access control methodology that leverages CP-ABE in conjunction with blockchain technology. The key contributions of this research work are demonstrated below:

- A secure CP-ABE scheme, named BloCPABE, is proposed to address security challenges in mHealth systems. The scheme enables efficient attribute revocation by updating the associated secret key and ciphertext, requiring minimal computational overhead.
- The proposed scheme establishes user privacy, backward and forward security, and is resistant to collusion attacks.
- Blockchain technology is leveraged to store outsourced data, ensuring immutability and tamper-proof integrity.
- The scheme incorporates the Linear Secret Sharing Scheme to formulate access structures, ensuring the confidentiality of the message and preventing private information leakage.
- Security measures and the proposed methodologies are rigorously evaluated and validated using the Scyther tool.

1.2 | Road Map

The systematic outline of this research paper is presented in this way: Section 2 gives preliminary studies on bilinear maps, the LSSS, and blockchain technology. The Section 3 surveys the pertinent literature. The Section 4 delineates the foundational construction of BloCPABE. The Section 5 details a concrete construction of BloCPABE. Section 6 offers a formal verification and security analysis of the scheme. Section 7 provides a

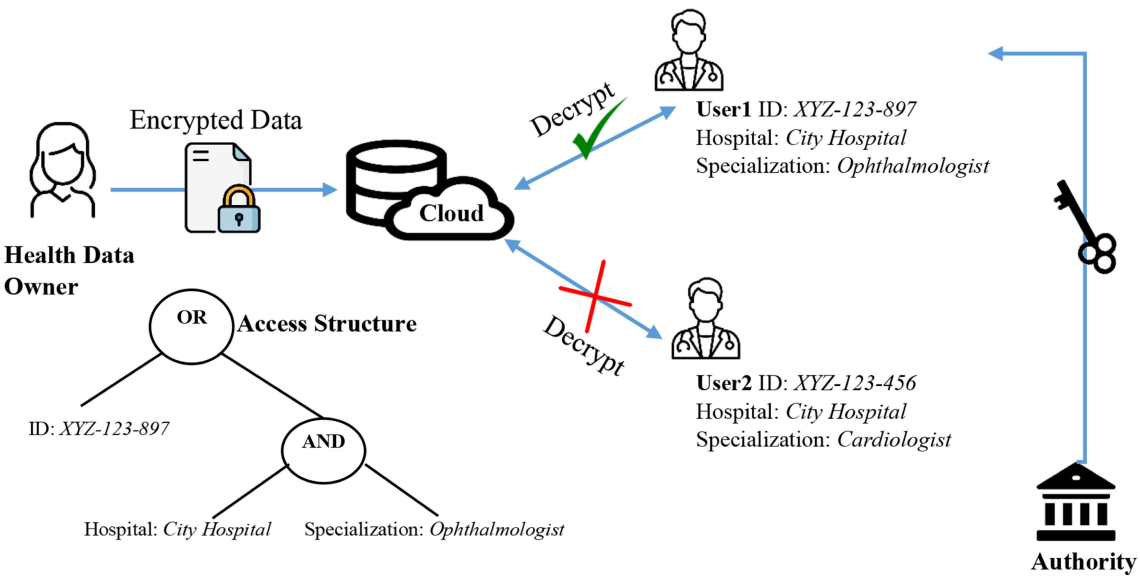


FIGURE 1 | A cloud-based health data storage system employing CP-ABE scheme.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70064 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

performance analysis, including experimental and simulation results. Lastly, Section 8 concludes the study and suggests paths for future research.

2 | Preliminaries

Here, we outline several basic and additional concepts that are employed in our scheme.

2.1 | Attribute-Based Encryption

ABE is a framework that utilizes user attributes for data encryption rather than relying solely on user identities. This approach enables any individual who meets a specified set of attributes to decrypt the data, eliminating the necessity for a unique key for each user. In ciphertext policy ABE, the access policy is incorporated within the ciphertext, and the attribute set is associated with the user's private key. This arrangement grants the data owner complete authority over who is permitted to decrypt their data. Also, Table 1 provides a list of the key notations used in this paper, along with their respective descriptions. CP-ABE comprises four key algorithms:

- **Setup**$(\lambda) \rightarrow (PP, MSK)$: Here, Setup takes the security parameters as input and produces the PP and the MSK.
- **Encrypt**$(PP, \mathbb{A}, data) \rightarrow CT$: This algorithm uses PP, the access structure $\mathbb{A}$, and the data file as inputs to produce the encrypted ciphertext (CT).
- **KeyGen**$(PP, MSK, S) \rightarrow SK$: This algorithm takes the PP, the MSK, and the attribute set (S) to produce private key (SK) of the user.
- **Decrypt**$(PP, CT, SK) \rightarrow data$: This algorithm uses the PP, the CT, and the SK to decrypt the ciphertext and recover the original data.

2.2 | Bilinear Pairing

Consider two multiplicative cyclic groups $\mathbb{G}_0$ and $\mathbb{G}_T$, both having the same prime order p and g which depicts the generator of $\mathbb{G}_0$. We say $e : \mathbb{G}_0 \times \mathbb{G}_0 \rightarrow \mathbb{G}_T$ is a bilinear pairing if the following properties hold:

- *Bilinearity*: $\forall a, b \in \mathbb{Z}_\rho$ *and* $g_1, g_2 \in \mathbb{G}_0$, $e(g_1^a, g_2^b) = e(g_1, g_2)ab$.
- *Non-degeneracy: There exists* $g_1 \in \mathbb{G}_0$, $e(g_1, g_1) \neq 1$
- *Computability*: $e(g_1, g_2)$ *is efficiently computable for all* $g_1, g_2 \in \mathbb{G}_0$ *in polynomial time.*

2.3 | Decisional Q-Parallel Bilinear Diffie-Hellman Exponent Assumption

The definition of the Decisional q-parallel BDHE problem is as described by Waters, B [9]. *Given a bilinear map* $(p, \mathbb{G}_1, \mathbb{G}_2, e, g)$, *choose random* $a, s, b_1, \ldots, b_q \in \mathbb{Z}_p$. *If the adversary* $\mathfrak{A}$ *is given*

TABLE 1 | Notations and their description.

Notation	Descriptions
CA	Certificate authority or Central authority
AA	Attribute authority (Hospital)
DO	Data owner (patient or health data owner)
DU	Data user (physician, researcher, etc.)
$\mathfrak{A}$	Adversary
$\mathfrak{C}$	Challenger
λ	Security Parameter
G_1, G_T	Multiplicative cyclic group of prime order p
g	Generator of G_1
e	Mapping function $e(G_1, G_1) \rightarrow G_T$
$\mathcal{H}$	Collusion resistant hash function
PP	Public Parameter
GPP	Global Public Parameter
MK	Master Key generated by the CA
uid_k	Unique identifier for user k
$Certificate(uid_k)$	User's certificate
$UPK_{uid_k}, UPK'_{uid_k}$	User Public Key
$USK_{uid_k}, USK'_{uid_k}$	User Secret Key
MSK	Master Secret Key generated by AA
PK	Public Key generated by AA
$\mathcal{U}$	Universe of attributes controlled by AA
VK_x	Version key of the attribute
AK_x	Public attribute key of the attribute
S	Attribute sets of user
SK	Decryption key for DU generated by AA
$\mathbb{A}$	Access structure
CT	Ciphertext
$\hat{x}$	Revoked attribute from user
$AUK_{\hat{x}}$	Attribute update key
$KUK_{\hat{x},k}$	user key update key
$CUK_{\hat{x}}$	Ciphertext key update

$$\vec{y} = g, g^s, g^a, \ldots g^{a^q}, g^{a^{q+2}}, \ldots, g^{a^{2q}}$$

$$\forall 1 \leq j \leq q, g^{s.b_j}, g^{a/b_j}, \ldots, g^{a^q/b_j}, g^{a^{q+2}/b_j}, \ldots, g^{a^{2q}/b_j}$$

$\forall 1 \leq j, k \leq q, \quad k \neq j, g^{a.s.b_k/b_j}, \ldots, g^{a^q.s.b_k/b_j}$ *it must be difficult to distinguish between* $e(g,g)^{a^{q+1}s} \in \mathbb{G}_2$ *and a random element* $R \in \mathbb{G}_2$. *An algorithm* $\mathfrak{C}$ *guesses* $z \in \{0, 1\}$ *with advantage* ε *in solving the Decisional q-parallel BDHE problem in* $\mathbb{G}_2$ *if*

$$\left| Pr[\mathfrak{C}(\vec{y}, T = e(g,g)^{a^{q+1}s}) = 0] - Pr[\mathfrak{C}(\vec{y}, T = R) = 1] \right| \geq \varepsilon$$

As in [9], *we can say that the Decisional q-parallel BDHE assumption holds if there is no polynomial time algorithm to solve the Decisional q-parallel BDHE problem with a non-negligible advantage.*

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70064 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.4 | Linear Secret-Sharing Schemes (LSSS)

A secret-sharing scheme P over a set of parties P is called linear over $\mathbb{Z}_\rho$ if the following conditions hold:

- *The shares for each party form a vector over* $\mathbb{Z}_\rho$.
- *There exists a share-generating matrix* $M \in \mathbb{Z}_\rho^{l\times n}$ *with* l *rows and* n *columns which can make shares for* $\prod$. *For all* $i \in \{1, \ldots, l\}$ *the* i^{th} *row of* M *is labeled via a function* ρ, *that associates* M_i *to the party* $\rho(i)$. *Considering the vector* $\overrightarrow{v} = (s, r_2, \ldots, r_n) \in \mathbb{Z}_\rho^n$, *where* $s \in \mathbb{Z}_\rho$ *is the secret to be shared, and* $r_2, \ldots, r_n$ *are randomly chosen from* $\mathbb{Z}_\rho$, *then* $M_{\overrightarrow{v}}$ *is the vector of* l *shares of the secret* s *according to* $\prod$. *The share* $M_{\overrightarrow{v}}$ *belongs to party* $\rho(i)$.

Waters, B. shows [9] that every linear secret-sharing scheme according to the above definition also enjoys the linear reconstruction property, defined as follows: Suppose that $\prod$ is an LSSS for the access structure $\mathbb{A}$. Let $S \in \mathbb{A}$ be any authorized set, and let $I \subseteq \{1, \ldots, l\}$ be defined as $I = \{i : \rho(i) \in S$. There exist constants $w_i \in \mathbb{Z}_\rho$ with $i \in I$ that satisfy that if λ_i are valid shares of any secret s according to $\prod$, then $\Sigma_{i\in I} w_i\lambda_i = s$ holds. According to the access control requirements of the paid data-sharing scenario, our system utilizes LSSS to support any monotonic access structure.

2.5 | Blockchain Technology

Blockchain is a decentralized digital ledger composed of interconnected blocks. Each block stores data, including transaction details, with a cryptographic hash connected it to the previous block as illustrated in Figure 2, ensuring immutability and data integrity. This structure provides transparency and security as all network participants maintain a copy of the ledger, collectively validating and recording data. Blockchain can be public, allowing anyone to participate, or private, restricted to authorized members. The core characteristics of this technology are decentralization, transparency, security, and immutability, enabling applications across various domains. Blockchain technology can be categorized into hybrid, private, and public blockchains.

A public blockchain is characterized by its open nature, where any node can generate transaction data and access all information contained within the blocks. In contrast, a private blockchain is managed by a single entity, with only a limited array of authorized nodes accessing the stored records. A hybrid blockchain integrates elements from both public and private blockchains, enabling organizations to balance control with the transparency and security inherent in public blockchains. This approach provides a way to harness the benefits of both models while maintaining a degree of governance. In this work, we have employed Ethereum, an open-source platform designed to facilitate smart contracts and decentralized applications. Ethereum offers a robust framework for developing distributed applications. It leverages the Ethereum Virtual Machine to execute smart contracts. A defining characteristic of Ethereum is its ability to support smart contracts, allowing developers to design, deploy, and execute intricate applications on the platform. These contracts are self-executing, following predefined conditions,

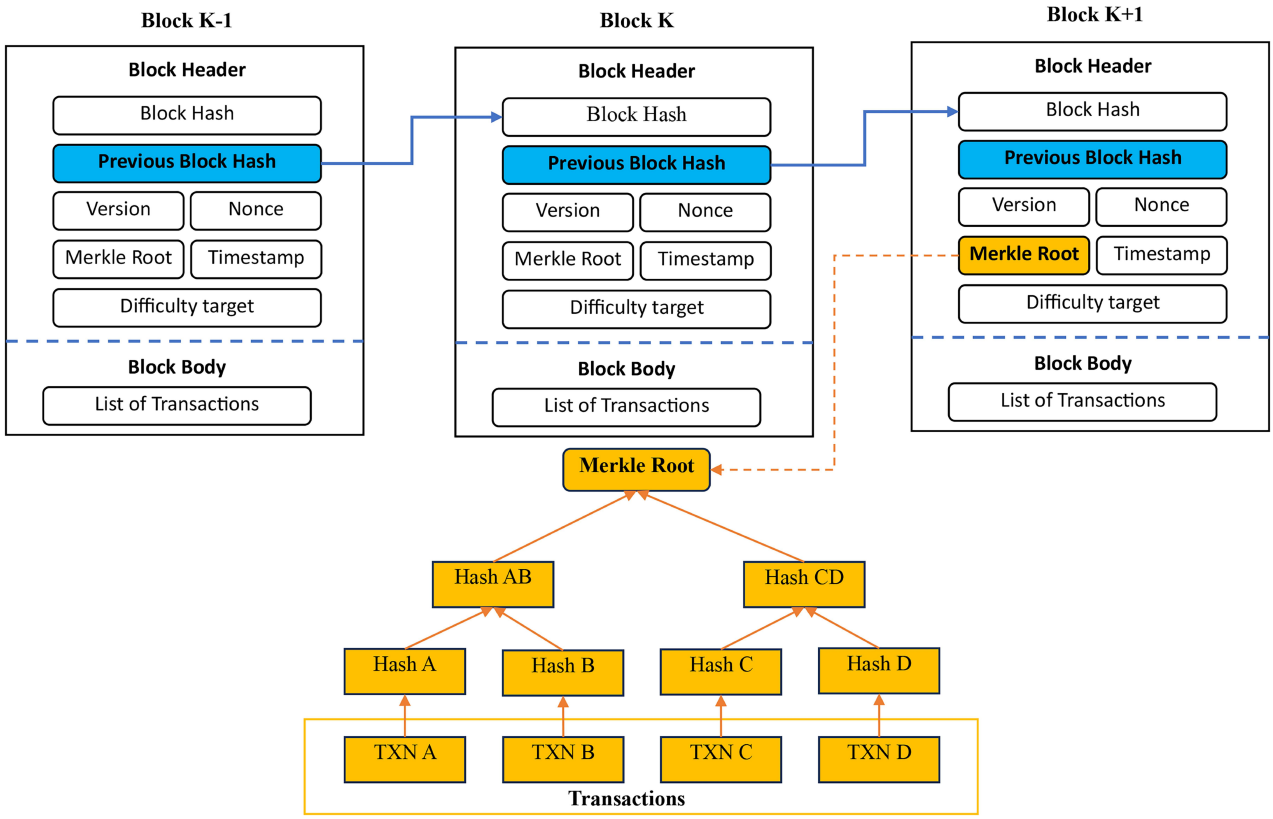


FIGURE 2 | Illustration of the block structure.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70064 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

which removes the reliance on intermediary entities, as detailed in Section 7.

3 | Related Works

In this section, we review the studies closely related to and relevant to our study. ABE, which Sahai et al. [10] introduced, has emerged as the dominant paradigm in the cryptographic access control for encrypted data. Bethencourt et al. [6] made a significant contribution to the field of CP-ABE by proposing a scheme that supports a tree-based access structure. This scheme is more adaptable and practical than KP-ABE. While ABE proposed by Bethencourt et al. [6], Waters, B. [9], Sahai et al. [10], and similar systems offer the advantages of one-to-many encryption, but inherent security concerns remain. Specifically, the single authority entrusted with key issuance to all eligible users can decrypt any ciphertext utilizing the distributed secret keys. Riepel et al. [11] implement FABEO, CP-ABE scheme, which has more concise ciphertexts, reduced pairings, and enhanced efficiency in both optimal and adaptive security scenarios.

A CP-ABE scheme that can support Non-Equivalence Logic (NEL) is developed by Yamada et al. [12] using a method for converting monotonic access structures into non-monotonic ones, which is originally proposed by Ostrovsky et al. [13]. Additionally, this CP-ABE scheme is capable of handling sets and access policies with unlimited sizes. This scheme exhibits greater complexity compared to conventional CP-ABE, resulting in increased overhead.

In the work of Chase, M. [14], a multi-authority CP-ABE (MCP-ABE) scheme is presented that leverages a central authority (CA). This CA possesses the capability to decrypt all ciphertexts within the system. This inherent decryption ability used by the CA introduces a vulnerability in the security position of decryption key storage.

Later, Chase et al. [15] introduced an MA-ABE scheme that operates without a centralized trusted authority. However, under this scheme, data consumers are required to obtain at least one attribute from each AA, thus limiting its practicality. In contrast, Lewko et al. [16] developed a CP-ABE scheme featuring fully decentralized multi-authorities, eliminating the necessity for attribute authorities to collaborate. Consequently, data consumers can acquire any desired number of attributes from any AA within this framework.

Yang et al. [17] address the critical challenge of data access control in cloud storage systems, particularly in the context of untrusted servers and the limitations of existing CP-ABE schemes. They propose a novel revocable multi-authority CP-ABE scheme that enhances expressiveness, efficiency, and security, allowing multiple authorities to independently issue attributes while ensuring effective attribute revocation with backward and forward security. However, the approach may face challenges related to the complexity of managing multiple authorities and the potential for increased overhead in communication and coordination. Additionally, while the proposed method improves efficiency by reducing the need for extensive ciphertext updates, it may still encounter scalability issues as the number of users and authorities grows.

To address the risks associated with semi-honest cloud servers, Miao et al. [18] propose a method that involves splitting secret keys associated with certain attribute sets. This approach allows healthcare providers and patients to securely exchange sensitive health information while only ensuring access is restricted to authorized individuals. The implementation challenges concerning scalability and the mechanism for attribute revocation related to the study require discussion.

Data consumers can leverage third-party servers for decryption computations to mitigate the computational burden associated with the decryption phase. Tu et al. [19] introduced a multi-authority CP-ABE scheme supporting decryption outsourcing. In this system, data consumers do not have to do any complicated bilinear pairing operations during the decryption phase. The security of ciphertext entrusted to a third-party custodian presents potential vulnerabilities to tampering, as documented by Guo et al. [20]. Blockchain technology mitigates this risk by demonstrably guaranteeing both data integrity and non-repudiation. These inherent cryptographic properties underpin the increasing adoption of blockchain across diverse application domains.

Jiang et al. [21] propose a data traceability generation method and a data access control scheme for edge clouds, ensuring EHRs integrity and confidentiality. The scheme uses ABE with a blockchain to provide secure data storage, fine-grained access control, and attribute revocation. However, the proof of correctness in the scheme is insufficient and may not satisfy the equations.

$$= \frac{e(C', K)}{e(\prod_{i\in I} C_i^{w_i}, L) \times \prod_{i\in I} e(D_i^{w_i}, K_{\rho(i)})}$$

$$= \frac{e(g^s, g^{\frac{a}{z}} g^t)}{e(\prod_{i\in I} (g^{a\lambda_i}(g^{V_{x_i}})^{-r_i})^{w_i}, g^t) \times \prod_{i\in I} e((g^{r_i})^{w_i}, (g^{V_{x_i}})^t)}$$

After simplifying the equation further, we get:

$$= \frac{e(g,g)^{\frac{s\alpha}{z}} e(g,g)^{st}}{e(g,g)^{ast}} \neq e(g,g)^{s\alpha/z} \quad where \sum_{i\in I} w_i\lambda_i = s$$

Leveraging advancements in cryptography, Guo et al. [22] sought to bolster user privacy and security through a novel encrypted data access control solution. Their architecture hinges on smart contracts orchestrating interactions between data owners (DOs), users, and attribute authorities. However, a key limitation of their approach lies in employing symmetric encryption by DOs for data encryption, resulting in a cumbersome key management overhead.

To alleviate this burden, Qin et al. [23] proposed an MA-ABE scheme that utilizes a consortium blockchain to establish trust bridges between attribute authorities. While their proposed solution effectively addresses the key management overhead, it also introduces certificate authorities as a potential point of vulnerability, raising concerns about single points of failure within the system's architecture.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70064 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

To enhance system security and address application requirements in resource-constrained environments, Wei et al. [24] proposed a revocable hierarchical scheme, RS-HABE. This scheme extends the original ABE by incorporating user revocation, secret key delegation, and ciphertext updating mechanisms. ABE has been widely applied in cloud services due to its rich expressiveness of ABE policies. However, traditional ABE schemes suffer from high computational overhead and lack verifiability, making them less efficient and less secure for cloud-based applications. To address this, Hou et al. [25] proposes a blockchain-based solution that improves decryption efficiency and allows users to verify the accuracy of decryption performed by external parties. This approach offers a more secure and efficient way to control access to data saved in the cloud.

The increasing reliance on cloud computing for managing PHRs has necessitated robust encryption methods to ensure data confidentiality. Zhang et al. [8] introduced ABE with keyword search, which addresses several issues related to key management and access control by enabling fine-grained access control. Enhancing usability in dynamic healthcare settings, the authors [2] introduce an innovative access control scheme known as PAFR-ABE, designed to safeguard users' identities during the secret key generation process, thereby thwarting unauthorized access to sensitive health data. The proposed framework incorporates a dynamic revocation system that enhances the effective administration of secret keys, thereby removing the necessity for updates for non-revoked users.

The authors [26] present a verifiable encryption scheme that incorporates blockchain technology to enhance security and accountability in data management. This scheme's multi-authorization mechanism enables various authorities to regulate user attributes, thereby improving flexibility and scalability in access control. The performance analysis indicates the scheme's effectiveness; however, empirical evidence is required to validate its efficiency and overall impact.

From the above literature, the research gaps are found:

- Some traditional ABE schemes rely on a central authority for key issuance, which can pose significant risk.
- Existing schemes often struggle with effective attribute revocation, particularly in dynamic environments where users frequently change roles or permissions.
- Conventional storage systems frequently depend on centralized designs that are susceptible to manipulation, lack immutability, and fail to offer comprehensive audit trails.

Based on the preceding discussion, this research paper introduces a secure access control mechanism utilizing blockchain and ABE. The proposed approach leverages the InterPlanetary File System (IPFS) for storing encrypted data, thereby mitigating blockchain resource limitations and centralized storage systems' reliability concerns. This method not only enhances security and adaptability but also facilitates attribute policy concealment and attribute revocation.

4 | Blocpabe'S Basic Construction

This section presents a concise overview of our proposed framework. The system model depicted in Figure 3 describes a secure

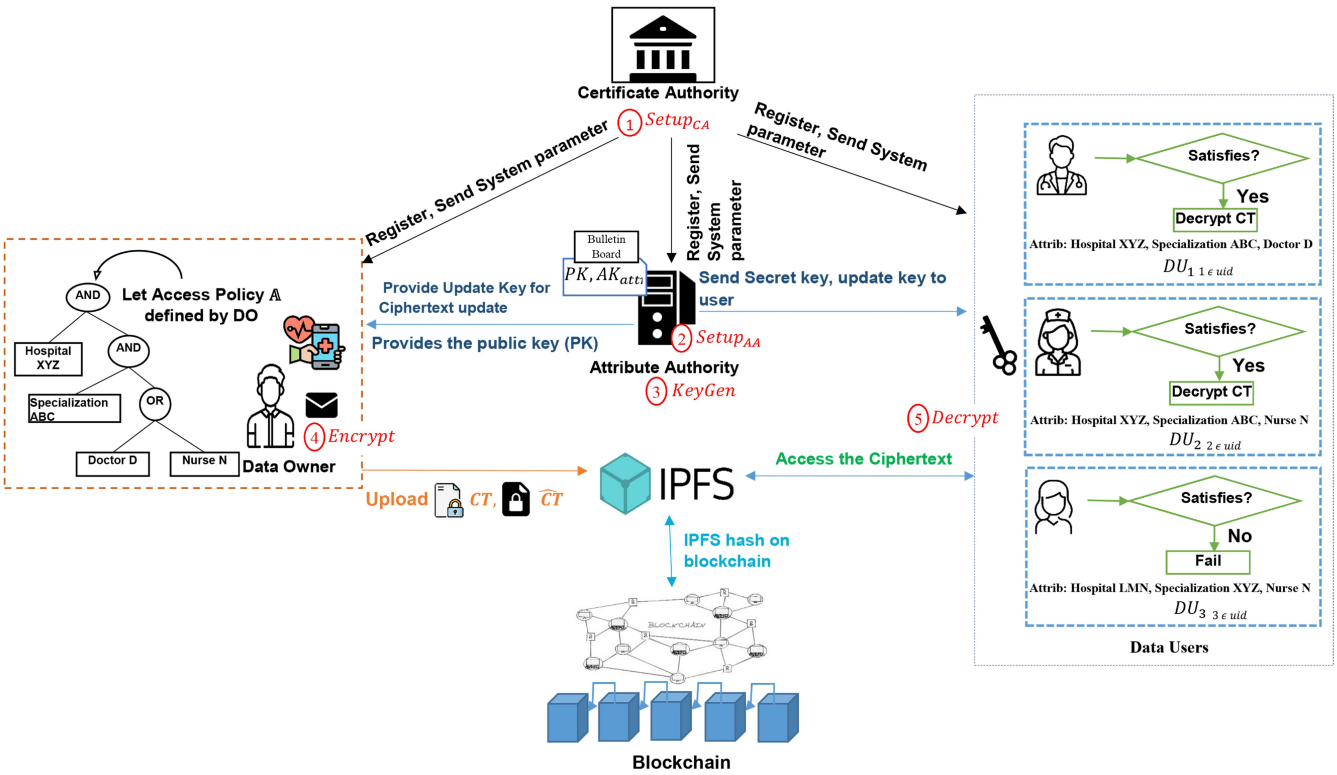


FIGURE 3 | Overview of system model.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70064 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and decentralized data-sharing framework that integrates ABE with IPFS and blockchain technology to enforce fine-grained access control over encrypted data. At the core, the Certificate Authority initializes the system by generating system parameters and registering participants, such as the Attribute Authority (AA) and Data Users. The AA is responsible for managing attributes and keys, including generating public keys and issuing secret keys to users based on their assigned attributes. A Data Owner (DO) defines an access policy that determines who can access the data. The DO encrypts the data using ABE under the policy and uploads the ciphertext to the IPFS for decentralized storage. To ensure data integrity and tamper-proof access, the IPFS hash of the ciphertext is stored on the blockchain. The proposed system model outlines the operational framework of the scheme, providing a detailed evaluation of its components as follows.

4.1 | System Model

Here, we have described all the participating entities and their contribution to the system model below:

- **Certificate Authority (CA):** CA operates as a globally recognized root of trust. Its primary function is to establish and manage identities within a system. It assigns unique identifiers to the users and authority. Additionally, the CA assumes responsibility for system initialization, including user and attribute authority registration and disseminating global public parameters and a master secret key. While possessing a privileged role, the CA is functionally constrained, acting as an "honest-but-curious" entity with no direct involvement in attribute management or generating attribute-associated secret keys.
- **Attribute Authority (AA):** AA exercises exclusive control over the management of attributes. It encompasses the issuance, updating, and revocation of attributes. AA is responsible for constructing the secret key for the users.
- **IPFS**: IPFS is a decentralized file storage network that uses content addressing to identify and retrieve files. Instead of traditional location-based addressing, where files are accessed by their storage location, IPFS retrieves files based on a unique hash derived from their content. The encryption process generates ciphertext and associated metadata, including a policy string that governs access control.
  The encrypted data is then serialized into JSON format before being uploaded to IPFS. IPFS, a decentralized peer-to-peer file system, assigns a unique content identifier (CID) to the serialized data. This CID is a reference for retrieving the data from the IPFS network, ensuring its availability across a distributed network of nodes.
- **Blockchain:** Blockchain is a decentralized and immutable ledger designed for recording transactions and tracking assets. The IPFS CID is recorded on the blockchain through a smart contract. This contract stores metadata related to the encrypted data, such as its IPFS hash, ensuring the data integrity and existence are verifiable. By recording this metadata on the blockchain, we leverage the blockchain's immutability to guarantee that the reference to the data remains accurate and tamper-proof.
- **Data Owner (DO):** In the system, DO is responsible for storing data on the decentralized storage and sharing with data consumers or users. The owner defines an access policy, encrypts the data according to this policy, and then sends the encrypted data to the centralized storage for secure storage and controlled access.
- **The Data User (DU):** DU is the data consumer who wants to get the data. CA assigns the UID to the user. If the attributes of the DU align with the stipulated access policy, the message can be successfully retrieved.

4.2 | Algorithm Description

Our BloCPABE framework is structured into distinct phases, each incorporating specific methods as briefly illustrated below:

Phase 1: Setup

- $Setup(1^{\lambda})$: 1^{λ} here is a security parameter which $Setup$ algorithm takes as input and generates global public parameter GPP and master key MK as output. Meanwhile, authority and users are registered by the CA, and AA also performs the setup to generate public (PK) and secret keys (MSK).

Phase 2: Secret Key Generation

- $KeyGen(GPP, MSK, S_k)$: The algorithm takes the GPP, MSK, and attribute assigned to the user and generates the corresponding secret key (SK_{uid_k}).

Phase 3: Encryption of message

- $Encrypt(GPP, PK, \mathbb{A}, msg)$: The algorithm requires the GPP, public key, an LSSS access structure $\mathbb{A}$ and message msg as input and generates ciphertext CT.

Phase 4: Decryption Process

- $Decrypt(CT, USK'_{uid_k}, SK_{uid_k})$: The algorithm takes CT, the global secret key of user, and SK_{uid_k} as input and returns the message.

Phase 5: Attribute Revocation

- $Revoke(\hat{x}, uid_k)$: This algorithm includes the update keygen algorithm ($UKeyGen$), secret key update ($Update_{SK}$), and ciphertext update algorithm ($Update_{CT}$). The attribute $\hat{x}$ of user uid_k is revoked, and a new secret key and ciphertext are generated for the non-revoked users.

4.3 | Threat Model and Security Goals

CA and AA are assumed to be fully trusted. In the case of a malicious DU, unauthorized DUs might collude to gain access to the encrypted data.

- **Fine-grained access control:** To control access for retrieving health data, the data owner creates a detailed access structure based on specific attributes. This framework guarantees that only authorized DUs have access to the encrypted

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70064 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

information. Unauthorized DUs, regardless of how they combine their attributes and keys, cannot extract confidential details from the encrypted data.

- **Confidentiality:** In this system, data access is restricted to devices that meet specific criteria, and only these authorized devices can receive a unique decryption key.
- **Non-repudiation:** It states that the origin of data cannot be denied. It provides irrefutable proof that a specific person or entity sent or created particular information. It is crucial in digital transactions and communications as it prevents parties from disavowing their actions or commitments.
- **Collusion resistance:** Malicious users could pool their decryption keys to access protected data that none can open independently. The system should be robust enough to prevent such collaborative attacks.
- **Data integrity:** The system must rigorously safeguard the integrity, consistency, and accuracy of data stored on the blockchain from its creation to its final state. Once data is recorded on a blockchain, cannot be altered or deleted without the network's consensus. It creates an unbreakable chain of records, making it virtually impossible to tamper with data.
- **Policy Privacy:** The access structure embedded within encrypted data should remain hidden from blockchain and unauthorized users to ensure data privacy.

4.4 | Security Model

Here, we provide a security model structured as an interactive game between the two entities here known as challenger $\mathfrak{C}$ and adversary $\mathfrak{A}$, following the framework from Waters, B. [9]. The security process unfolds as follows:

Setup: The *Setup* algorithm is run to produce the parameters of the system. The $\mathfrak{C}$ uses the $Setup_{AA}$ algorithm to formulate the PK and the MSK. The challenger keeps the MSK and shares the public key PK with the $\mathfrak{A}$.

*Phase*1: The $\mathfrak{A}$ can make adaptive queries to the $\mathfrak{C}$ to obtain secret keys. Following these queries, $\mathfrak{C}$ executes the *KeyGen* algorithm and provides the corresponding secret key to $\mathfrak{A}$.

Challenge: The adversary $\mathfrak{A}$ presents two messages, msg_0 and msg_1, of equal length, along with a challenge access structure $\mathbb{A}^*$ that cannot be satisfied by any of the attributes inquired in *Phase*1. $\mathfrak{C}$ flips a coin $\zeta \in 0, 1$ and generates a ciphertext $CT^* \leftarrow Encrypt(GPP, PK, msg_\zeta, \mathbb{A}^*)$, which is then given to $\mathfrak{A}$.

*Phase*2: The $\mathfrak{A}$ continues to request secret keys from $\mathfrak{C}$ under the same conditions as in *Phase*1.

Guess: The $\mathfrak{A}$ guesses the value of ζ as ζ'. The adversary's advantage is calculated as $|Pr[\zeta = \zeta'] - \frac{1}{2}|$.

Definition 1. Blo*C*ABE is secure if all polynomial-time adversaries have at most a negligible advantage in the above game.

5 | Concrete Construction

Section 5 provides a thorough construction of the BloCPABE scheme. BloCPABE is constructed by taking inspiration and combining the idea of Waters, B. [9] and Yang et al. [17]. The scheme consists of algorithms such as system initialization ($Setup_{CA}, RegisterUser, Setup_{AA}$), $KeyGen$, $Encrypt$, $Decrypt$, $UKeyGen$, $Update_{SK}$, and $Update_{CT}$ (illustrated in Figure 4).

5.1 | Proposed Blocpabe Scheme

Phase 1: System Initialization (Setup): In this phase, algorithms such as $Setup_{CA}$, $RegisterUser$, and $Setup_{AA}$ are executed.

- **Global Initialization** ($Setup_{CA}$): In this sub-algorithm, the CA utilizes a security parameter, denoted as λ, which dictates the length of the input group. The CA selects two multiplicative cyclic groups, $\mathbb{G}_1$ and $\mathbb{G}_T$, each of prime order p. Here, g serves as the generator of $\mathbb{G}_1$, and $e : \mathbb{G}_1 \times \mathbb{G}_1 \rightarrow \mathbb{G}_T$ represents a bilinear map. CA selects a collusion-resistant hash function $\mathcal{H}$, which matches the binary string $\{0,1\}^*$ to the random element in $\mathbb{G}_1$, expressed as $\mathcal{H} : \{0,1\}^* \rightarrow \mathbb{G}_1$. CA randomly selects two numbers $a, b \in_R \mathbb{Z}_p$, constructs the $MK = (a, b)$, and publishes GPP to all the system's members.
  $GPP = (g, g^a, g^b, \mathcal{H})$.
  Furthermore, both AA and users are required to undergo registration with CA to acquire globally unique identifiers uid_k (for users) and aid (for AA) that validate their legal identities.
- ***RegisterUser***: Here in our system, DU and DO are referred to as users (with different roles) (Subsection 4.2). It is assumed that CA runs the $RegisterUser$ algorithm within a secure environment, ensuring the confidentiality of all transmitted data. Upon system initialization, each user submits a registration request to the CA. The CA verifies the user's identity and subsequently assigns a uid_k and also creates $Certificate(uid_k)$ for uid_k. The identifier uid_k is a randomly generated value within the prime field $\mathbb{Z}_p$, that is, $uid_k \in \mathbb{Z}_p$. After this, it randomly chooses the two numbers $\gamma_k, \gamma'_k \in_R \mathbb{Z}_p$ to generate the public/private key pair as $UPK_{uid_k} = g^{\gamma_k}, UPK'_{uid_k} = g^{\gamma'_k}$ and $USK_{uid_k} = \gamma_k, USK'_{uid_k} = \gamma'_k$.
- **Attribute Authority Setup** ($Setup_{AA}$): This sub-algorithm is managed by an AA to generate the MSK and PK. The AA within the system manages a universe of attributes $\mathcal{U}_{aid}$ and for each attribute $x_i \in_R \mathcal{U}_{aid}$, AA generates the public attribute key AK_{x_i} along with the version key VK_{x_i}. AA chooses two random exponents $\alpha, y \in \mathbb{Z}_p$ and compute the PK and a secret key MSK as:

  $$PK = e(g, g)^\alpha, MSK = (\alpha, y).$$

  For each attribute $x_i : (x_i \in \mathcal{U}_{aid})$, AA chooses version key as $VK_{x_i} = \upsilon_{x_i}$ and get the corresponding attribute key $AK_{x_i} = (g^{\upsilon_{x_i}} \mathcal{H}(x_i))^y$. Attributes are kept confidential and known only to AA, whereas PK and AK_{x_i} are published publicly with system entities.

Phase 2: Key generation (PHR user authorization): When the user seeks access to the data, it requests the secret key SK_{uid_k}

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70064 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

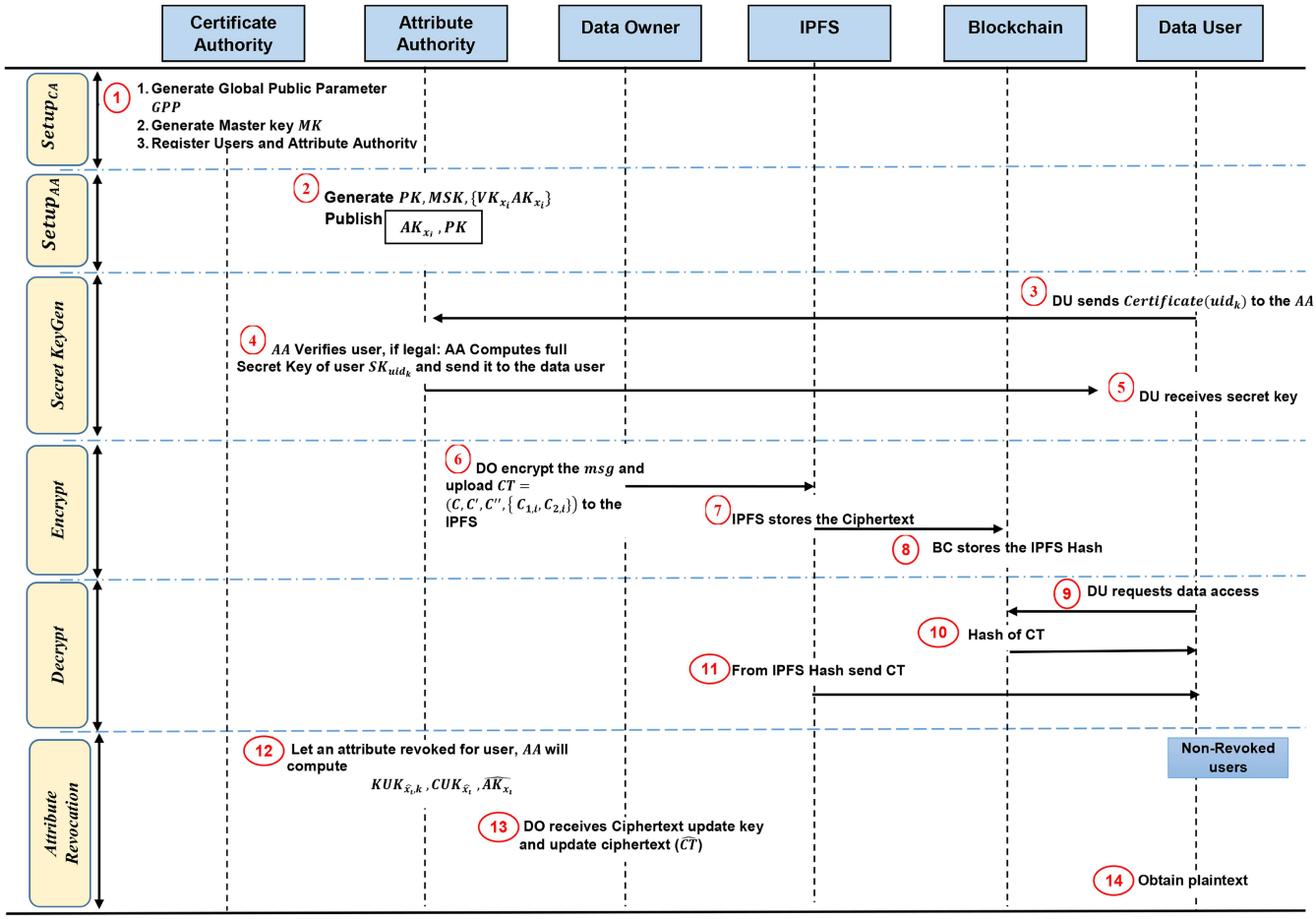


FIGURE 4 | Illustration of the proposed scheme.

from the AA. AA assigns a set of attributes $S_k \in \mathcal{U}_{aid}$ to the legal user according to its identity and role in the system. Subsequently, uid_k gives $Certificate(uid_k)$ to the AA, and upon verifying the user's identity, the authority executes the $KeyGen$ algorithm. It selects $t_k \in_R \mathbb{Z}_p$ and for each user uid_k it generates

$$K_k = g^{\alpha} g^{a\gamma_k} g^{bt_k},$$

$$K'_k = g^{t_k}, K'' = g^{\gamma t_k},$$

$$\forall x_i \in S_k \ : \ K_{x_i,k} = g^{\gamma' {}_k \gamma t_k} . (g^{\upsilon_{x_i}} \mathcal{H}(x_i))^{\gamma \gamma_k}$$

$$SK_{uid_k} = \left\{ K_k, K'_k, K''_k, \{K_{x_i,k}\}_{x_i \in S_k} \right\}$$

Phase 3: Encryption Process (Outsourcing PHR on decentralized storage): Access to a data owner PHR is granted to authorized healthcare professionals, including physicians and nurses, upon satisfaction of specific conditions. Before outsourcing data to the decentralized storage, the PHR owner encrypts the data using an encryption algorithm. As discussed in Subsection 4.2, it takes the public parameter, message, public keys of relevant authorities, and LSSS access structure. (M, ρ) is an access structure where M is a matrix with $l \times n$ elements, and ρ maps the rows of M into the attributes set. The DO subsequently selects a random column vector $\vec{\upsilon} = (s, \upsilon_2, \upsilon_3, \ldots, \upsilon_n)$, where $\upsilon_2, \upsilon_3, \ldots, \upsilon_n$ are applied to share s. DO computes $\lambda_i = M_i . \vec{\upsilon}$ for $\forall i \in [1, l]$, where M_i is a vector corresponding to the ith row of matrix M. DO randomly choose $\tau_1, \tau_2, \ldots, \tau_l \in \mathbb{Z}_p$ and compute the CT as

$$C = msg.e(g, g)^{\alpha s}$$

$$C' = g^s, C'' = g^{bs},$$

$$\forall i \in [1, l] \ :$$

$$C_{1,i} = g^{a\lambda_i} . ((g^{\upsilon_{\rho(i)}} \mathcal{H}(\rho(i)))^y)^{-\tau_i}, C_{2,i} = g^{\tau_i}, \rho(i) \in \mathcal{U}_{aid}$$

$$CT = \{C, C', C'', \{C_{1,i}, C_{2,i}\}\}$$

Phase 4: Decryption Process (Accessing the PHR): To access the encrypted data, a DU submits a request to the decentralized storage. Those DUs that meet the access policy defined in the CT can decrypt the information. The DU runs the local decryption algorithm as briefly illustrated in Subsection 4.2.

Case 1: If user attributes necessary of decryption $\notin \mathbb{A}$ then $\perp$.

Case 2: If user attributes necessary of decryption $\in \mathbb{A}$ then algorithm recovers the msg as: $I = \{i : p(i) \in \mathcal{U}_{aid}\} \subseteq [1, \ldots, l]$ and compute the coefficient $\{w_i \in \mathbb{Z}_p\}_{i \in I}$ so that $\Sigma_{i \in I} w_i \lambda_i = s$. The decryption algorithm computes

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70064 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$\frac{e(C', K_k)e(C'', K'_k)^{-1}}{\prod_{i\in I}\left(e(C_{1,i}, UPK_{uid_k})e(C_{2,i}, K_{\rho(i),k})e(C_{2,i}, K''^{USK'_{uid_k}}_k)^{-1}\right)^{w_i N_A}}$$

to obtain the original message (msg), the user computes

$$msg = \frac{C}{e(g,g)^{\alpha s}}$$

Phase 5: Attribute Revocation: When an attribute $\hat{x}_i$ is revoked for user uid'_k by AA, the system should implement a mechanism to invalidate the user's ability to decrypt ciphertexts containing that attribute. It prevents the revoked user from accessing PHRs protected by policies that require the revoked attribute. For users whose attributes remain valid, AA should update their secret key to maintain system security and integrity. The attribute revocation of our work is somewhat similar to the Yang et al. [17].

1. *Update Key by AAs*: When $\hat{x}_i$ is revoked from the user, the AA runs the $UKeyGen$ algorithm as illustrated in Subsection 4.2. It firstly chooses a random versions key $v'_{\hat{x}_i}$ and generate the attribute update key $AUK_{\hat{x}_i} = y\left(v'_{\hat{x}_i} - v_{\hat{x}_i}\right)$, then AA compute the $KUK_{\hat{x}_i,k}$ and $CUK_{\hat{x}_i}$.
$KUK_{\hat{x}_i,k} = g^{\gamma_k.y(v'_{\hat{x}_i} - v_{\hat{x}_i})}$
$CUK_{\hat{x}_i} = -\left(y\left(v'_{\hat{x}_i} - v_{\hat{x}_i}\right)\right)$
The AA update the attribute key of revoked attribute $\hat{x}_i$ by computing
$\widehat{AK}_{\hat{x}_i} = AK_{\hat{x}_i}.g^{AUK_{\hat{x}_i}}$
and send the message about $\hat{x}_i$ updates to the data owner under the safe channel.

2. *Secret Key Update by Non-revoked Users*: Upon receiving the $KUK_{\hat{x}_i,k}$ from the AA, user will run the $Update_{SK}$ algorithm to update the secret key as
$\widehat{SK}_{uid_k} = (\widehat{K}_k = K_k, \widehat{K'_k} = K'_k, \widehat{K''_k} = K''_k,$
$\widehat{K}_{\hat{x}_i,k} = K_{\hat{x}_i,k}.KUK_{\hat{x}_i,k},$
$\forall x_i \in S_k, x_i \neq \hat{x}_i : \widehat{K}_{x_i,k} = K_{x_i,k})$
The $KUK_{\hat{x},k}$ is associated with revoked uid_k so that revoked users can easily be identified and differentiated. Consequently, revoked users cannot obtain the updated key.

3. *Ciphertext Update by the Data Owner*: Upon receipt of the key $CUK_{\hat{x}}$ from the AA, the owner will execute the $Update_{CT}$ algorithm to refresh the ciphertext. Initially, the data owner gathers the ciphertext components $(C_{1,i}, C_{2,i})$ that includes the attribute $\hat{x}$ from the blockchain. The updated ciphertext is then calculated as follows:
$\widehat{CT} = (C = msg.e(g,g)^{\alpha s}, C' = g^s, C'' = g^{bs},$
$\forall i \in [1,l]:$
if $\rho(i) = \hat{x}_i : \widehat{C}_{1,i} = C_{1,i}.(C_{2,i})^{CUK_{\hat{x}_i}},$
if $\rho(i) \neq \hat{x}_i : \widehat{C}_{1,i} = C_{1,i}, \widehat{C}_{2,i} = C_{2,i}).$

5.2 | Correctness Proof

Definition 2. Our scheme is correct because the following equations are satisfied.

Proof.

$$\frac{e(C', K_k)e(C'', K'_k)^{-1}}{\prod_{i\in I}\left(e(C_{1,i}, UPK_{uid_k})e(C_{2,i}, K_{\rho(i),k})e(C_{2,i}, K''^{USK'_{uid_k}}_k)^{-1}\right)^{w_i N_A}}$$
$$= e\left(C', K_k\right) e\left(C'', K'_k\right)^{-1}$$
$$= e\left(g^s, g^{\alpha} g^{a\gamma_k} g^{bt_k}\right).e\left(g^{bs}, g^{t_k}\right)^{-1}$$
$$= e(g,g)^{s\alpha}.e(g,g)^{sa\gamma_k}.e(g,g)^{sbt_k}.e(g,g)^{-bst_k}$$

canceling

$$e(g,g)^{sbt_k}.e(g,g)^{-bst_k}$$

we get,

$$= e(g,g)^{sa\gamma_k}.e(g,g)^{s\alpha}$$

for each $i \in I$, assume $p(i) \in \mathcal{U}_{aid}$:

$$= \prod_{i\in I}\left(e\left(C_{1,i}, UPK_{uid_k}\right) e\left(C_{2,i}, K_{p(i),k}\right) e\left(C_{2,i}, K''^{USK'_{uid_k}}_k\right)^{-1}\right)^{w_i N_A}$$

simplifying the equation:

$$= e(g^{a\lambda_i}.((g^{v_{p(i)}}\mathcal{H}(p(i))^y)^{-\tau_i}, g^{\gamma_k}).e(g^{\tau_i}, g^{\gamma'_k yt_k}.$$
$$(g^{v_{p(i)}}\mathcal{H}(p(i)))^{y\gamma_k}).e\left(g^{\tau_i}, g^{yt_k\gamma'_k}\right)^{-1}$$
$$= e\left(g,g\right)^{a\lambda_i\gamma_k}.e\left(g, g^{v_{p(i)}}\mathcal{H}(p(i))\right)^{-\gamma_k y\tau_i}.e\left(g,g\right)^{\tau_i\gamma'_k yt_k}.$$
$$e\left(g, g^{v_{p(i)}}\mathcal{H}(p(i))\right)^{\tau_i y\gamma_k}.e\left(g,g\right)^{-\tau_i yt_k\gamma'_k}$$
$$= e(g,g)^{a\lambda_i\gamma_k}$$

after we get,

$$\prod_{i\in I}\left(e(g,g)^{a\lambda_i\gamma_k}\right)^{w_i N_A} = e(g,g)^{sa\gamma_k}, \quad \Sigma_{i\in I} w_i\lambda_i = s,$$

then it computes,

$$\frac{e(g,g)^{sa\gamma_k}.e(g,g)^{\alpha s}}{e(g,g)^{sa\gamma_k}} = e(g,g)^{\alpha s}$$

to obtain the message

$$= \frac{C}{e(g,g)^{\alpha s}}$$
$$= \frac{msg.e(g,g)^{\alpha s}}{e(g,g)^{\alpha s}}$$
$$= msg$$

□

6 | Security Analysis

Here in this Section 6, a detailed security analysis of BloCPABE is discussed:

Theorem 1. *The proposed system model is protected against repudiation attack.*

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70064 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Proof. By permanently recording data in a tamper-proof manner and verifying the identity of transaction participants through digital signatures, blockchain ensures that actions cannot be denied. □

Theorem 2. *BloCPABE ensures the integrity and resistance to tampering.*

Proof. BloCPABE ensures the integrity and resistance to tampering with encrypted data by utilizing the inherent properties of blockchain technology. Since the hash of encrypted data is stored on a blockchain, any attempts to tamper with or alter the data would require a consensus among the majority of nodes on the network, making it virtually impossible to manipulate the data without being detected. By its very design, blockchain is a decentralized and immutable ledger that records transactions across a network of computers. Once data is entered into the blockchain, it is cryptographically linked to the previous blocks, forming a chain resistant to modification. Any attempt to alter a previous block would require recalculating and altering all subsequent blocks, which is computationally infeasible in a well-distributed network. □

Theorem 3. *Our scheme protects and maintains the confidentiality of data, ensuring access is limited to authorized users only.*

Proof. BloCPABE facilitates robust, fine-grained access control, ensuring that sensitive data is securely delivered to authorized users while protecting it from unauthorized access. It guarantees that only users who satisfy the predefined access policy can access the encrypted data. Furthermore, the data owner encrypts the data and stores it in decentralized storage while sharing the corresponding secret key with authorized users through a secure channel, preventing potential adversaries from accessing the plaintext. □

Theorem 4. *The proposed framework is resistant to the collusion attack.*

Proof. In scenarios where multiple users collaborate, they might potentially decrypt the ciphertext by combining the attributes. When malicious users collude, they may attempt to combine their attributes to decrypt the ciphertext. However, since the secret key is linked to each user's unique identifier uid, collusion between two users with different identifiers uid and uid' will not enable them to recover the factor $e(g,g)^s$. Consequently, such collusion will not allow these users to gain unauthorized privileges. □

Theorem 5. *The scheme provides both backward and forward security.*

Proof. Forward security guarantees that once a user's decryption key (secret key) is revoked, they can no longer access newly encrypted data. Backward security, on the other hand, ensures that a newly generated decryption key cannot be used to access ciphertext that was encrypted prior to the key's creation. It prevents a user from accessing sensitive information that was encrypted before they were granted access, thereby maintaining the confidentiality of historical data. The proposed scheme incorporates attribute revocation to guarantee both backward and forward security. Upon attribute revocation, only users possessing valid attributes can update their secret keys to the latest version. Additionally, ciphertexts are updated using a ciphertext update key, preventing newly added users from accessing previously encrypted data. □

Theorem 6. *BloCPABE scheme is IND-CPA secure if the Decisional q-PBDHE assumption maintains, that is, No polynomial-time adversary $\mathfrak{A}$ attempting to selectively break BloCPABE system with a challenge matrix of size $l^* \times n^*$ will succeed if the Decisional q-PBDHE assumption holds.*

Proof. The details of the proof are referred to Waters, B. [9]. □

6.1 | Formal Verification

To perform the formal verification of our protocol, we have employed the Scyther Tool [27] as shown in Figure 5. Other than Scyther, there are numerous tools such as AVISPA [28], Proverif [29], and TAMARIN [30] that are used for the formal security analysis [31]. The Scyther tool analyzes a protocol described in Security Protocol Description Language (SPDL) and formulates security claims for each role specified in the protocol. It then verifies that these assertions are accurate; in other words, it checks the correctness of these claims. Scyther tool is designed to function with an unlimited number of sessions, though this capability does not ensure indefinite correctness. In situations where this limitation applies, Scyther will indicate, "No attacks within bounds." If the tool identifies no specific attacks or attack patterns, it will report "No attacks."

The implementation code of the entire BloCPABE protocol is very large, but to provide a brief understanding of the protocol, a code snippet illustrating the system initialization of BloCPABE is given in Figure 5a. The code snippet showcases the exchange of messages between the CA, AA, and users (where "users" refers to both DO and DU). Each entity exchanges data using the *send* and *recv* actions. The *send* action describes a participant in the protocol sending a message to another participant, while the *recv* action implies that a participant is prepared to accept an incoming message.

The implementation of the BloCPABE protocol in Scyther contains five participants: CA, AA, DU, DO, and BC. Here, the Scyther tool also assesses claims such as $Weakagree$, which verifies that all parties involved are actively executing the protocol. $Nisynch$ ensures that data is handled consistently throughout the execution of the protocol. $Niagree$ guarantees that the sender believes the intended receiver actively participates in the one-way authentication protocol. $Alive$ confirms the receiver's active status after the protocol's execution and verifies that they have initiated the expected response. Finally, $Commit$ confirms that the protocol has successfully established a valid claim.

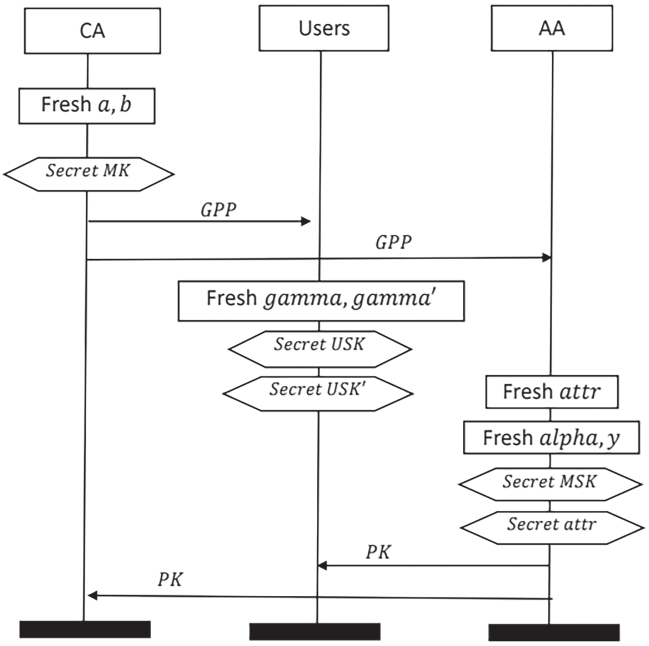
The security is evaluated over five runs with adversarial parameters to determine the most suitable attack, with a maximum of 10 patterns per claim, as shown in Table 2.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70064 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

```
protocol BloCPABE(CA,Users,AA){
role CA{
claim(CA,Secret,MK);
send_1(CA,Users,GPP);
send_2(CA,AA,GPP);
recv_4(AA,CA,PK);
}

role Users{
claim(Users,Secret, USK);
claim(Users,Secret, USK');
recv_1(CA,Users,GPP);
recv_3(AA,Users,PK);
}
role AA{
claim(AA,Secret,MSK);
claim(AA,Secret,attr);
recv_2(CA,AA,GPP);
send_3(AA,Users,PK);
send_4(AA,CA,PK);
}
}
```

(a)



(b)

FIGURE 5 | Formal verification of the system initialization phase: (a) a code snippet of the setup phase and (b) a sequence diagram.

TABLE 2 | Parameter setting in Scyther.

Parameters	Values
Maximum number of runs	5
Matching type	Typed matching
Search Pruning	Find best attack
Maximum number of patterns per claim	10
Attack graph font size	14

In the setup phase (system initialization phase), we investigate the attacks between the transmission from the CA to users, from the CA to the AA, and from the AA to the users. The *Secret* claim ensures the protocol parameters remain confidential and protected from unauthorized observers or intruders. Scyther evaluates the following protocol claims: $Claim(CA, Secret, MK)$ for the CA, $claim(Users, Secret, USK)$ and $claim(Users, Secret, USK')$ for the users, and $claim(AA, Secret, MSK)$ and $claim(AA, Secret, attr)$ for the AA. MK is the master key, and is defined in blocpabe.spdl as $macro\ MK = keys(a, b)$, where $keys$ is a function. Similarly, USK and USK' defined as $macro\ USK = gamma$ and $macro\ USK' = gamma'$. MSK is defined here as $macro\ MSK = keys(alpha, y)$. The sequence diagram Figure 5b illustrates that role CA performs a *send* action to send message/data to Users using $send_1(CA, Users, GPP)$ and to AA using $send_2(CA, AA, GPP)$. The Users claim the secrecy of USK, USK'. The role AA sends the message to the CA and the Users with $send_3(AA, Users, PK)$ and $send_4(AA, CA, PK)$ respectively, where PK is defined as $macro\ PK = map(g, g)power(alpha)$. The SPDL language is used to implement our suggested scheme, and the verification protocol results (Table 3) have confirmed its correctness.

7 | Performance Analysis

This section presents a theoretical analysis that includes a comparison of features or functionalities, along with an evaluation of computation (Subsubsection 7.1.2), communication (Subsubsection 7.1.3), and storage (Subsubsection 7.1.4) overhead of relevant existing schemes. Additionally, it offers an experimental analysis of encryption and decryption times across varying numbers of attributes.

7.1 | Theoretical Analysis

7.1.1 | Features Comparison

To evaluate the scalability of our scheme, we conducted a comparative analysis with several established CP-ABE schemes. We have assessed the functionality of BloCPABE in comparison with the schemes described in [6], [17], [24], [1], [23], [19], [5], [11], [12] and [25], as detailed in Table 4.

As outlined in Table 4, we conduct an analysis based on several key aspects, including type, group type, and the access structure utilized. It also encompasses the level of renovation, whether at the attribute or user level, as well as support for key updates and ciphertext updates. Furthermore, the analysis considers who is responsible for updating the ciphertext, the provision of non-repudiation, and protection against collusion attacks. We also assess whether the scheme ensures both forward and backward security, integrates blockchain technology for immutability and data integrity, and provides formal verification to confirm the absence of potential attack vectors. A "✓" indicates support for the technology, a "✗" indicates a lack of support, and "×" denotes that it is not explicitly mentioned.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70064 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 | Scyther verification results.

Claim				Status	Comment
BloCPABE	CA	BloCPABE, CA1	Secret keys(a,b)	Ok Verified	No attacks
	Users	BloCPABE, Users1	Secret gamma	Ok Verified	No attacks
		BloCPABE, Users2	Secret gamma	Ok Verified	No attacks
	AA	BloCPABE, AA1	Secret keys(alpha,y)	Ok Verified	No attacks
		BloCPABE, AA2	Secret attr	Ok Verified	No attacks
Done					

This comparison focuses on various functional aspects critical for the practical deployment of these schemes in sensitive data-sharing environments. All the considered schemes are of the CP-ABE type and utilize the LSSS access structure, enhancing each scheme's expressiveness. Bethencourt et al. [6] scheme utilizes the tree access structure, and Riepel et al. [11] scheme uses the boolean formulae and MSP access structure. The scheme in [6], [17], [12], [24], [23], [19], [11], [5], [25] and our take prime order group, while [1] adopts composite order group.

In terms of revocation mechanisms, the schemes proposed by Yang et al. [17], Tu et al. [19], and our system support attribute-based revocation, while Wei et al. [24] scheme implements user-based revocation. Other schemes do not incorporate any form of revocation. Regarding key and ciphertext updates, the scheme [17], [24], [19], and our facilitate both the key update and ciphertext update. The server handles the ciphertext update in [17], [24], and [19], while the owner is responsible for ciphertext updates in our scheme. When evaluating security characteristics, such as user collusion, forward/backward security, and non-repudiation, the schemes by [5], [25], and our scheme support non-repudiation and leverage blockchain technology for the secure storage of encrypted data, ensuring data integrity. The schemes [17], [24], [19] and our support the forward and backward security as well. Additionally, [17], [23], [19], and our scheme provides protection against user collusion. Furthermore, our scheme is formally verified using the Scyther tool.

7.1.2 | Computation Overhead

Before assessing the computational overhead, we have examined of the execution time for each operation as shown in Table 5. The duration for each operation has been estimated utilizing the Charm framework [32] on a supersingular curve SS512. The analysis presented in Table 5 indicates that hash operations, exponentiation in $\mathbb{G}_1$ and $\mathbb{G}_T$, and pairing operations are significant contributors to the computational cost associated with ABE schemes. Further, Table 6 illustrates the computational overhead across different schemes, highlighting the costs associated with key generation, encryption, and decryption processes. We include the following schemes in our evaluation: Yang et al. [17], BSW07 [6], FABEO [11] and Yamada et al. [12]. Let T_E denote the time required for exponentiation operations in $\mathbb{G}_1$ and $\mathbb{G}_T$. Specifically, T_E is measured as 0.82201ms in $\mathbb{G}_1$ and 0.07321ms in $\mathbb{G}_T$. The $|S|$ represents the number of attributes possessed by user uid_k, which are utilized in constructing uid_k's secret key. Additionally, let T_P signify the time for pairing operations and H' the time for hash operations. The m denotes the number of attributes in the ciphertext. In schemes employing LSSS access structure, m corresponds to the number of rows in the Linear Secret-Sharing Schemes (Subsection 2.4) matrix $M_{l\times n}$. The term I_e represents the minimum number of interior nodes required to satisfy an access structure, and $|I|$ indicates the number of attributes involved in the decryption of ciphertext.

This paper utilizes these time measurements (Table 5) in Table 6 to construct the computational times for key generation, encryption, and decryption with a varying number of attributes. The computation cost for the 100 number of attributes in BSW07 can be calculated as $(2|S|+2)T_E + |S|H'$, that is, $(2\times 100+2).0.82201 + 100\times 1.8171 \approx 348.12$ ms, let for Yamada et al. scheme keygen time is $(6\times 100+2)\times 0.82201 = 494.85 \approx 498.26$ ms. Our scheme's computation cost for keygen 100 attributes is $(2\times 100+2)\times 0.8171 \approx 162.88$ ms. Similar to the keygen, the computation cost for encryption for schemes such as Yang et al., for 100 number of attributes can be calculated as $(5\times 100+3)\times 0.82201 + 0.537239 = 414 \approx 424.21$ ms including an additional overhead Δ associated with the encryption process. The results depicted in Figure 8a,b illustrate that the proposed scheme, along with the schemes BSW07, Yamada et al., Yang et al., and Fabeo, demonstrates superior efficiency in secret key generation and encryption compared to other approaches. The computation cost associated with the decryption (shown in Figure 8c) of our scheme is comparatively higher than FABEO, BSW07, and Yamada et al. and lower than Yang et al.(discussed in detail in Subsection 7.2). Addressing the reduction of decryption costs while maintaining the security of the scheme is an area we plan to explore in future work.

7.1.3 | Communication Overhead

The Subsubsection 7.1.3 provides an overview of the communication overhead analysis, including both the count of communicated messages and the overall communication cost incurred. Specifically, Subsubsection 7.1.3 indicates that the number of public parameters for the schemes proposed by BSW07, Yamada et al., Yang et al., FABEO, and our own are 5, 7, 4, 3, and 4, respectively.

Table 7 outlines the communication overhead of various CP-ABE schemes and presents the public parameter, secret key, and ciphertext used in the schemes by BSW07, Yamada et al., Yang et al., and FABEO. These schemes utilize public parameters, as follows: BSW07 includes $\{\mathbb{G}_0, g, h, f, e(g,g)^{\alpha}\}$, Yamada et al.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70064 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 4 | Features comparison of proposed work with the existing solutions.

Ref	Type	Order groups	Access structure	Revocation level	Key update	Ciphertext update	Ciphertext updater	Non-repudiation	Forward/backward security	User collusion resistant	Blockchain	Formal verification
[6]	Ciphertext-Policy ABE	Prime	Tree	✗	✗	✗	✗	✗	×	×	✗	✗
[17]	Ciphertext-Policy-ABE	Prime	LSSS	Attribute	✓	✓	Server	✗	✓	✓	✗	✗
[12]	Ciphertext-Policy-ABE	Prime	LSSS	✗	✗	✗	✗	✗	✗	✗	✗	✗
[24]	Ciphertext-Policy-ABE	Prime	LSSS	User	✓	✓	Server	✗	✓	✗	✗	✗
[1]	Ciphertext-Policy-ABE	Composite	LSSS	✗	✗	✗	✗	✗	✗	✗	✗	✗
[23]	Ciphertext-Policy-ABE	Prime	LSSS	✗	✗	✗	✗	✗	✗	✓	✗	✗
[19]	Ciphertext-Policy-ABE	Prime	LSSS	Attribute	✓	✓	Server	✗	✓	✓	✗	✗
[11]	Ciphertext-Policy-ABE	Prime	Boolean Formulae and MSP	✗	✗	✗	✗	✗	×	×	✗	✗
[5]	Ciphertext-Policy-ABE	Prime	LSSS	✗	✗	✗	✗	✓	✗	✗	✓	✗
[25]	Ciphertext-Policy-ABE	Prime	LSSS	✗	✗	✗	✗	✓	✗	✗	✓	✗
BloCPABE	Ciphertext-Policy-ABE	Prime	LSSS	Attribute	✓	✓	Owner	✓	✓	✓	✓	✓

Note: ✓ = The scheme has the feature, ✗=absent, ×=Unknown.

TABLE 5 | Average execution time for each operation.

Operation	Time
Multiplication time in $\mathbb{Z}_p$	0.001129 ms
Exponentiation time in $\mathbb{G}_1$	0.82201 ms
Exponentiation time in $\mathbb{G}_T$	0.07321 ms
Multiplication time in $\mathbb{G}_1$	0.00479 ms
Multiplication time in $\mathbb{G}_T$	0.001039 ms
Inverse Operation time in $\mathbb{G}_1$	0.000509 ms
Inverse Operation time in $\mathbb{G}_T$	0.00323 ms
Time for hash operation	1.8171 ms
Time for pairing operation	0.537239 ms

employ $\{g, H, U, V, V', W, e(g,g)^{\alpha}\}$, Yang et al. use $\{g, g^a, g^b, \mathcal{H}\}$, and FABEO incorporates $\{\mathcal{G}, \mathcal{H}, e(g_1, g_2)^{\alpha}\}$.

The number of messages in the secret key for each scheme is 3, 6, 3, and 3, respectively. These are as follows: BSW07 includes D, D_j, D'_j, Yamada et al. incorporates $\{D_1, D_2, \{K_{i,1}, K_{i,2}, K'_{i,1}, K'_{i,2}\}\}$, Yang et al. uses $\{K_{uid,aid}, K'_{uid,aid}, \{K_{x_{aid},uid}\}\}$, and FABEO employs $\{sk_1, \{sk_2\}, sk_3\}$. Similarly, in the ciphertext, the communicated messages are as follows: BSW07 includes $\{\widetilde{C}, C, \{C_y, C'_y\}\}$, Yamada et al. use $\{C_0, C_1, \{C_{i,1}, C_{i,2}, C_{i,3}\}\}$, Yang et al. employ $\{C, C', C'', \{C_i, C'_i, D_i, D'_i\}\}$, and FABEO includes $\{ct_1, \{ct_{2,j}\}, \{ct_{3,i}\}\}$.

As elaborated in Section 5, our scheme utilizes public parameters (PP) $g, g^a, g^b, \mathcal{H}$. Furthermore, the secret key (SK) in our scheme comprises the set $\{K, K', K'', \{K_{x_i}\}\}$, as detailed in Figure 6. The total number of communicated messages in ciphertexts is 5, represented as $CT = \{C, C', C'', \{C_{1,i}, C_{2,i}\}\}$. When compared to the schemes by Yamada et al. and Yang et al., the communication overhead of our scheme remains relatively affordable, as illustrated in Figure 6.

7.1.4 | Storage Overhead

Figure 7 presents a comparison of storage overhead across the BSW07, Yamada et al., Yang et al., FABEO, and our proposed scheme. In this analysis, we evaluate the secret key size and ciphertext size for different CP-ABE schemes as shown in Table 8. Let $|L_{\mathbb{G}_1}|$ represent the bit size of an element in the source group $\mathbb{G}_1$, and $|L_{\mathbb{G}_T}|$ represent the bit size of an element in the target group $\mathbb{G}_T$. These schemes are based on the symmetric pairing SS512, where in the elliptic curve SS512, the sizes of the fields corresponding to $|\mathbb{Z}_p|$, and $|L_{\mathbb{G}_1}|$ are 20 bytes, and 64 bytes respectively. These field sizes are applied to calculate the size of ciphertext and secret key in Figure 7, and the results are depicted in Subsection 7.1.4.

It can be seen from Figure 7a that our scheme results in significantly lower storage overhead compared to the Yamada et al. $(4|S|+2)|L_{\mathbb{G}_1}|$ and BSW07 schemes $(2|S|+1)|L_{\mathbb{G}_1}|$. Additionally, the difference in storage overhead between the FABEO $(|S|+1)|L_{\mathbb{G}_1}|$, Yang et al. $(|S|+2)|L_{\mathbb{G}_1}|$ and our proposed

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70064 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 6 | Computation overhead of CP-ABE schemes.

Schemes	Key generation	Encryption	Decryption
BSW07 [6]	$(2\|S\|+2)T_E + \|S\|H'$	$(2m+2)T_E + mH' + T_P$	$(2\|I\|+1)T_P + I_eT_E$
Yamada et al. [12]	$(6\|S\|+2)T_E$	$(4m+2)T_E + T_P$	$(3\|I\|T_P)$
Yang et al. [17]	$(3\|S\|+2)T_E$	$(5m+3)T_E + T_P$	$(4\|I\|T_P + \|I\|T_E)$
FABEO [11]	$(\|S\|+2)T_E + (\|S\|+1)H'$	$(2m)T_E + (m+1)H' + T_P$	$3T_P$
Proposed	$(2\|S\|+3)T_E$	$(3m+3)T_E + T_P$	$(3\|I\|T_P) + \|I\|T_E$

TABLE 7 | Total communication overhead of CP-ABE schemes.

Schemes	Public parameter	Secret key	Ciphertext	Total overhead
BSW07	5	3	4	12
Yamada et al.	7	6	5	18
Yang et al.	4	3	7	14
FABEO	3	3	3	9
Proposed	4	4	5	13

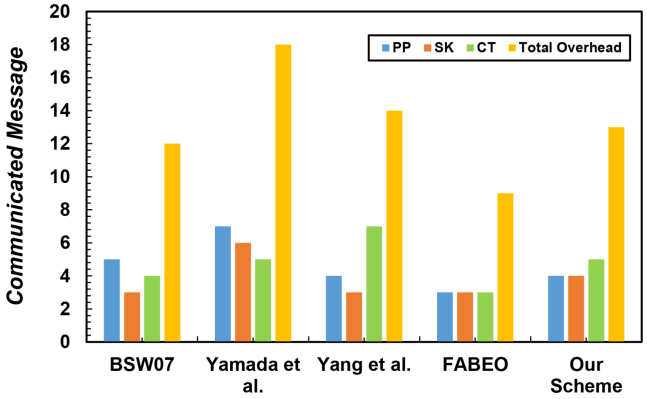


FIGURE 6 | Communication overhead of various schemes.

scheme $(|S|+3)$ $|L_{\mathbb{G}_1}|$ is minimal. Regarding ciphertext storage, FABEO has the smallest cost, expressed as $(m+1)|L_{\mathbb{G}_1}| + 1|L_{\mathbb{G}_T}|$. The BSW07 scheme also has a relatively low overhead at $(2m+1)|L_{\mathbb{G}_1}| + 1|L_{\mathbb{G}_T}|$, which is close to the size of our ciphertext $(2m+3)|L_{\mathbb{G}_1}| + 1|L_{\mathbb{G}_T}|$. In comparison to the schemes by Yang et al. and Yamada et al., our scheme demonstrates lower storage overhead (Figure 7b).

7.2 | Experimental Analysis

This section outlines a simulation-based experimental study aimed at comparing the proposed scheme with state-of-the-art algorithms in the field. The experiment is conducted across three common aspects: Key generation, data encryption, and data decryption. All the experiments are carried out on a personal computer equipped with an AMD Ryzen 5 5500U processor featuring Radeon Graphics, RAM 16 GB, and an SSD 512 GB [33], running Ubuntu 22.04 LTS as the operating system. Our algorithms are implemented in Python 3.10.0 utilizing the "Charm-Crypto" [32] framework on a 512-bit supersingular elliptic curve (SS512) with 160-bit order.

Figure 8 demonstrates the time efficiency of the CP-ABE schemes concerning key generation, encryption, and data decryption. Note that some of the schemes do not address attribute revocation; therefore, only the common aspects are compared in this analysis. In the key generation stage, we evaluated the time consumption of the key generation algorithm by varying the number of attributes, as illustrated in Figure 8a. As the number of attributes increases, we observe a corresponding rise in key generation time for all five schemes. While Yang et al.'s scheme is designed for multi-authority ABE, we evaluated it in a single-authority for computational time comparison. Our proposed scheme demonstrates a notable advantage in key generation time compared to BSW07, Yamada et al., Yang et al., and FABEO. When evaluating with 10 attributes, the key generation times for BSW07, Yamada et al., Yang et al., FABEO, and our scheme are 36.22, 51.95, 28.21, 30.78, and 16.24 ms, respectively. Similarity evaluating with 100 attributes, the time is 348.12, 498.26, 248.78, 268.84, and 162.88 ms, respectively.

To guarantee consistency in the encryption and decryption process, each ciphertext is produced based on "AND" access policies, such as $ATTR_1$ AND $ATTR_2$ AND $ATTR_3$ AND ... AND $ATTR_N$, where N ranges from 2 to 100. Also, to ensure the reliability of execution time data, each experiment is conducted 20 times. The average of these trials is used to calculate the final execution time. As depicted in Figure 8b, the encryption algorithm's execution times for access policies with 10, 50, and 100 attributes are measured across five schemes: BSW07, Yamada et al., Yang et al., FABEO, and our proposed scheme. The results indicate that encryption under these policies consumed 36.69, 35.25, 44.98, 40.03, and 36.46 ms, respectively, for 10 attributes; 177.36, 170.84, 216.48, 181.84, and 169.80 ms for 50 attributes; and 352.86, 340.32, 424.21, 363.48, and 337.70 ms for 100 attributes. Our proposed encryption scheme demonstrates a notable improvement in computational efficiency compared to existing approaches. While there is a marginal performance gap between our scheme and Yamada et al.'s, our solution exhibits better encryption performance.

As depicted in Figure 8c, the execution time of the decryption algorithm for access policies with 10, 50, and 100 attributes are measured across five schemes: BSW07, Yamada et al., Yang et al., FABEO, and our proposed scheme. The results indicate that decryption under these policies consumed 11.64, 18.35, 33.93, 2.37, and 27.79 ms, respectively, for 10 attributes; 54.24, 86.23, 160.34, 3.04, and 130.20 ms for 50 attributes; and 107.16, 171.23, 315.20, 3.50, and 130.20 ms for 100 attributes. The decryption time results presented in Figure 8c indicate that our scheme takes more time compared to BSW07, Yamada et al., and FABEO.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70064 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

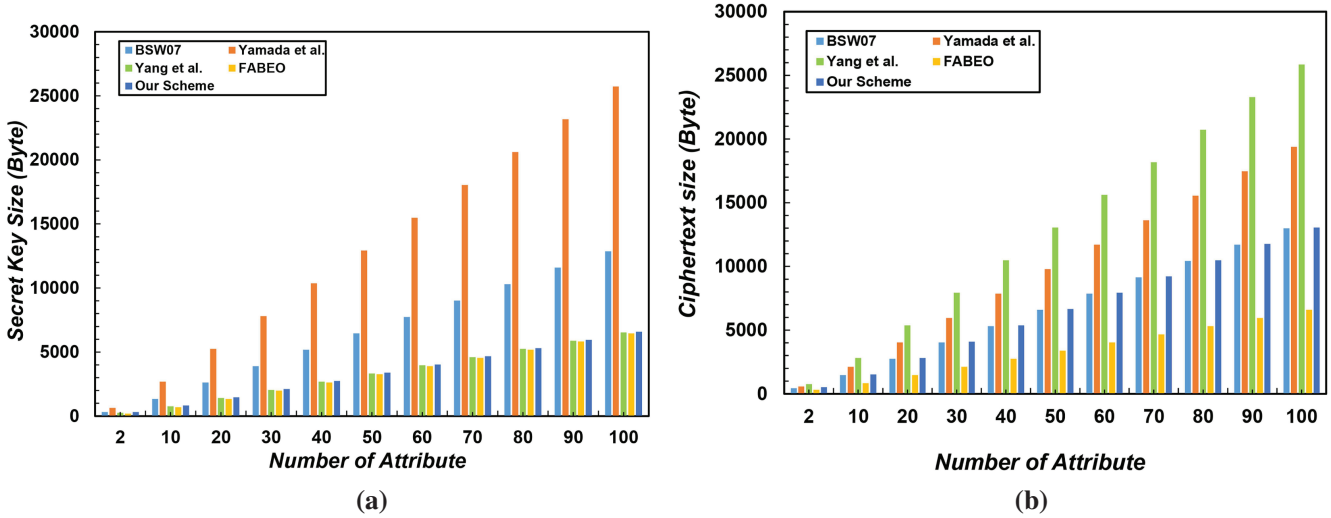


FIGURE 7 | The storage overhead computed for various ABE schemes: (a) key size and (b) ciphertext size.

TABLE 8 | Storage overhead comparison of proposed work with the existing schemes.

Schemes	Secret key size	Ciphertext size
BSW07	$(2\|S\|+1)\|L_{\mathbb{G}_1}\|$	$(2m+1)\|L_{\mathbb{G}_1}\|+1\|L_{\mathbb{G}_T}\|$
Yamada et al.	$(4\|S\|+2)\|L_{\mathbb{G}_1}\|$	$(3m+1)L_{\mathbb{G}_1}+1\|L_{\mathbb{G}_T}\|$
Yang et al.	$(\|S\|+2)\|L_{\mathbb{G}_1}\|$	$(4m+2)\|L_{\mathbb{G}_1}\|+1\|L_{\mathbb{G}_T}\|$
FABEO	$(\|S\|+1)\|L_{\mathbb{G}_1}\|$	$(m+1)\|L_{\mathbb{G}_1}\|+1\|L_{\mathbb{G}_T}\|$
Proposed	$(\|S\|+3)\|L_{\mathbb{G}_1}\|$	$(2m+3)\|L_{\mathbb{G}_1}\|+1\|L_{\mathbb{G}_T}\|$

This additional computational overhead arises from pairing operations and secret key processing tasks associated with each attribute that the user possesses, which increases with the number of attributes stated in the access policy. Through rigorous theoretical and empirical analysis, our proposed scheme demonstrates notable advantages in functionality and computational efficiency compared to existing approaches. The detailed simulation of our proposed CP-ABE scheme with blockchain is analyzed in Subsection 7.3.

7.3 | Blockchain Simulation

To evaluate the effectiveness of the proposed model, a prototype illustrated in Figures 3 and 4 is tested on a system specification given in Subsection 7.2. We developed the smart contract using Solidity and tested it on RemixIDE. For blockchain simulation, we utilized Ethereum on Ganache-Truffle Suite [34], leveraging Web3.js as the communication interface. In this simulation, we assessed our CP-ABE scheme when integrated with IPFS 0.20.0 (InterPlanetary File System) and the Ethereum blockchain. The primary objective is to evaluate the efficiency of the ABE scheme by leveraging IPFS for decentralized storage and Ethereum for recording metadata.

Initially, a smart contract is deployed on a local Ethereum blockchain (Ganache), designed to store IPFS hashes associated with encrypted data. Encrypted content is serialized and uploaded to IPFS, ensuring a decentralized and immutable storage solution. Figure 9 details the overhead associated with the entire process, including the encryption time, the outsourcing of encrypted data to IPFS, and the recording of the hash on the blockchain. The time taken increases with the number of attributes. With 2 attributes, the observed overhead is 0.256 s; with 50 attributes, it rises to 3.882s; and with 100 attributes, it grows to 8.447 s. This increase in processing time highlights the challenges that arise as the array of attributes increases; the complexity and time required for encryption, data outsourcing to IPFS, and blockchain interactions become significant.

8 | Conclusion and Future Scope

The security, privacy, and access control challenges in mHealth systems are addressed in this paper by the proposal of the BloC-PABE scheme. Additionally, the system facilitates the efficient revocation of attributes with minimal computational latency, guaranteeing that user access rights are promptly updated in response to attribute modifications. The results demonstrate that the BloCPABE is computationally more efficient than existing ABE methods. In particular, we observed a reduction in key generation costs by 35% to 67% and encryption costs by 26% to 39%, which underscores the efficiency enhancements of the approach. The BloCPABE scheme is designed to resolve critical concerns in healthcare data sharing by ensuring user privacy, forward/backward security, and resistance to collusion attacks. The solution further enhances security by guaranteeing the immutability and tamper-proof integrity of outsourced data through the integration of blockchain technology. These findings reveal that the BloCPABE scheme is highly secure and privacy-preserving, rendering it appropriate for real-time, fine-grained access control in mHealth systems. This work has substantial implications, as it offers a solution for enhancing the security and privacy of patient health records and enabling the efficient exchange of sensitive health data. Through theoretical complexity analysis and experimental experiments, we verified the efficiency and feasibility of our designs. The algorithm's scalability and versatility will be enhanced by refining it to

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70064 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

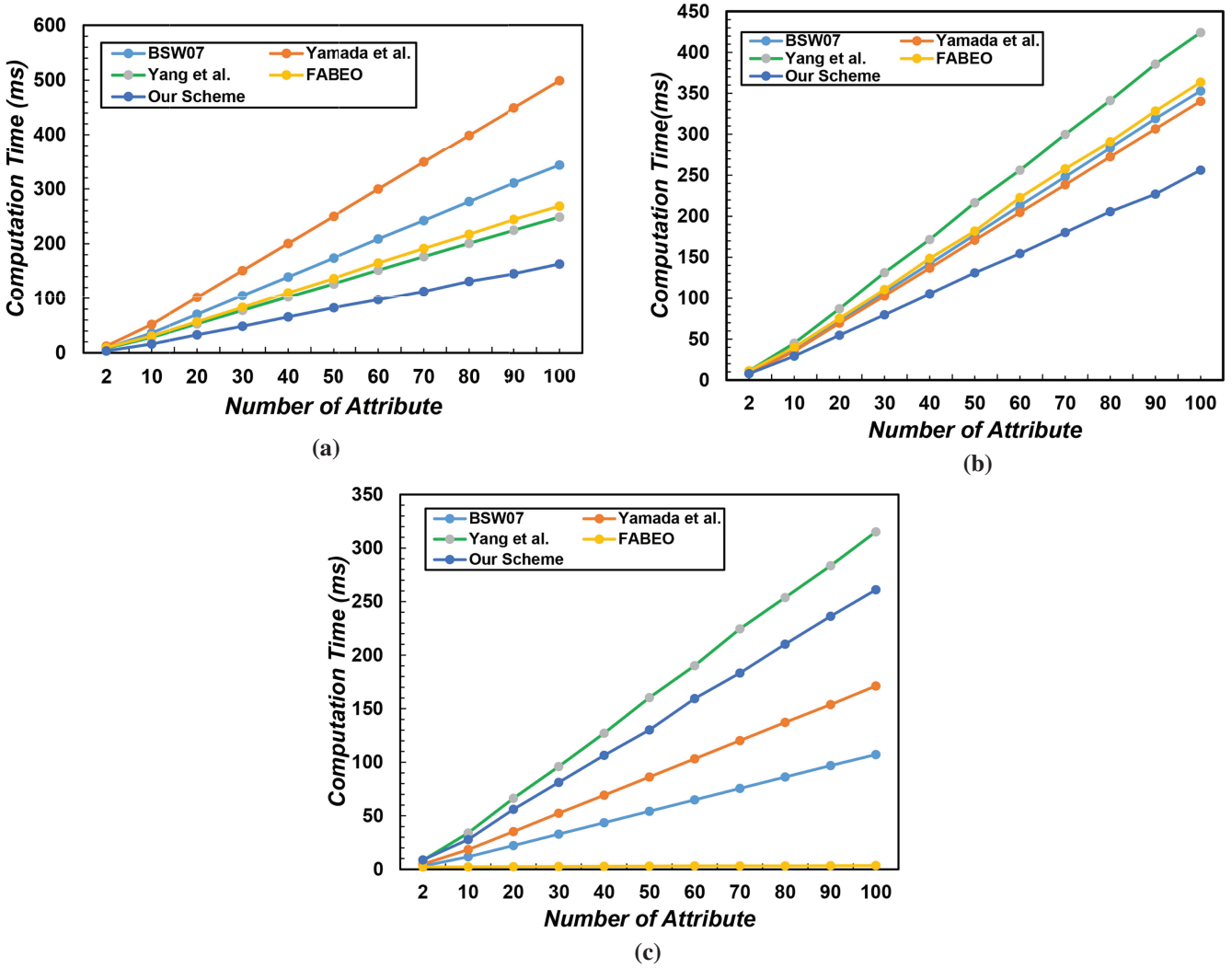


FIGURE 8 | The computation time of CP-ABE schemes for keygen, encrypt, and decrypt algorithm. (a) Key generation, (b) encryption, and (c) decryption.

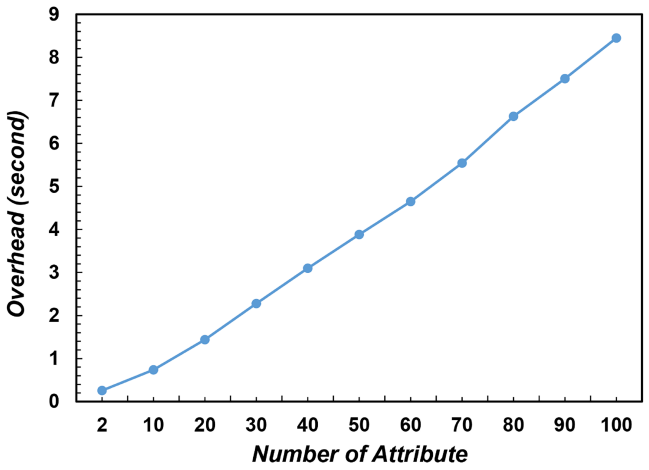


FIGURE 9 | Overhead analysis with blockchain interaction.

accommodate a multi-authority framework in future developments. In parallel, its computational efficiency will be improved while simultaneously ensuring adherence to rigorous security.

Disclosure

The authors have nothing to report.

Ethics Statement

The authors have nothing to report.

Consent

All authors have their full consent for its publication.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

References

1. Q. Li, Y. Zhang, T. Zhang, H. Huang, Y. He, and J. Xiong, "HTAC: Fine-Grained Policy-Hiding and Traceable Access Control in mHealth," *IEEE Access* 8 (2020): 123430–123439.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70064 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2. X. Liang, Y. Liu, and J. Ning, "An Access Control Scheme With Privacy-Preserving Authentication and Flexible Revocation for Smart Healthcare," *IEEE Journal of Biomedical and Health Informatics* 28 (2024): 3269–3278.

3. R. Arul, R. Alroobaea, U. Tariq, A. H. Almulihi, F. S. Alharithi, and U. Shoaib, "IoT-Enabled Healthcare Systems Using Block Chain-Dependent Adaptable Services," *Personal and Ubiquitous Computing* 28, no. 1 (2024): 43–57.

4. K. Zala, H. K. Thakkar, N. Dholakia, M. Shukla, and D. Thumar, "Designing an Attribute-Based Encryption Scheme With an Enhanced Anonymity Model for Privacy Protection in E-Health," *SN Computer Science* 5, no. 2 (2024): 203.

5. S. Xu, J. Zhong, L. Wang, D. He, S. Zhang, and W. Shao, "A Privacy-Preserving and Efficient Data Sharing Scheme With Trust Authentication Based on Blockchain for mHealth," *Connection Science* 35, no. 1 (2023): 2186316.

6. J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-Policy Attribute-Based Encryption," in *2007 IEEE Symposium on Security and Privacy (SP'07)* (IEEE, 2007), 321–334.

7. Y. Tao, Y. Zhu, C. Ge, et al., "ORR-CP-ABE: A Secure and Efficient Outsourced Attribute-Based Encryption Scheme With Decryption Results Reuse," *Future Generation Computer Systems* 161 (2024): 559–571.

8. B. Zhang, W. Yang, F. Zhang, and J. Ning, "Efficient Attribute-Based Searchable Encryption With Policy Hiding Over Personal Health Records," *IEEE Transactions on Dependable and Secure Computing* 22, no. 2 (2025): 1299–1312.

9. B. Waters, "Ciphertext-Policy Attribute-Based Encryption: An Expressive, Efficient, and Provably Secure Realization," in *International Workshop on Public Key Cryptography* (Springer, 2011), 53–70.

10. A. Sahai and B. Waters, "Fuzzy Identity-Based Encryption," in *Advances in Cryptology–EUROCRYPT 2005: 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Aarhus, Denmark, May 22–26, 2005. Proceedings 24* (Springer, 2005), 457–473.

11. D. Riepel and H. Wee, "FABEO: Fast Attribute-Based Encryption With Optimal Security," in *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security* (ACM, 2022), 2491–2504.

12. S. Yamada, N. Attrapadung, G. Hanaoka, and N. Kunihiro, "A Framework and Compact Constructions for Non-Monotonic Attribute-Based Encryption," in *Public-Key Cryptography–PKC 2014: 17th International Conference on Practice and Theory in Public-Key Cryptography, Buenos Aires, Argentina, March 26–28, 2014. Proceedings 17* (Springer, 2014), 275–292.

13. R. Ostrovsky, A. Sahai, and B. Waters, "Attribute-Based Encryption With Non-Monotonic Access Structures," in *Proceedings of the 14th ACM Conference on Computer and Communications Security* (ACM, 2007), 195–203.

14. M. Chase, "Multi-Authority Attribute Based Encryption," in *Theory of Cryptography: 4th Theory of Cryptography Conference, TCC 2007, Amsterdam, the Netherlands, February 21–24, 2007. Proceedings 4* (Springer, 2007), 515–534.

15. M. Chase and S. S. Chow, "Improving Privacy and Security in Multi-Authority Attribute-Based Encryption," in *Proceedings of the 16th ACM Conference on Computer and Communications Security* (ACM, 2009), 121–130.

16. A. Lewko and B. Waters, "Decentralizing Attribute-Based Encryption," in *Annual International Conference on the Theory and Applications of Cryptographic Techniques* (Springer, 2011), 568–588.

17. K. Yang and X. Jia, "Expressive, Efficient, and Revocable Data Access Control for Multi-Authority Cloud Storage," *IEEE Transactions on Parallel and Distributed Systems* 25, no. 7 (2013): 1735–1744.

18. Y. Miao, F. Li, X. Li, et al., "Verifiable Outsourced Attribute-Based Encryption Scheme for Cloud-Assisted Mobile e-Health System," *IEEE Transactions on Dependable and Secure Computing* 21, no. 4 (2024): 1845–1862.

19. S. Tu, M. Waqas, F. Huang, G. Abbas, and Z. H. Abbas, "A Revocable and Outsourced Multi-Authority Attribute-Based Encryption Scheme in Fog Computing," *Computer Networks* 195 (2021): 108196.

20. Y. Guo, Z. Lu, H. Ge, and J. Li, "Revocable Blockchain-Aided Attribute-Based Encryption With Escrow-Free in Cloud Storage," *IEEE Transactions on Computers* 72, no. 7 (2023): 1901–1912.

21. Y. Jiang, X. Xu, and F. Xiao, "Attribute-Based Encryption With Blockchain Protection Scheme for Electronic Health Records," *IEEE Transactions on Network and Service Management* 19, no. 4 (2022): 3884–3895.

22. H. Guo, W. Li, M. Nejad, and C. C. Shen, "A Hybrid Blockchain-Edge Architecture for Electronic Health Record Management With Attribute-Based Cryptographic Mechanisms," *IEEE Transactions on Network and Service Management* 20, no. 2 (2022): 1759–1774.

23. X. Qin, Y. Huang, Z. Yang, and X. Li, "A Blockchain-Based Access Control Scheme With Multiple Attribute Authorities for Secure Cloud Data Sharing," *Journal of Systems Architecture* 112 (2021): 101854.

24. J. Wei, X. Chen, X. Huang, X. Hu, and W. Susilo, "RS-HABE: Revocable-Storage and Hierarchical Attribute-Based Access Scheme for Secure Sharing of e-Health Records in Public Cloud," *IEEE Transactions on Dependable and Secure Computing* 18, no. 5 (2019): 2301–2315.

25. Z. Hou, J. Ning, X. Huang, S. Xu, and L. Y. Zhang, "Blockchain-Based Efficient Verifiable Outsourced Attribute-Based Encryption in Cloud," *Computer Standards & Interfaces* 90 (2024): 103854.

26. T. Feng, F. Kong, C. Liu, and Y. Lu, "Multi-Authorization Attribute-Based Verifiable Encryption Scheme Based on Blockchain," *Mobile Networks and Applications* 28, no. 5 (2023): 1617–1624.

27. C. J. Cremers, "The Scyther Tool: Verification, Falsification, and Analysis of Security Protocols: Tool Paper," in *International Conference on Computer Aided Verification* (Springer, 2008), 414–418.

28. A. Armando, D. Basin, Y. Boichut, et al., "The AVISPA Tool for the Automated Validation of Internet Security Protocols and Applications," in *Computer Aided Verification: 17th International Conference, CAV 2005, Edinburgh, Scotland, UK, July 6–10, 2005. Proceedings 17* (Springer, 2005), 281–285.

29. B. Blanchet, "Automatic Verification of Security Protocols in the Symbolic Model: The Verifier Proverif," in *International School on Foundations of Security Analysis and Design* (Springer, 2012), 54–87.

30. S. Meier, B. Schmidt, C. Cremers, and D. Basin, "The TAMARIN Prover for the Symbolic Analysis of Security Protocols," in *Computer Aided Verification: 25th International Conference, CAV 2013, Saint Petersburg, Russia, July 13–19, 2013* (Springer, 2013), 696–701.

31. A. Thakur, V. Ranga, and R. Agarwal, "Attribute-Based Encryption Scheme for Secure and Efficient Access in Blockchain," in *IEEE International Conference for Women in Innovation Technology & Entrepreneurship (ICWITE)* (IEEE, 2024), 653–658.

32. J. A. Akinyele, C. Garman, I. Miers, et al., "Charm: A Framework for Rapidly Prototyping Cryptosystems," *Journal of Cryptographic Engineering* 3 (2013): 111–128.

33. A. Thakur, V. Ranga, and R. Agarwal, "Workload Dynamics Implications in Permissioned Blockchain Scalability and Performance," *Cluster Computing* 27, no. 8 (2024): 11569–11593.

34. T. Suite, "Ganache-Truffle Suite" (2024), https://archive.trufflesuite.com/ganache/.

15320634, 2025, 9-11, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/cpe.70064 by Delhi Technological University, Wiley Online Library on [30/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Materials Research Proceedings 49 (2025) 549-558 https://doi.org/10.21741/9781644903438-55

Rice straw as the potential raw material for biogas production at variable temperature and feed ratio

Shubham KUMAR[1], Sawan KAPOOR[1] and Neeraj BUDHRAJA[1,a]*

[1]Department of Mechanical Engineering, Delhi Technological University, Delhi 110042, India

[a]neeraj_budhraja@yahoo.com

Keywords: Biogas, Rice Straw, Anaerobic Digestion, Feed Ratio

Abstract. The use of agricultural wastes for the production of renewable energy is becoming more popular as communities look for long-term solutions to problems related to waste management and energy. An abundant residue of rice farming, rice straw offers a promising fuel to produce biogas by anaerobic digestion. The aim of this research paper is to analyze the method of converting rice straw into biogas. A culture of bacteria breaks down complex organic compounds present in rice during the anaerobic digestion process to produce methane-rich biogas. To improve the yield and quality of biogas, optimization strategies like adjusting process parameters are investigated. To reduce global warming and advance sustainable development, the results highlight rice straw's potential as a renewable energy source and call for its usage.

Introduction

Recently, the international community has become more determined to find sustainable sources of power; this is because they are worried about the environment and need to prevent global warming [1]. Among all kinds of renewable energy, biogas is considered as a good choice due to its flexibility, efficiency and being eco-friendly. In fact, making use of rice straws for biogas production can completely change how people generate electricity and deal with waste. The benefits, drawbacks, prospects, and risks associated were examined with using various portions of wheat and rice straw as a feedstock for the bio-based economy [2,3].

Rice is consumed by almost half of the world's population so it is no surprise that it is grown in many countries [4]. Two studies of the trials wherein straw was consumed, using field probing straw consolidation and ignition chamber analysis to examine the impact of rice straw consumption on methane and nitrous oxide discharges were conducted. By a factor of 2.1, the Emission Factor of methane obtained from dry rice straw (>10% moisture content) exceeded the default IPCC level. Methane was the primary driver of all outflows and GWP, despite the fact that the nitrous oxide Emission Factor was nearly identical to the IPCC estimate. This commitment will even be problematic in light of the regular increase in methane Emission Factors as a dampness substance component, as described in several analyses [5]. However, large quantities of agricultural waste known as rice straw are produced during rice farming activities such as burning them off or letting them rot where they lie which causes pollution through release into the atmosphere as smog or greenhouse gasses like methane. Various estimates of the air pollutants released by burning crop wastes in India's fields, broken down by state [1]. This not only harms our surroundings but also means we ignore important fuel resources available to us. The conversion of rice straw to biogas offers an attractive avenue towards addressing the two issues.

The research provided a thorough understanding of the detrimental effects of crop residue mismanagement in India, specifically with regard to stubble burning. It also discusses other, more promising methods of managing crop residues, such as the use of bioenergy if widely used, will not only lessen the effect of crop residue management on the environment but also create additional value for the agricultural industry worldwide [6]. The different options for transforming a "waste"

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 license. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. Published under license by Materials Research Forum LLC.

object into a valuable one were also examined [7]. It could be accomplished by producing electricity from biomass, serving as a hotspot for domesticated animal feed, assisting in the management of disintegration, or even producing necessary end products like medicines, health products, and composts. This essay aims to investigate optional environmentally friendly disposal methods that India could implement.

The rice straw recycling, with particular attention to the energy circumstances in China and around the world, managing biogas digestate [8]. In contrast, *Mothe and Polisetty* examined how using rice straw in combination with other co-substrates and the right pretreatment could be a sustainable way to reduce risks to the health and environment while increasing the production of biomethane [9]. In contrast, rice straw was pretreated chemically and biologically (with NaOH, HCl, $CO(NH_2)_2$, and cellulase) at room temperature (approximately 20℃) to enhance its biodegradability and boost the production of anaerobic biogas.

Biogas, made up of CH_4 and CO_2, is produced by anaerobic digestion through fermentation of organic materials by microorganisms. Rice straw is made up of cellulose and hemicellulose, which makes it an excellent fuel for biogas production due to its renewability and availability. The effectiveness of three pretreatments to increase the production of biogas from rice straw: microwave, ammoniation, and microwave-assisted ammoniation (MAA) were examined [10]. In order to reduce environmental pollution, an overview of the most promising options for managing paddy straw and converting biogas to bio-CNG is provided in this article [11]. The performance of fresh and dry rice straw was evaluated in this study at 35°C during batch anaerobic digestion. The findings revealed that the biogas produced from fresh rice straw was 400.5 mL/g of volatile solids, which was more than 40.9% produced from the substrate made of dry rice straw [12,13]. In contrast, a 1 m^3 floating drum anaerobic digester was used to determine the ideal parameters for the production of biogas from the anaerobic digestion of rice straw using Response Surface Methodology (RSM) [14].

This research paper seeks to explain and analyze the conversion process of rice straw and other agricultural wastes into useful biofuel like biogas which can be used as a source of energy. Complete process of anaerobic digestion used for the conversion was studied. This paper also analyzed various process parameters that affect the production of biogas through anaerobic digestion. The effect of various changes of process parameters were examined and the consequence of each parameter on production of biogas was compared to find the best possible solution for getting maximum production of biogas.

Biogas from Rice Straw Good for Sustainable Energy

Making biogas from rice straw using anaerobic digestion is a good way to make sustainable energy. It helps to reduce carbon in the energy industry and fight climate change by releasing less greenhouse gases. Moreover, it answers the growing need for clean and renewable power supply[15]. In addition to this, the process gives opportunities for employment creation in rural areas. Another thing that can be done towards sustainability is refraining from burning rice straw on fields. It has been estimated that about 300 million tons of waste rice straw are burned in fields across Asia every year thereby adding to greenhouse gasses emitted into the atmosphere [16]. Furthermore, respiratory diseases as well as premature deaths are caused by burning rice straws. Nonetheless, biogas production from rice straw has been proved successful but much still remains untapped on this technology. With Biogas energy made out of rice husk it can be used for green power generation or cook with heat produced from it can also serve heating and electricity purposes while being supplied through power systems or other means like fuel cells etcetera. Converting it into biogas contributes to sustainable energy production through utilization of residues left after harvesting rice crops too [17]. Mechanical pretreatment may help overcome some challenges encountered during biogas production from this type of feedstock i.e., straw. The THDI Equation may be utilized in assessing how far these processes would go towards ensuring

sustainable development goals get achieved within our societies more especially those living near such facilities should take note so that they do not complain later on when things fail them again [18].

MATERIALS AND METHOD

Rice straw

Rice straw is a residual material from rice harvesting. As the primary staple food, rice provides half of the world's population with both protein and energy. As a staple crop, rice is widely consumed across India. In the financial year 2023 India produced 135 million metric tons of rice. This indicates steady rise in rice production since 2017. Whereas more than any other nation, China produced approximately 146 million metric tonnes of rice in the 2022–2023 crop year as observed from Fig. 1. With more than 135 million metric tonnes of rice produced in that crop year, India ranks second.

In the past, rice straw was harvested manually using traditional methods. After that, it served as a raw ingredient for pulp and paper, fertilizer, and animal feed industries. In spite of this, technology development leads to rise in rice straw collection expenses and time-consuming parameters. As a consequence, a very low percentage of the rice straw, up to 20%, is further used, while farmers burn the rest amounting to 50% for the preparation of the fields for the subsequent crop. Burning techniques release a variety of harmful pollutants, such as carbon monoxide (CO), sulfur dioxide (SO_2), and nitric oxide (NO_X), as well as greenhouse gasses (GHGs) including carbon dioxide (CO_2), methane (CH_4) resulting in respiratory diseases.

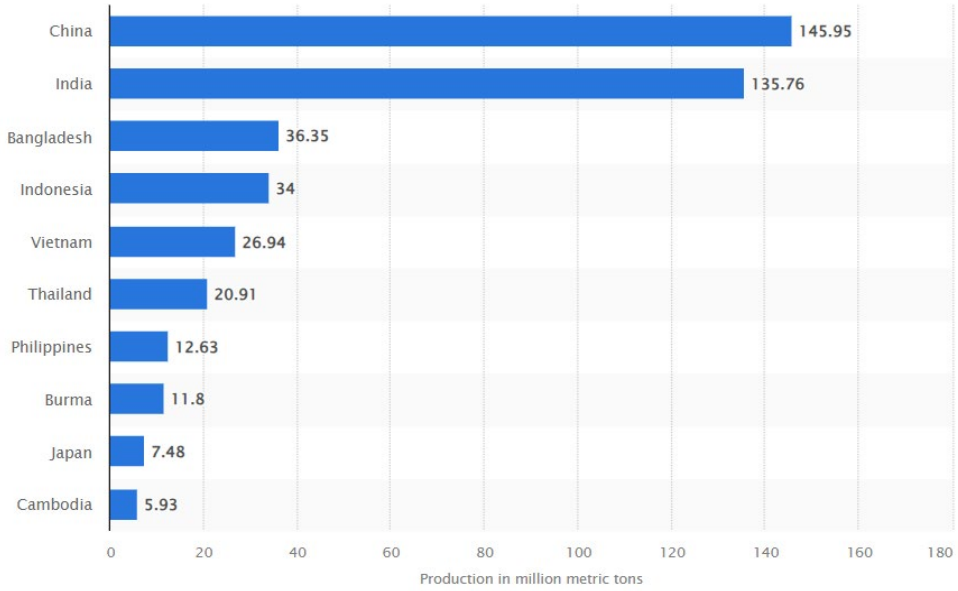


Figure 1. Rice production by top 10 nations (Source: US Department of Agriculture 2024)

Properties and Composition of Rice Straw

The properties of rice straw can be divided into three main areas, to determine its utility: (i) physical characteristics; (ii) thermal qualities; and (iii) chemical composition.

Physical characteristics include bulk density, heat capacity, and thermal conductivity. The most important consideration for managing and storing rice straw is density. Thermal properties and heating value are crucial in the process of turning biomass into energy. Lignin, cellulose, hemicellulose/carbohydrates, and nutritional contents are among the components of chemical composition that have an impact on application-specific aspects including soil fertility and livestock feed.

Materials Research Proceedings 49 (2025) 549-558 https://doi.org/10.21741/9781644903438-55

Physical properties

The dry matter density of loose rice straw can range from 14-18 kg/m³. Chopped straw with lengths ranging from 1-10 mm can have densities in range from 50 to 120 kg/m³, depending on the apparatus used. Depending on the machinery used, the straw's size, mass density, and compression ratio will vary. A circular rice straw bale of 65 cm in length and 55 cm in diameter has a mass density between 65 and 85 kg/m³. The density of rice straw briquettes, with a diameter of 85 mm and a thickness ranging from 8-16 mm, is 360–440 kg/m³. Pellets of rice straw, measuring 9 mm in diameter and varying in height from 40 -60mm have density ranging from 600-700 kg/m³. The moisture content of the straw affects its heating value, which is important if the waste is intended to be used as bioenergy.

Thermal properties

The calorific value of rice straw is 14.2 MJ/kg is comparable to that of rice husks. As the first components to be driven off during combustion, the volatile carbon, combined with water, nitrogen, sulfur and hydrogen are referred to as volatiles in the proximate analysis. The range of 60.55 to 69.70% for the volatile matter (VOM) concentration of rice straw is proportional to the biomass of sugarcane bagasse, corn straw, and wheat straw. The final examination reveals the elemental oxygen, hydrogen, carbon, nitrogen and sulfur content of rice straw. In comparison to fossil fuels, rice straw biomass contains more oxygen and hydrogen but less carbon content. The ratio of H:C in rice straw is 0.94:1.06 and O:C is 1.1:1.36.

Table 1. Composition of Rice straw

S. No	Component Found in Rice Straw	Quantity (wt. %)
1.	Cellulose	43
2.	Hemicellulose	25
3.	Lignin	12
4.	Ash and Silica	34 - 44
5.	Glucose	19 - 22
6.	Arabinose	2 - 4
7.	Mannose	1.8 - 2
8.	Galactose	0.4
9.	Non-Structural Components	16 - 20

Chemical properties

The three main components of lignocellulose biomass are hemicellulose, cellulose and lignin. Glycoside linkages bind the subunits of monomeric D-glucose together to form cellulose, an essential component. Depending on how the structure is organized, cellulose is further divided into two forms: crystalline, which is more organized, and amorphous, which is less organized and is typically more relevant for fermentation. The cellulose and lignin molecules are held together by hemicellulose, giving the structure its tensile strength as depicted in Table 1. The mosaic structure made up of alcohols—is protected by lignin.

Anaerobic Digester

Anaerobic digesters are basically enclosed areas where anaerobic process breaks down organic waste from compost. The anaerobic bacteria converts the organic waste into fuel-grade biogas.

Components of Anaerobic Digesters-

i. Tank or Vessel: An anaerobic digester's primary component is a tank or vessel that contains the organic substrate and creates a controlled environment for the process of anaerobic digestion.
ii. Mixing System: To guarantee substrate and microbial population uniformity throughout the digestion, many anaerobic digesters use a mixing system.
iii. Gas Collection System: During anaerobic digestion, biogas is created and builds up in the digester vessel's headspace. The biogas is collected and captured for further processing or usage using a gas collecting system, which typically consists of pipelines, valves, and a gas storage tank.
iv. Effluent Removal System: Liquid effluent is regularly or continuously released from the digester during the anaerobic digestion process. The digestate is removed from the digester vessel using devices like pumps or gravity drainage systems.

ANAEROBIC DIGESTION

Anaerobic digestion is the process by which anaerobic bacteria decomposes organic materials in the absence of oxygen, such as animal dung, plant waste, and food wastes [19,20]. A sealed vessel known as a digester or reactor—which is built in various sizes and shapes depending on the site and feedstock conditions—is where the process of anaerobic digestion for the production of biogas takes place [21,22]. Complex microbial populations found in these digesters break down the waste material to produce digestate, or the solid and liquid byproducts of the anaerobic digestion process, and biogas that is discharged from the digester [23]. The reactions involved in the anaerobic digestion process are depicted in Table 2. The various stages involved in the anaerobic digestion are depicted in Fig. 2.

Hydrolysis

The goal of the hydrolysis process is to convert organic macromolecules into smaller pieces so that acidogenic bacteria can use them. While hydrolysis in anaerobic digestion can occur as an electrochemical process, it primarily occurs as a biological one. Hydrolytic bacteria release extracellular enzymes during the hydrolysis process, which can change lipids, proteins and carbohydrates into long-chain fatty acids and amino acids. Acidogenic bacteria cell membranes allow the hydrolysis products to permeate through after enzymatic cleavage. To improve the breakdown of these sugars, enzymes are frequently added.

Acidogenesis

Acidogenic bacteria produce volatile fatty acids (VFAs) and other compounds by absorbing the products formed during hydrolysis through their cell membranes. Larger organic acids like acetates and other organic acids make up VFAs. VFA concentrations can vary greatly for digesters operating at varied pH and temperature levels. The precise concentrations of intermediates produced in the acidogenesis stage depends on the digester's circumstances. It is commonly accepted that acidogenesis proceeds more quickly than any other stage of anaerobic digestion, with acidogenic bacteria regenerating in less than 36 hours. It is commonly known that digester failure is caused by VFA acidification.

Acetogenesis

A fraction of the initial substrate becomes eligible for acetoclastic methanogenesis when acetate is produced during acidogenesis. The process that turns highly volatile fatty acids (VFAs) and other intermediates into acetate is called acetogenesis. Lipids go through two distinct processes of acetogenesis at the same time: acidogenesis produces acetate from glycerol and β-oxidation produces acetate from LCFAs.

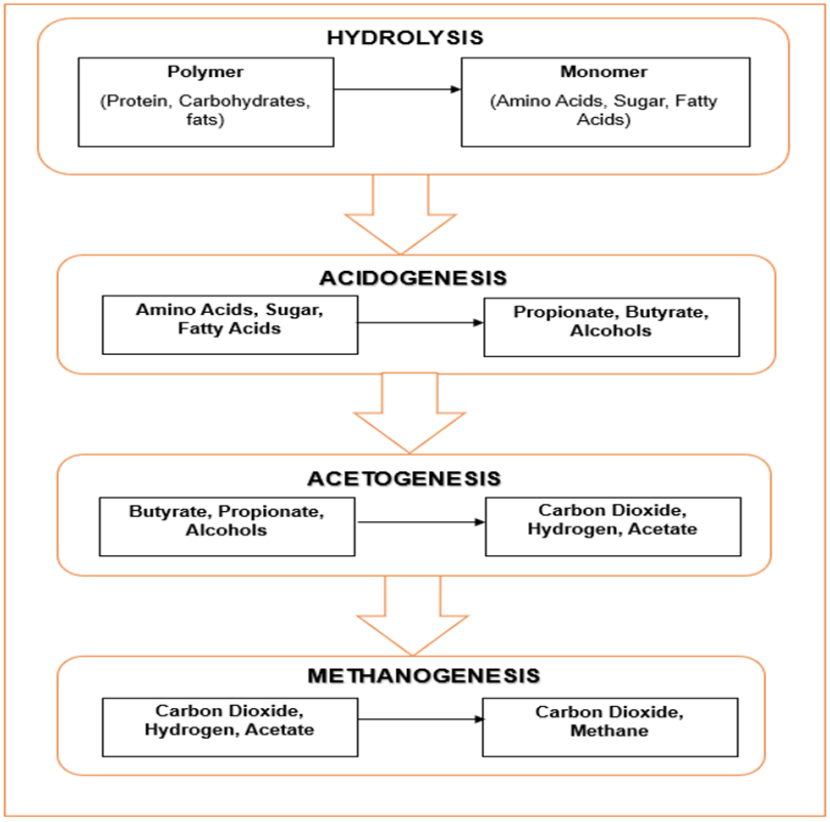


Figure 2. Stages involved in anaerobic digestion process

Methanogenesis

The last phase of anaerobic digestion is methanogenesis, in which methanogenic bacteria consumes available intermediates to generate methane. A class of oxygen-sensitive obligatory anaerobic archaea are known as methanogenic microorganisms. Methanogenic microbes need a higher pH and a lower redox potential in order to carry out methanogenesis. Methanogenesis ends when the production of biogas ceases, which can happen after 40 days or so. The volatile solids concentration and capacity of a sludge can be used to assess the degree of digestion.

Table 2. Reactions involved in anaerobic digestion process

Reactants	Products	Organisms Involved
$4H_2 + HCO_3^- + H^+$	$CH_4 + 3H_2O$	Most Methanogens
$4HCO_2 + H^+ + H_2O$	$CH_4 + 3HCO_3^-$	Many Hydrogenotrophic Methanogens
$2CH_3COOH + HCO_3^-$	$2CH_3COO^- + H^+ + CH_4 + H_2O$	Methanosarcina and Methanothrix
$4CH_3OH$	$3CH_4 + HCO_3^- + H_2O + H^+$	Methanosarcina and other Methylotrophic Methanogens

RESULTS AND DISCUSSION

Feed Ratio Adjustment

Rice straw is cut into smaller pieces, which reduces the particle size of rice straw breaking down the cell walls, increasing the organic substrate's availability for microbial breakdown. Size

reduction of rice straw increases surface area and disrupts its polymer structure, which boosts the yield and rate of hydrolysis during digestion.

After cutting the rice straw into smaller pieces making it more suitable for the anaerobic digestion process in the digester, rice straw is mixed with water making a mixture of water and rice straw (v/v) to feed into the digester. Feed is taken in different ratios of water and rice straw to check production of biogas (methane) from different feed ratios in digester from anaerobic digestion, to check at what ratio maximum biogas production occurs.

Three different feed ratios are taken for this experiment:

- 1:4 (1 unit rice straw and 4 units water)
- 1:5 (1 unit rice straw and 5 units water)
- 1:6 (1 unit rice straw and 6 units water)

Test results were conducted on a daily basis to check the biogas production on daily basis at different percentages of feed added to the anaerobic digester (see Table 3).

Table 3. Methane Production at different hydraulic retention time

Days	Digester 1 (1:4)	Digester 2 (1:5)	Digester 3 (1:6)
Day 1	34.6%	43.3%	45.9%
Day 3	35.2%	44.1%	46.3%
Day 5	36.9%	45.3%	48.1%
Day 7	37.7%	45.4%	56.6%
Day 9	40.1%	48.2%	57.1%

In Fig. 3, the methane production at different feeding days was depicted. It showed that higher days of feeding has more potential of methane production.

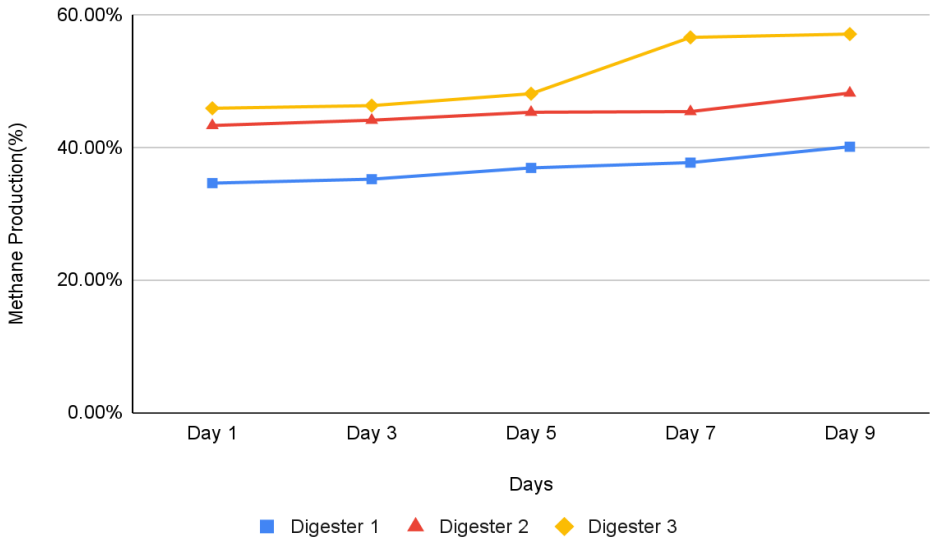


Figure 3. Methane production versus feeding days

From the experiment, it is concluded that maximum production of biogas (methane) from rice straw through anaerobic digestion occurs at 1:6 feed ratio producing 57.1% methane. The results for methane production at different feed ratios are represented in Table 4.

Table 4. Methane production at different feed ratios

FEED RATIO in kg (Rice Straw: Water)	1:4	1:5	1:6
METHANE PRODUCTION	40.1%	48.2%	57.1%

Temperature Variation of feed

Rice Straw feed is taken at different temperature ranges, Psychrophilic (35-40℃), Mesophilic (45-50℃) and Thermophilic (55-60℃) for different feed ratios (1:4, 1:5, 1:6) in biogas digesters to check variations in production of biogas from anaerobic digestion at different temperature and find what is the effect of temperature on overall production and working of bio-enzymes (Table 5).

Temperature has a big impact on methane-forming bacteria's activity. In general, the generation of biogas rises with temperature. Digesters are heated in colder climates because of this reason the ideal temperature range is between 30℃ and 40℃. Methanogen activities are weakened by low temperatures, sudden systemic changes, or both. An essential factor in the effective anaerobic digestion of rice straw is temperature. It is necessary for both economic input and maximized methane production. From the experiment it can be concluded that maximum production of biogas (methane) from rice straw through anaerobic digestion occurs at 45℃ producing 58.5% methane.

Table 5. Methane production at different temperatures

Temperature	Digester 1 (1:4)	Digester 2 (1:5)	Digester 3 (1:6)
35℃	38.40%	46.50%	52.00%
45℃	52.30%	56.80%	58.50%
55℃	45.60%	49.80%	54.70%

CONCLUSIONS

From the experiments conducted it can be determined that maximum production of biogas (methane) occurs when the feed ratio of rice straw and water is 1:6. Production of biogas increases by approx. 9% when ratio is changed from 1:5 to 1:6 and by 17% when it is changed from 1:4 to 1:6. This concludes that changing feed ratio has a big impact on biogas production by anaerobic digestion.

Similarly, analysis of the experiments also shows that maximum production of biogas (methane) occurs when temperature of the sludge is kept between 40-50℃. Raising the temperature has a number of benefits, including making organic compounds more soluble, which makes them easier to accessible to microorganisms, accelerating biological and chemical reactions, enhancing chemical-physical characteristics like the diffusivity of dissolved substrates, boosting the rate at which liquids transfer to gasses, quickening the pace at which pathogenic bacteria die, and generating a more vigorous oxidation reaction at high temperatures. But a significant increase in temperature is also harmful as it tends to kill the bacteria responsible for biogas production. Experiment shows that production of biogas increases by 10-15% when temperature is increased from 35℃ to 45℃. There is also a decrease in production when temperature is further increased to 55℃, production decreases by 7-8% at high temperatures.

REFERENCES

[1] N. Jain, A. Bhatia, H. Pathak, Emission of Air Pollutants from Crop Residue Burning in India, Aerosol Air Qual Res 14 (2014) 422–430. https://doi.org/10.4209/aaqr.2013.01.0031

[2] D. Sietske Boschma, I. Kees, W. Kwant, Rice straw and Wheat straw Potential feedstocks for the Biobased Economy, 2013. www.wageningenur.nl/fbr

[3] S.J. Thomas, A.G. Kumar, S. Thomas, S.S. Sahoo, Sustainability Analysis On Rural Livelihood Systems With Interventions In Energy-Water-Food Nexus, International Journal of Energy for a Clean Environment 23 (2022) 41–52. https://doi.org/10.1615/InterJEnerCleanEnv.2021038177

[4] O.M. Buraimoh, O.I. Olusola, I. Humphrey, I.A. Raheem, U.K. Onyekachi, O.T. Ogundipe, Bioconversion Of Organic Wastes For Sustainable Biogas Production Using Locally Sourced Materials, International Journal of Energy for a Clean Environment 21 (2020) 201–219. https://doi.org/10.1615/InterJEnerCleanEnv.2020033074

[5] R.R. Romasanta, B.O. Sander, Y.K. Gaihre, Ma.C. Alberto, M. Gummert, J. Quilty, V.H. Nguyen, A.G. Castalone, C. Balingbing, J. Sandro, T. Correa, R. Wassmann, How does burning of rice straw affect CH4 and N2O emissions? A comparative experiment of different on-field straw management practices, Agric Ecosyst Environ 239 (2017) 143–153. https://doi.org/10.1016/j.agee.2016.12.042

[6] G.K. Porichha, Y. Hu, K.T.V. Rao, C.C. Xu, Crop Residue Management in India: Stubble Burning vs. Other Utilizations including Bioenergy, Energies (Basel) 14 (2021) 4281. https://doi.org/10.3390/en14144281

[7] N. Fatima Scholar, A. Chandramauli Assistant Professor, N. Chaddha Assistant Professor, Analyzing Disposal Methods of Crop Residue (Parali), 2020. www.ijisrt.com

[8] A. Alengebawy, Y. Ran, N. Ghimire, A.I. Osman, P. Ai, Rice straw for energy and value-added products in China: a review, Environ Chem Lett 21 (2023) 2729–2760. https://doi.org/10.1007/s10311-023-01612-3

[9] S. Mothe, V.R. Polisetty, Review on anaerobic digestion of rice straw for biogas production, Environmental Science and Pollution Research 28 (2021) 24455–24469. https://doi.org/10.1007/s11356-020-08762-9

[10] H.H.A. Matin, S. Syafrudin, S. Suherman, Rice Husk Waste: Impact on Environmental Health and Potential as Biogas, Jurnal Kesehatan Masyarakat 18 (2023) 431–436. https://doi.org/10.15294/kemas.v18i3.42467

[11] S. Kaur, D. Kumar, M. Singla, R. Dogra, Biogas to Bio-CNG from paddy straw: A review, Int J Chem Stud 8 (2020) 1833–1838. https://doi.org/10.22271/chemi.2020.v8.i3y.9475

[12] Q. Liu, S. Pan, Z. Long, Z. Li, L. Du, Y. Wei, Assessment of Fresh and Dry Rice Straw for Biogas Potential by Anaerobic Digestion, Bioenergy Res 13 (2020) 845–852. https://doi.org/10.1007/s12155-020-10106-x

[13] Q. Liu, S. Pan, L. Zhou, L. Feng, Z. Li, L. Du, Y. Wei, Improving the Biogas Potential of Rice Straw Through Microwave-Assisted Ammoniation Pretreatment During Anaerobic Digestion, Bioenergy Res 15 (2022) 1240–1250. https://doi.org/10.1007/s12155-021-10299-9

[14] S. Sathish, S. Vivekanandan, Parametric optimization for floating drum anaerobic bio-digester using Response Surface Methodology and Artificial Neural Network, Alexandria Engineering Journal 55 (2016) 3297–3307. https://doi.org/10.1016/j.aej.2016.08.010

[15] Raju, Y. Maharani, Comparison of biogas production from various blends of biomass, IOP Conf Ser Earth Environ Sci 1241 (2023) 012067. https://doi.org/10.1088/1755-1315/1241/1/012067

[16] S. Meraj, R. Liaquat, S. Raza Naqvi, Z. Sheikh, A. Zainab, A.H. Khoja, D. Juchelkova, A. Atabani, Enhanced Methane Production from Anaerobic Co-Digestion of Wheat Straw Rice Straw and Sugarcane Bagasse: A Kinetic Analysis, Applied Sciences 11 (2021) 6069. https://doi.org/10.3390/app11136069

[17] D. Ahmed, R. Wagdy, N. Said, Evaluation of biogas production from anaerobic co-digestion of sewage sludge with microalgae and agriculture wastes, Bioresources 14 (2019) 8405–8412. https://doi.org/10.15376/biores.14.4.8405-8412

[18] N. Budhraja, A. Pal, R.S. Mishra, Various treatment methods for enhanced bio-hydrogen production, Mater Today Proc (2023). https://doi.org/10.1016/j.matpr.2023.03.198

[19] M. Ilbas, L.O.A. Antari, M. Sahin, Biogas Production From Domestic Resources As An Alternative Energy Source: A Comprehensive Feasibility Study, 2022. www.begellhouse.com

[20] U.U. Ndubuisi-Nnaji, U.A. Ofon, N.U. Asamudo, V.M. Ekong, Enhanced Biogas and Biofertilizer Production from Anaerobic Codigestion of Harvest Residues and Goat Manure, J Sci Res Rep (2020) 1–13. https://doi.org/10.9734/jsrr/2020/v26i330231

[21] N. Budhraja, Simulation and optimization for biohydrogen production potential of various organic waste via anaerobic digestion, Fuel 360 (2024) 130563. https://doi.org/10.1016/j.fuel.2023.130563

[22] S. Lalhriatpuia, N. Budhraja, K. Pal, Optimization of Liquid Metal Nanocomposites and Biogas Addition Rate Using ANN-GA, Journal of Polymer and Composites 11 (2023) 12–27. https://doi.org/10.37591/JoPC

[23] N. Budhraja, A. Pal, R.S. Mishra, Simulation and Optimization of Biohydrogen Production from Biomass Feed via Anaerobic Digestion, Chem Eng Technol 47 (2024) 706–715. https://doi.org/10.1002/ceat.202300395.

Check for updates

Transactions on Emerging Telecommunications Technologies

WILEY

RESEARCH ARTICLE

SDHO-KGNN: An Effective Knowledge-Enhanced Optimal Graph Neural Network Approach for Fraudulent Call Detection

Pooja Mithoo | Manoj Kumar

Department of Computer Science and Engineering, Delhi Technological University, New Delhi, India

Correspondence: Manoj Kumar (mkumarg@dce.ac.in)

Received: 26 July 2024 | **Revised:** 3 January 2025 | **Accepted:** 15 February 2025

Keywords: CDR data | fraudulent call detection | knowledge-enhanced optimal graph neural network | sheepdog hunt optimization

ABSTRACT

Rapid advancements in mobile communication technologies have led to the progression of telecom scams that not only deplete individual fortunes but also affect social income. Hence, fraudulent call detection gains significance, which not only aims to proactively recognize the frauds, but also alleviate the fraudulent activities to manage external losses. Though the traditional methods, such as rule-based systems and supervised machine learning techniques, actively engage in detecting such fraudulent activities, they fail to adapt to the evolving fraud patterns. Therefore, this research introduces a sheepdog hunt optimization-enabled knowledge-enhanced optimal graph neural network classifier (SDHO-KGNN) approach for detecting fraudulent calls accurately. The effectiveness of the proposed SDHO-KGNN approach is achieved through the combination of the power of graph representation learning with expert insights, which allows the proposed SDHO-KGNN approach to capture complex relationships and patterns within telecom data. Additionally, the integration of the SDHO algorithm enhances model performance by optimizing the discrimination between legitimate and fraudulent calls. Moreover, the SDHO-KGNN classifier captures the intricate call patterns and relationships within dynamic call networks, thereby attaining a better accuracy, precision, and recall of 93.8%, 95.91%, and 95.53% for 90% of the training.

1 | Introduction

Over the past few years, the swift expansion of mobile communication technology has ushered in numerous benefits, revolutionizing the way to connect and communicate. Wireless communication networks have undergone significant development during the past few decades, progressing from 1G in the 1980s to 5G in 2019. Every generation sought to overcome the shortcomings of its predecessor, primarily about latency and mobility, density, security and privacy, and data rates and capacity. Consequently, the future 5G/6G networks intend to provide great flexibility to the network operators concerning the allocation of coverage, speed, and the capacity of diverse use cases. Although the 5G deployment phase has already started [1], its full potential has not yet been reached in mobile communication. Meanwhile, 6G intends to transform communication by integrating the digital, biological, and physical realms to give people new experiences via immersive virtual worlds and augmented intelligence [2]. Specifically, artificial intelligence (AI) and machine learning (ML) techniques are expected to drive 5G by continuing the network softwarization and cloudification initiatives toward the development of intelligent telecommunication networks that led to 6G [3]. To fully realize the potential and benefits of 6G, cyber resilience, privacy, and trust are essential prerequisites. Further, these prerequisites will propel 6G development and are essential to attaining even faster and more dependable

© 2025 John Wiley & Sons Ltd.

connectivity than 5G [4]. Telecommunication fraud [1], often shortened to telecom fraud, involves the fraudulent acquisition of significant public and private assets through deceitful means, such as fabricated narratives conveyed through messages, calls, and various communication channels [5]. As science and technology have progressed, the methods of telecommunications network fraud have evolved as well, becoming increasingly precise and efficient and spreading more extensively [2, 6]. Contemporary fraudsters often adopt behaviors that closely resemble those of typical users, operating in well-coordinated groups. This shift in their approach presents a formidable challenge when it comes to identifying telecom frauds within the vast volume of regular call records [7]. Fraudsters may have various motives, one of which could be to evade service charges or minimize them, ensuring they do not bear the full expense of the actual service. In addition, their intentions may extend beyond cost reduction, with the fraudster's ultimate goal being to exploit the provider's network for profit. These fraudulent activities result in substantial financial losses for the telecom industry [8]. Fraudsters are continuously adopting more advanced fraud tactics to evade detection. They employ various fraud tools and concepts, including Benford's theory, Z scores, computer assisted auditing tools (CAATs), and ratio analysis, to detect financial fraud. In contrast, social network analysis (SNA) serves as a nonfinancial fraud detection method that assists organizations in identifying unusual connections among individuals and entities [9, 10].

Presently, telecom operators have implemented diverse fraud detection systems aimed at mitigating losses stemming from fraudulent activities. Those systems analyze customer data within telecom networks using a range of mathematical techniques to identify potential fraudulent customers; also, those detection methods examine the information formed by the users to discover the suspected user and notify managers, in that way, the loss caused by the frauds is reduced by the telecom operators [11]. Various ML and deep learning approaches are utilized in the field of fraudulent call detection [12]. ML algorithms are employed to analyze call data and recognize patterns indicative of fraudulent activity. This can include supervised learning for classification or anomaly detection techniques. Creating user profiles and monitoring their calling behavior to detect deviations from normal usage, such as sudden changes in call volume or international call activity, is also used to identify fraud. In recent years, research efforts have primarily focused on crafting intricate features for phone call data using expert knowledge and subsequently applying ML algorithms for classification. These encompass conventional ML approaches such as support vector machine (SVM) [13, 14], random forest (RF) [15], as well as deep learning methods like convolutional neural networks (CNN) [16], long short-term memory (LSTM) [12], and recurrent neural networks (RNN) [17]. More recently, graph neural networks (GNN) have come into play for extracting interactive insights among users and making predictions [18].

Most existing fraud detection approaches [19–21] have conducted their experiments using either synthetic data or undisclosed real-world data, making it difficult for other researchers to reproduce their results. Additionally, as CDR typically consists of unstructured data, employing a grid-based data processing method for their analysis can result in the loss of significant structural information, potentially leading to a substantial reduction in detection accuracy [22]. In the GNN [23] model, if the issue of class imbalance is not taken into account during the design of the model, the majority class could disproportionately influence the loss function [24]. This could cause the trained GNN to excessively classify instances from these majority classes, leading to inaccuracies in predicting samples from the minority class, which is our primary concern. When applying these models to graphs with imbalanced class distributions, existing GNN approaches tend to excessively prioritize the majority class, resulting in less effective embedding results for the minority class [25]. Additionally, the Neural Factorization Autoencoder (NFA) based method [26] is employed to detect fraudulent call activity, in which the execution time for the detection system is high. The traditional methods for outlier detection primarily depend upon clustering-based [27] and distance-based [28] perceptions. Nevertheless, these methods of telecom fraud detection depend deeply on the applicability of features to the task of categorization. When the feature space is huge, the occurrence of unrelated, unwanted, unacceptable, or noising features in the data may automatically affect the model's performance [29]. Moreover, the recent approaches have data complication, which typically comprises multilevel, multi-granularity, multidimensional data that make the processing and application of data intricate and miscellaneous [25]. Next, ambiguity, changeability, and repetition of the data make difficulties in telecom fraud detection. Also, mixed data, which frequently encompass non-single heterogeneous data such as categorical and arithmetical data, make the detection process challenging to process the data efficiently [30].

This research introduces an effective framework to identify fraudulent calls using SDHO-knowledge-enhanced optimal GNN in the 5G/6G networks. Initially, the data collected from the CDR graph dataset are verified with the data present in the blacklist table, which provides information about the data as normal or abnormal. When data are deemed to normal, the data are transmitted to its intended destination via the data transmission module. This module verifies the information using knowledge-enhanced optimal GNN, which enriches the graph representation by adding semantic information about nodes, edges, and relationships. The SDHO algorithm is employed to enhance the model's performance and efficiently fine-tune the classifier by incorporating the characteristics of sheepdog and coati optimization. When the information is cross-verified, the classifier determines whether the call is normal or abnormal. If it is determined that the call is abnormal, the data are forwarded to the prevention module, where the process is stopped by the prevention module, and the blacklist table is updated. The primary contribution of this research is further detailed in the ensuing discussion.

1.1 | Sheepdog Hunt Optimization Algorithm (SDHO)

The SDHO algorithm is formulated by combining the hunting traits observed in coatis [31], together with the herding attributes of collies [32]. These characteristics combination eradicates issues like overall instability and computational complexity, and they offer advantages by accelerating convergence speed and elevating classification accuracy.

21613915, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ett.70101 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

1.2 | Sheepdog Hunt Optimization Algorithm-Enabled Knowledge-Enhanced Optimal Graph Neural Network(SDHO-KGNN)

The integration of the SDHO algorithm with the KGNN enables the model to accurately detect fraudulent calls. The model can adapt to evolving fraud tactics by continuously updating its knowledge base, ensuring it remains effective in detecting emerging threats. KGNN significantly enhances fraud call detection by leveraging external knowledge to improve accuracy, reduce false positives, adapt to evolving threats, provide interpretability, and handle large-scale call networks.

The manuscript is further divided into the following sections. Section 2 explains the advantages and challenges of existing research, Section 3 includes the proposed methodology and block diagram, Section 4 describes the experimental results and comparative discussion of this research, and Section 5 expands the conclusion of the research.

2 | Literature Review

Existing methods in the research of fraudulent call detection, with advantages as well as drawbacks, are elaborated in this section.

Xuxu Zheng et al. [33] initiated a technique for detecting fraud, employing similarity and multi-view GNN. In this approach, a weight parameter is employed to increase the importance of fraud samples that have been flagged or labeled. However, there is a risk of overfitting the training data, especially if not enough labeled fraud examples are available. Xinxin Hu et al. [25] deployed a GAT-COBO model for fraud detection which can adapt to evolving fraud tactics and patterns in real-time, making them suitable for dynamic environments. However, GNNs can be challenging to interpret, making it difficult to grasp the logic behind the model's decisions. Yan Jiang et al. [7] designed a HESM approach for fraud detection, which enhances the superiority of the model. However, the model can be challenging to interpret, making it difficult to comprehend the reasoning behind the model's choices.

Zhen Z. and Gao J. [22] suggested a CNN method for fraud detection using CDR data. In the suggested method, an end-to-end model was offered to recognize new distinctive features without automatically designed features. Also, the suggested model achieved satisfying outcomes on the challenges in the detection part of telecommunication. The developed method extracted behavior data from raw CDR effectively, nevertheless, combining it with communicating data was still an unresolved issue. John C. Daka and Mayumbo Nyirenda [34] used a SIM Box model for fraud detection in telecommunication. The system can operate in real-time, allowing for immediate identification and response to fraudulent activities as they occur. However, ANNs can be challenging to interpret. Jianbing Yan [35] presented a C-BiLSTM algorithm for fraud detection. The utilization of intelligent algorithms allows for the automated and efficient identification of fraudulent call information, potentially reducing the workload for human operators. The output of intelligent algorithms may be challenging to interpret, making it difficult to understand the reasoning behind specific decisions. Gao P. et al. [36] suggested a reinforced cost-sensitive graph network (RCGN) to detect the head of fraud in telecom fraud. The deep deterministic policy gradient (DDPG) technique was utilized to optimize the coefficients of mass dynamically, and a graph was constructed using CDR, and three base classifiers were united to predict the results and final results of classification. The integration of three base classifiers in the RCGN made the detection of fraud effective, however, the usage of a small-size dataset in the RCGN method was the limitation, which lacked the experiments of comparison.

2.1 | Challenges

- In the automated fraudulent call detection approach [12], the LSTM model attained limited performance due to the formation of data, which is not a series at all times to acquire the arrangement relationship between each dimension. Moreover, the DNN model has a major risk of overfitting, which was exposed using the drift evaluation concept.
- Training and using ANN-based models may require substantial computational resources, which could be a limitation for some organizations. The system's efficiency relies significantly on the quality and comprehensiveness of the call traffic data. Inaccurate or partial data can result in false alarms or overlooked instances of fraud [34, 37].
- Mobile social network fraud detection datasets often suffer from class imbalance, where genuine user activities significantly outnumber fraudulent ones. Handling this imbalance effectively is crucial [22].
- GAT-COBO heavily relies on the quality and availability of data. Noisy or incomplete data can affect its performance. GNNs can be challenging to interpret, making it difficult to comprehend the reasoning behind the model's choices [25].
- Detecting temporal patterns in telecom fraud, and enabling real-time detection and adaptation also face challenges related to data quality, complexity, interpretability, and resource requirements, which need to be carefully considered when applying the approach in practice [38].

3 | Sheepdog Hunt Optimization Algorithm-Enabled Knowledge-Enhanced Optimal Graph Neural Network Classifier

Detecting and mitigating fraudulent actions like telecom fraud, phishing calls, and robocalls that can result in financial setbacks for both service providers and customers is essential to enhance cyber resilience, privacy, and trust in the evolving 5G/6G networks [33]. The major goal of this research is to detect fraudulent calls utilizing the SDHO-knowledge-enhanced optimal GNN in the evolving 5G and 6G networks that often give network operators a significant deal of flexibility in allocating speed, coverage, and capacity for enhancing telecommunication. Initially, data will be gathered from the CDR graph dataset, preprocessed, and subsequently cross-referenced with the information contained in the Blacklist table. In the realm of telephone calls, a blacklist table serves as a repository of specific phone numbers or entities regarded as undesirable or potentially harmful. Telecommunication systems use this table as a point of reference to

21613915, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ett.70101 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

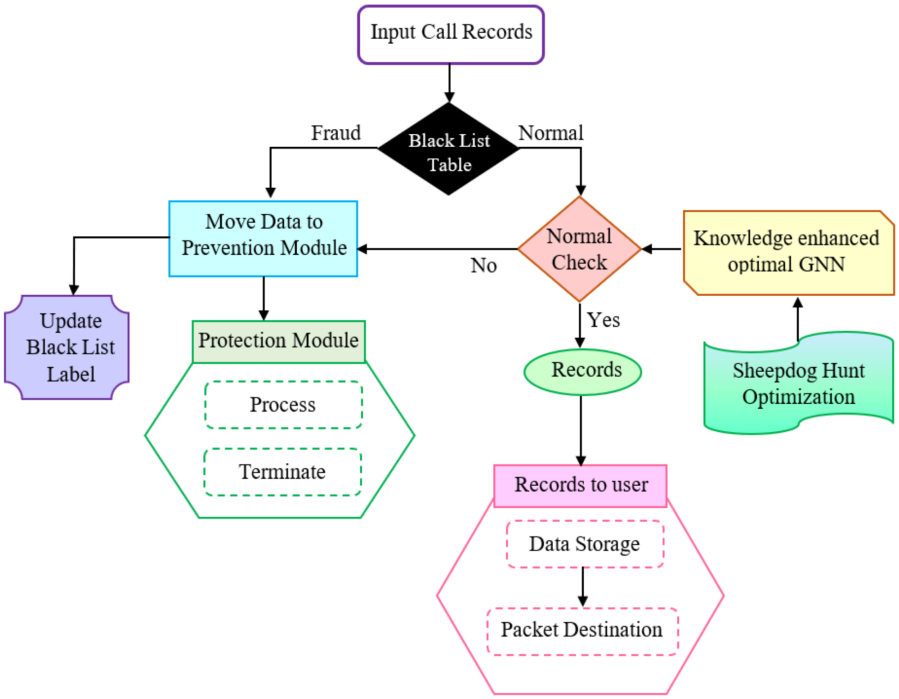


FIGURE 1 | Schematic representation of the fraudulent call detection utilizing SDHO-enabled knowledge-enhanced optimal GNN model.

block or restrict calls originating from or directed to the numbers listed within it. The blacklist table categorizes the data as either normal or abnormal. When data are deemed normal, this is transmitted to its intended destination via the data transmission module. This module validates the information using the knowledge-enhanced optimal GNN [38], which represents an innovative approach that merges the capabilities of GNN with external knowledge sources to augment the model's learning and inference abilities. The integration of external knowledge into GNNs via the knowledge-enhanced GNN framework offers several advantages by enriching the graph's representation with semantic information about nodes, edges, and relationships. This deepens the model's comprehension of the data it processes, leading to improved feature extraction and representation learning. Furthermore, an SDHO algorithm is developed by combining the traits of the sheepdog and the coati. This algorithm facilitates effective tuning of the classifier. The classifier cross-validates the information, and if data are identified as normal, the call is recorded and forwarded to the intended destination. In contrast, if the data are classified as abnormal, it is directed to the prevention module, where the process is terminated, and the blacklist table is updated accordingly. Figure 1 demonstrates the schematic illustration of the fraudulent call detection mechanism utilizing the SDHO-enabled knowledge-enhanced optimal GNN model.

3.1 | Input Telecom Records

Here the data are collected from the CDR graph dataset [39] typically includes structured data representing various aspects of telecommunications activities in 5G/6G networks. These call records serve as the foundational data for analysis. These call records are typically organized as nodes within a graph dataset, with connections (edges) representing relationships or interactions between them.

3.2 | Preprocessing of Call Records

Preprocessing is required to enhance the quality of the data since the raw data typically contain instances that are weakly regulated and may have out-of-range and missing values. Here, the missing values from the CDR graph dataset are handled by removing the rows with non-values, ensuring a cleaner dataset.

3.3 | Blacklist Table

A blacklist table is a database or list containing specific items, such as phone numbers, email addresses, IP addresses, or entities that are considered undesirable, potentially harmful, or suspicious. The primary purpose of a blacklist is to restrict or block access to or communication with the items within it. In telecommunications, blacklists are used to prevent fraudulent activities and are often used to protect individuals' privacy and security. Unwanted behavior serves as the basis for blacklist inclusion, and its dynamic nature keeps it up to date against new threats. The adoption of blacklists is primarily motivated by security and privacy concerns, offering people and organizations a strong defense against known and unknown threats. A proactive strategy, connection with larger security systems, and frequent updates all enhance a blacklist table's ability against possible threats.

3.4 | Training Phase

The knowledge graph neural network (KGNN) is trained on a CDR graph dataset that involves particular procedures designed for telecom data. Here the preprocessed data are ready for the KGNN by encoding call kinds, stamps, and phone numbers into numerical representations. Embedding layers for nodes, capturing distinct representations for every phone number, and

21613915, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ett.70101 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

attention techniques to identify intricate links in call patterns are common features of the model architecture. In training, the KGNN compares and predicts associations inside the CDR graph to minimize a loss function.

3.5 | Knowledge-Enhanced Optimal Graph Neural Network Classifier

Graph-based fraud detection has gained popularity and considerable attention due to the prevalence of fraud incidents that manifest as graph structures. One category of research in this domain centers on identifying unusual nodes [40, 41] and specific abnormal subgraph configurations [42, 43] within static networks. A knowledge-enhanced graph neural network (KGNN) for fraud call detection typically extends a standard GNN architecture with the integration of domain-specific knowledge graphs, which can capture valuable domain-specific information from knowledge graphs which can lead to better performance in fraud call detection. By incorporating external knowledge, KGNNs can generalize better to unseen or rare patterns. KGNNs can transfer knowledge from one domain to another through shared knowledge graphs. This transfer learning capability is valuable in situations where data from one domain can inform predictions in another.

Constructing a graph where nodes represent subscribers in the mobile network, the node features represent the behavior of these subscribers, and the edges represent communication interactions between subscribers enables us to build a network representation. Fraudulent subscribers within the network are potentially identified using GNNs [44].

A user behavior graph is denoted as $U = (B, \chi, A_d, E, N)$ established within a mobile network context. Here, B represents the set of users in the mobile network as nodes, where $B = \{w_1, w_2, w_3, \ldots\ldots\ldots\ldots\ldots w_n\}$. The edges in the graph represent as $E = \{\alpha_1, \alpha_2, \alpha_3 \ldots\ldots\ldots\ldots\ldots \alpha_m\}$. $\alpha_j = (w_{cj}, w_{zj}) \in E$ is an edge between w_{cj} and w_{zj}, where $w_{cj}, w_{zj} \in B$. Additionally, $\chi = \{s_1, s_2, s_3, \ldots\ldots\ldots\ldots\ldots\ldots s_n\}$ denotes the collection of original behavioral features of subscribers, with each $\chi_i \in \Re^d$ representing the behavioral feature vector of the user w_i. These feature vectors are aggregated into the feature matrix $\chi \in \Re^{N\times D}$ for the graph U. The adjacency matrix $A_d \in \Re^{N\times N}$ encodes the relationships within the graph, which $a_{i,j} = 1$ indicates the presence of an edge between nodes i and j, while $a_{i,j} = 0$ signifies a connection. Lastly, $N = \{r_1, r_2, r_3, \ldots\ldots\ldots r_n\}$ constitutes the labels associated with nodes in the set B. In the context of a given subscriber behavior graph, the utilization of GNNs facilitates the learning of node embedding representation. The general formulation of GNNs is given as

$$h^l_{\text{out}} = \text{update}\Big(\text{Agg}\Big(\{h^{l-1}_v \forall v \in N(v)\}W^l_{\text{agg}}\Big), h^l_{\text{in}}, W^l_{\text{update}}\Big) \quad (1)$$

where h^l_{in} and h^l_{out} represents the input and updated node embeddings at the lth layer, correspondingly. h^{l-1}_v refers to the embeddings of neighboring nodes from the previous layer. W^l_{agg} and W^l_{update} denote trainable matrices for the aggregation function and update function at the lth layer, correspondingly. These notations are commonly used to describe the elements and operations in GNN architecture for node embedding and updating processes.

$$h^l_w = \sigma\big([C_l \cdot \text{agg}(\{h^{(l-1)}_v \forall v \in N(v)\}), D_l h^{l-1}_w]\big) \quad (2)$$

where the notations and process described can be summarized as follows: h^l_w represents the node w at the current layer l and h^{l-1}_w represents the previous layer $(l-1)$, $h^{(l-1)}_v$ denotes the node embeddings of neighboring nodes from the previous layer, C_l and D_l are trainable weight matrices, σ represent a nonlinear activation function, agg stand for the aggregator.

The generation of h^l_w follows these steps:

Initially, each node aggregates features from its nearby neighbors using the AGG aggregator, resulting in a unified vector $h^l_{N(w)}$. Then graph SAGE then appends the node's prior representation to its neighborhood feature vector. The concatenated vector is subsequently processed through an MLP with a softmax activation function. The output produced by the MLP becomes the updated representation for the node, which is utilized as input in the subsequent layers. The domain-specific knowledge graph is integrated into the model. The attention mechanism is employed to determine the impact of knowledge nodes on the message-passing procedure, which is denoted as

$$\begin{aligned} h^{(l+1)}_w = {} & \text{propagate}\left(h^{(l)}_w, \{h^{(l)}_v \forall w \in N(v)\}\right) \\ & + Q\left(h^{(l)}_w, \{k_j \forall j \in \text{knowledge graph}\}\right) \end{aligned} \quad (3)$$

where k_j represents the knowledge nodes in the knowledge graph.

The attention mechanism for knowledge graph integration is as follows:

$$\begin{aligned} & Q\left(h^{l-1}_w, \{k_j \forall j \in \text{knowledge graph}\}\right) \\ & \quad = \text{soft max}\left(\{\alpha_j * f\left(h^{(l)}_v, k_j\right) \forall j \in k\text{nowledge graph}\}\right) \end{aligned} \quad (4)$$

The softmax is utilized for binary fraud/non-fraud predictions. Figure 2 shows the architecture of knowledge-enabled optimal GNN.

3.6 | Sheepdog Hunt Optimization

The proposed SDHO algorithm draws inspiration from combining the hunting traits of coati [31] and herding traits of the Border Collie [32] to find the optimal solution that promotes detection accuracy. By incorporating these characteristics, the SDHO algorithm can prioritize making smart decisions based on evolving patterns in the call data.

3.6.1 | Motivation

The SDHO algorithm is utilized to deal with complex high-dimensional problems, which potentially reduce computational bottlenecks in large-scale fraud detection systems. This can lead to faster and more scalable fraud detection. The SDHO algorithm efficiently generated the optimal solution by preventing the algorithm from getting trapped in the local

21613915, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ett.70101 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

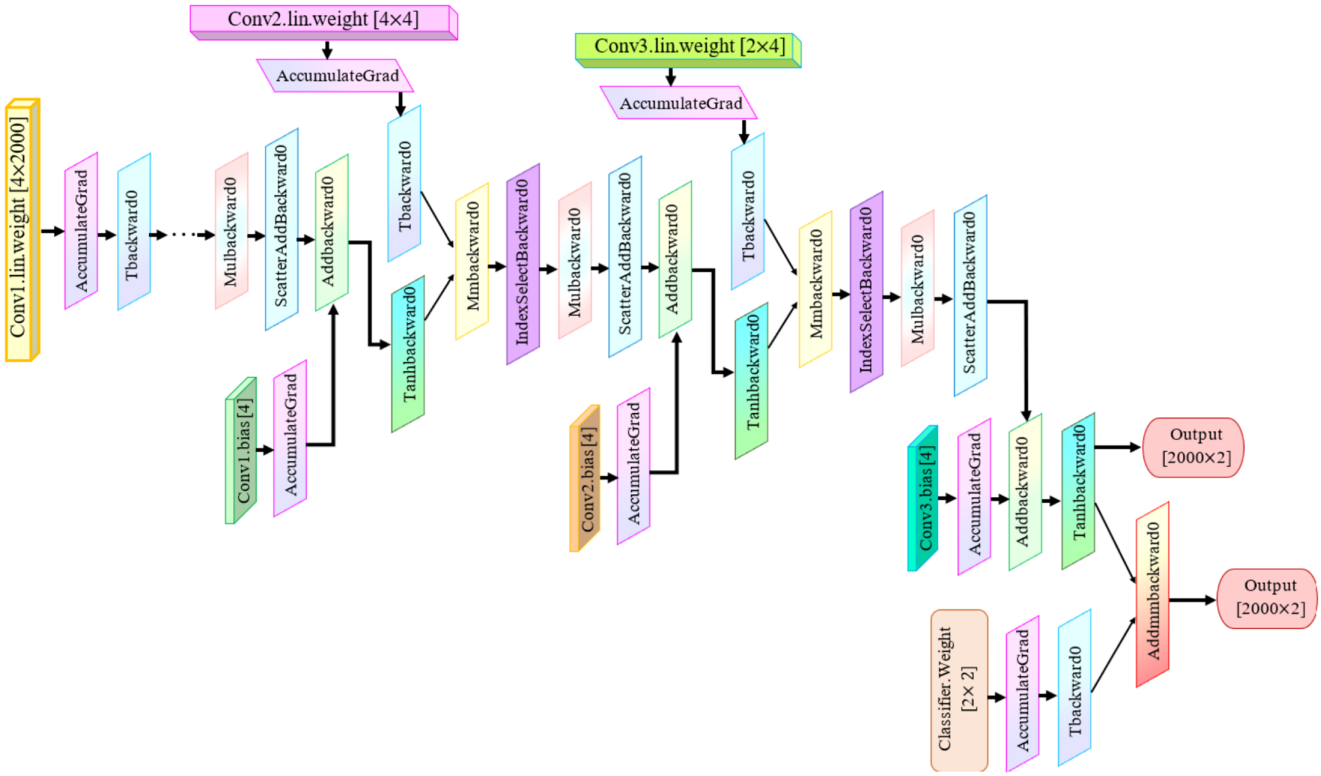


FIGURE 2 | Architecture of knowledge-enabled optimal GNN.

optima. The SDHO algorithm shows its great ability in balancing the search process, this can be particularly useful when dealing with complex call data among various fraud detection mechanisms. Additionally, the SDHO algorithm allows high-speed convergence to offer appropriate values for decision variables in detection tasks.

3.6.2 | Inspiration

Coatis, referred to as coatimundis, belong to the Procyonidae family and are categorized under the *Nasua* and *Nasuella* genera [45]. They are mammals with an omnivorous diet, which includes the consumption of both invertebrate and small vertebrate prey. The diet of these arboreal animals often includes iguanas, which are typically located in trees. Their hunting strategy is characterized by coordinated group efforts, wherein some individuals ascend trees to induce the iguana to descend while others engage in rapid pursuit and capture [46]. Notwithstanding their proficient tactics, they are susceptible to predators and avian predators of substantial size. This behavior can inspire an optimization algorithm that involves multiple agents collaborating to solve complex problems, such as detecting fraudulent calls, where different pieces of information must be gathered and analyzed collectively. *Canis lupus familiaris*, commonly known as sheepdog [32], is an extraordinary dog breed that is renowned for its exceptional herding capabilities and strong work ethic and exhibits a remarkable aptitude for situational assessment and adaptive decision-making [47]. This inspired the development of a metaheuristic algorithm that emulates their behavioral traits. In the context of herding, these canines employ a unique strategy, opting for lateral and frontal approaches rather than traditional rearward herding methods. In the herding context, canines employ a gathering technique where they exert control from the flanks and the front to collect and guide livestock toward a designated location, commonly referred to as gathering. In the process of controlling livestock, canines exhibit a behavior akin to that of wolves. This includes crouching, bowing their heads, raising their hindquarters, and angling their tails downward, a conduct referred to as stalking. Canines emulate the victim selection behavior observed in wolves, referred to as giving an eye or eyeing. In instances where individuals within the group exhibit deviant behavior, these highly intelligent canines employ an intense gaze directed at the individuals in question, applying psychological pressure to guide the collective movement in the desired direction. The algorithm could draw inspiration from their speed and efficiency to develop optimization techniques that operate swiftly and effectively in processing and analyzing large volumes of call data.

3.6.3 | Initialize the Population

Population initialization is the process of creating an initial set of potential solutions that the algorithm will evolve and improve upon during the optimization process. The key goals of population initialization are diversity and coverage of the search space to increase the chances of finding a global optimum.

$$X^{tH} = X^t + r_1\left(I_g - I \cdot X^t\right) + r_2(u - l) \tag{5}$$

where r_2 represents a random number, u and l is denoted as upper and lower bound.

$$r_2 = \frac{r_{\max} - r_{\min}}{r_{\max}} \in [0, 1] \tag{6}$$

21613915, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ett.70101 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$r_1 = e^{\left(\frac{X_{\text{worst}} - X^t}{\sum X_{\text{best}}}\right)} \tag{7}$$

3.6.4 | Fitness Evaluation

The fitness function assesses the solution's quality for every individual within the population; apply the fitness function to calculate its fitness score. This step involves solving or simulating the problem using the individual's parameter values or configuration and measuring its performance based on the fitness function. Higher fitness scores indicate better solutions.

$$\text{Fit} = \max\left(\text{acc}\left(X^t\right)\right) \tag{8}$$

3.6.5 | Claim the Population Gaps

Split the population equally based on fitness. Solutions with higher fitness are regarded as Group 1 (guard), while solutions with fitness below $\text{fit}\left(X^{\text{best}}\right)$ are grouped under Group 2 (herd).

3.6.6 | Herd Population Update

The herd population represents a set of potential solutions or individuals that are being evolved or optimized toward finding an optimal solution for a given problem. These are herd candidates, which obey or follow the best solutions for sooner convergence.

$$X^{tH} = V^{tH} \cdot D^{tH} + \frac{1}{2} A^{tH} \times D^{tH} + r_3\left(X^t_{\text{per}} + X^t\right) \tag{9}$$

where V^{tH} denotes velocity, D^{tH} is the duration of travel, A^{tH} is represented as acceleration.

$$V^{tH} = \left(\sqrt{V^t \cdot \left(\tan\theta_1\right)^2 + 2 \cdot A^t \times X^t}\right) + \delta\left(X_{\text{global}} - X^t\right) \tag{10}$$

where $\delta = 0.5$ when the herd is away from the guard, $\delta = 1$ when the herd is near to guard.

3.6.7 | Guard Population Update

The purpose of the guard population is to dynamically adapt the search behavior of the algorithm, ensuring a proper balance between exploration and exploitation. It helps in achieving an optimal convergence rate while maintaining the diversity of the solutions. The update process of the guard population is typically performed periodically to ensure that the algorithm continuously adapts to the evolving search landscape. By updating the guard population in each iteration, the algorithm can adjust its exploration and exploitation based on the current state of the search.

$$X^{tH} = X^t + \frac{1}{2}\alpha X^{t-1} + \frac{1}{2}\alpha(1-\alpha)X^{t-2} + V^{tH} \times D^t + \frac{1}{2} A^t \cdot \left(T_r\right)^2 + V^{tH} * L \tag{11}$$

Here is the loudness factor to assemble the herd and represents traversal time. A guard should be aware of enemies to protect the herd.

$$V^{tH} = V^t + r_4 c_1 \cdot \left(X_{\text{pee}} - X^t\right) +_5 c_2 \cdot \left(X_{\text{global}} - X^t\right) + D^t c_3 - X^t * \partial^t \tag{12}$$

Distance direction, alertness factor c_2, and awareness coefficient c_1 help guard to protect the herd from enemies.

3.6.8 | Enemy Modulation

The enemies follow a tricky process of the fox, and they try to betray the herd. The vision radius is defined through the observation angle and the scaling parameter, which is represented as,

$$g = \begin{cases} a\frac{\sin(\phi_0)}{\phi_0} & \text{if } \phi_0 \neq 0 \\ \theta & \text{if } \phi_0 = 0 \end{cases} \tag{13}$$

where g represents the vision radius, $\phi_0 \in (0, 2\pi)$ is the observation angle, which is selected for every individual, θ indicates the arbitrary value among 0 and 1, which is set at the beginning of the algorithm, $a \in (0, 0.2)$ represents the parameter of scaling which is set after the iteration of every individual in the inhabitants to model arbitrarily varying the distance from the solution during the enemy approaching. The position of enemies is modulated as

$$\begin{cases} X_0^{\text{new}} = \text{ag} \cdot \cos(\phi_1) + X_0^{\text{actual}} \\ X_1^{\text{new}} = \text{ag} \cdot \sin(\phi_1) + \text{ag} \cdot \cos(\phi_2) + X_1^{\text{actual}} \\ X_2^{\text{new}} = \text{ag} \cdot \sin(\phi_1) + \text{ag} \cdot \sin(\phi_2) + \text{ag} \cdot \cos(\phi_3) + X_2^{\text{actual}} \\ \ldots. \\ X_{n-2}^{new} = ag. \sum_{k=1}^{n-2} \sin(\varphi_k) + ag. \cos(\phi_{n-1}) + X_{n-2}^{actual} \\ X_{n-2}^{new} = ag. \sin(\phi_1) + ag. \sin(\phi_2) + \ldots.. + ag. \sin(\phi_{n-1}) + X_{n-1}^{actual} \end{cases} \tag{14}$$

The location is updated based on the position of the enemy to attain the optimal solution. This model signifies the characteristics of an enemy after it spots a prey and attempts to approach as near as possible to strike, and if it fails when exposed to attempts to access the alternate one likewise.

3.6.9 | Termination

If $t < t_{\max}$, the loop ends, else the process continues to fitness evaluation population split and so on. Figure 3 shows the overall flowchart for the SDHO optimization algorithm and Algorithm 1 demonstrates the pseudocode of the proposed SDHO-KGNN model.

4 | Result and Discussion

This section presents the research findings of the SDHO-enabled knowledge-enhanced optimal GNN classifier for fraudulent call detection clearly and objectively.

4.1 | Experimental Setup

The experimental setup employed the PYTHON software on a Windows 10 operating system equipped with 8GB of RAM. These selections of software and the operating system were deliberately chosen to align with the needs of the implementation. The system's 8GB of RAM provided the necessary computational resources for the successful execution of the experiments.

4.2 | Dataset Description

CDR graph dataset [39]: A CDR (Call Details Record) graph dataset is a collection of structured data that represents telecommunications activities, particularly call and message

21613915, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ett.70101 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

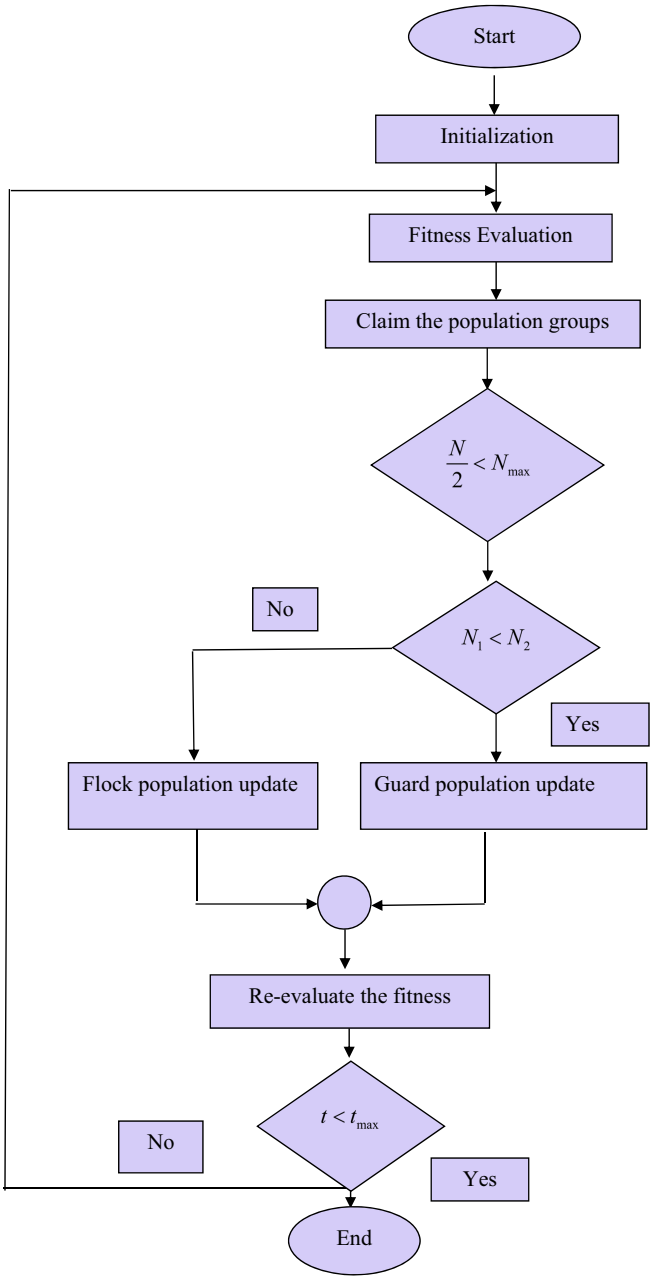


FIGURE 3 | Flowchart for SDHO optimization.

ALGORITHM 1 | Pseudocode for the proposed SDHO-KGNN model.

```
Input: Data obtained from the CDR Graph Dataset
Output: Fraud Call or Non-Fraud Call
Read the Call logs
 Call Pre-Process (D_C) #Pre-processing step initiated.
    {
    p_{D_C} = Imputter(D_C)
    return p_{D_C}
    }
    Call blacklist table #verifies the Blocklist table
    Execute Feature Extraction(D_C) #If incoming is not available
      in the Blocklist table
       {
       FE = (f_{V1}, f_{V1}, ... f_{Vn})
       return FE
       }
                  Call GraphNeuralNetwork (FE) #Features as
                  Input- Fraudulent call detection.
                  {
                  GCN Layer (FE, Hidden Size)
                   GCN Layer (Hidden Size, 2)
                    }
                  Call GNN Training (FE, Label) #Generate
                  Trained Model
                   {
                    Model = GraphNeuralNetwork ()
                    Model.Fit (train data, train label)
                    Call Optimization (Model)
                     {
                      Wg = Model.get weights ()
                      Model = Update hyperparameters (Wg)
                      Return model
                      }
Call Evaluation (model, testing_data) #Model Performance
    Evaluation
{
 Prediction = model.predict(testing data)
 if prediction==Fraud
    {
    Update blacklist table
     }
        Return
          End-}
```

interactions between users. This dataset provides valuable insights into the communication patterns and behaviors of individuals or entities in a network. The dataset comprises information on 101,174 customers, encompassing 17 attributes, including an indicator of whether a customer has churned. Among the customers, there are a total of 8830 churners. Some of the variables in the dataset include state, account length, phone number, international plan, mail plan, number of voice mail messages, total day minutes, and so on. They comprise various data fields, including called number, caller number, time, and date of the call. The attributes of the dataset are Gmail message, day charge, eve mins, night charge, CustServ Calls, eve calls, Intl charge, Intl Mins, night calls, day mins, eve charge, night mins, and Intl calls.

4.3 | Performance Metrics

The metrics utilized to assess the effectiveness of fraud call detection are accuracy, precision, and recall, which are formulated as follows:

$$\text{Accuracy} = \frac{C_{\text{TP}} + C_{\text{TN}}}{C_{\text{TP}} + C_{\text{TN}} + C_{\text{FP}} + C_{\text{FN}}} \tag{15}$$

$$\text{Precision} = \frac{C_{\text{TP}}}{C_{\text{TP}} + C_{\text{FP}}} \tag{16}$$

$$\text{Recall} = \frac{C_{\text{TP}}}{C_{\text{TP}} + C_{\text{FN}}} \tag{17}$$

21613915, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ett.70101 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

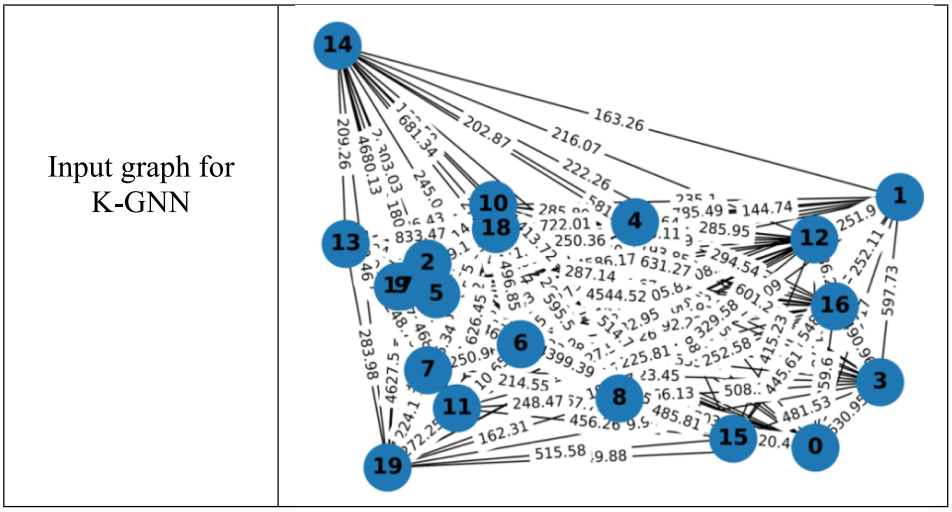


FIGURE 4 | Input graph for KGNN model.

where C_{TP} denotes the true positive, C_{TN} indicates the true negative, C_{FP} is false positive, and C_{FN} is the false negative.

4.4 | Experimental Results

In this section, Figure 4 displays the input for KGNN, this shows the initial formation of the model process, potentially identifying the key factors influencing its predictions. Figure 4 comprehends the nature of input data in a more complex manner, improves transparency, and makes it easier to grasp the KGNN's following outputs.

Figure 5 demonstrates the image results obtained from the experiment. Green points represent the sending nodes, while blue points represent the receiving nodes. Initially, calls from the sender are analyzed using the blacklist table. If a call is identified as fraud, it is routed to the protection module and terminated; otherwise, it proceeds to reach its destination. This identification of fraud calls indicates the model's ability to effectively detect fraudulent calls. This visual representation makes it easier to quickly and thoroughly comprehend the effectiveness of the fraud call detection mechanism. The system's ability to successfully route valid calls to their intended recipients while blocking or ending the fraudulent ones using the blacklist table can be seen in Figure 5.

4.5 | Performance Analysis for Sheepdog Hunt Optimization Algorithm-Enabled Knowledge-Enhanced Optimal Graph Neural Network

The efficacy of SDHO-enabled KGNN is assessed using a CDR dataset, employing the TP metric, and employing various *k*-fold cross-validation schemes with folds labeled 6, 8, and 10. Furthermore, TP values ranging from 50 to 90 are taken into account, along with different epoch sizes of 100–500.

4.5.1 | Performance Analysis for Sheepdog Hunt Optimization Algorithm-Enabled Knowledge-Enhanced Optimal Graph Neural Network With Training Percentage

The performance outcomes of SDHO-enabled KGNN in fraud call detection, considering various epoch values, are visualized in Figure 6. In Figure 6a, the accuracy is measured in terms of TP 90 for epochs 100, 200, 300, 400, and 500, yielding values of 66.83%, 73%, 74.47%, 78.91%, and 93.8%. Similarly, Figure 6b shows precision for TP 90, where the SDHO-enabled KGNN achieves values of 71.17%, 77.61%, 78.98%, 84.88%, and 95.91% for respective epochs. Moreover, in Figure 6c, the SDHO-enabled KGNN shows specificity values of 68.28%, 72.17%, 76.82%, 78.02%, and 95.53% for TP 90, corresponding to the different epoch values.

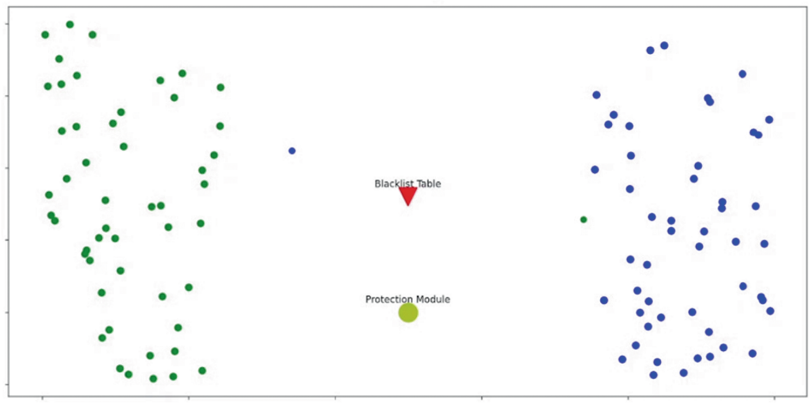
4.5.2 | Performance Analysis for Sheepdog Hunt Optimization Algorithm-Enabled Knowledge-Enhanced Optimal Graph Neural Network With *k*-Fold

The results depict the SDHO-enabled KGNN's performance in detecting fraud calls based on various epoch values as shown in Figure 7. Figure 7a presents the evaluation of SDHO-KGNN with epoch values 100, 200, 300, 400, and 500 attains an accuracy of 68.27%, 68.67%, 70%, 71.65%, and 94.25%, respectively for *k*-fold 10. Figure 7b presents the evaluation of SDHO-enabled KGNN with epoch values of 100, 200, 300, 400, and 500, focusing on precision within the context of a 10 *k*-fold cross-validation. The corresponding sensitivity values achieved are 65.61%, 65.99%, 67.79%, 68.46%, and 96.4%. Moreover, Figure 7c demonstrates the SDHO-enabled KGNN's performance, the attained recall values are 60.15%, 69.12%, 73.43%, 74.31%, and 83.15% in terms of *k*-fold 10 for respective epochs.

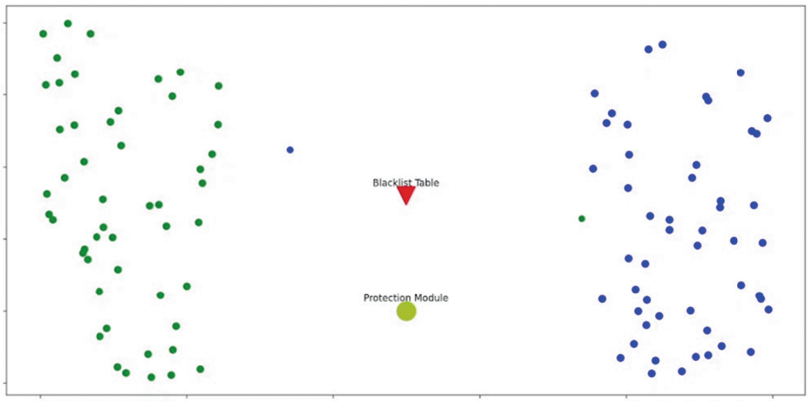
4.6 | Comparative Methods

The following techniques are used in the comparison: neural network (NN) [48], RNN [17], BiLSTM [35], deep CNN (DCNN) [49], GNN [50], graph neural network with BCO

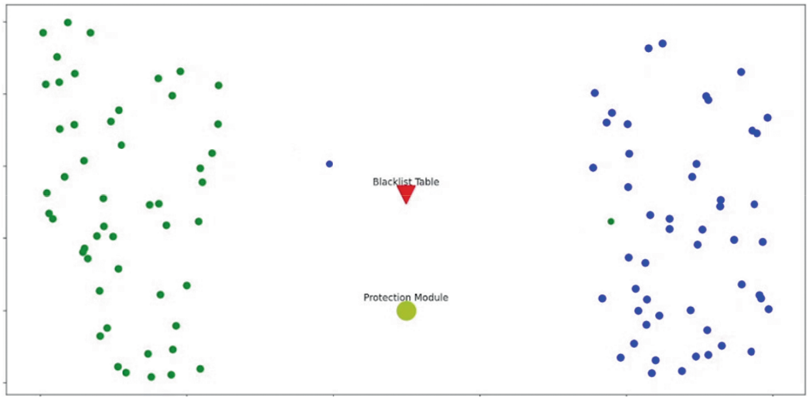
21613915, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ett.70101 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



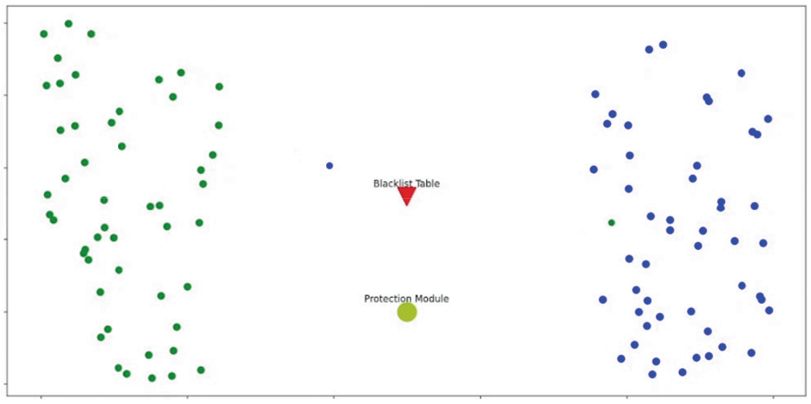
Frame 1



Frame 2



Frame 3



Frame 4

FIGURE 5 | Experimental results of the developed model

21613915, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ett.70101 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

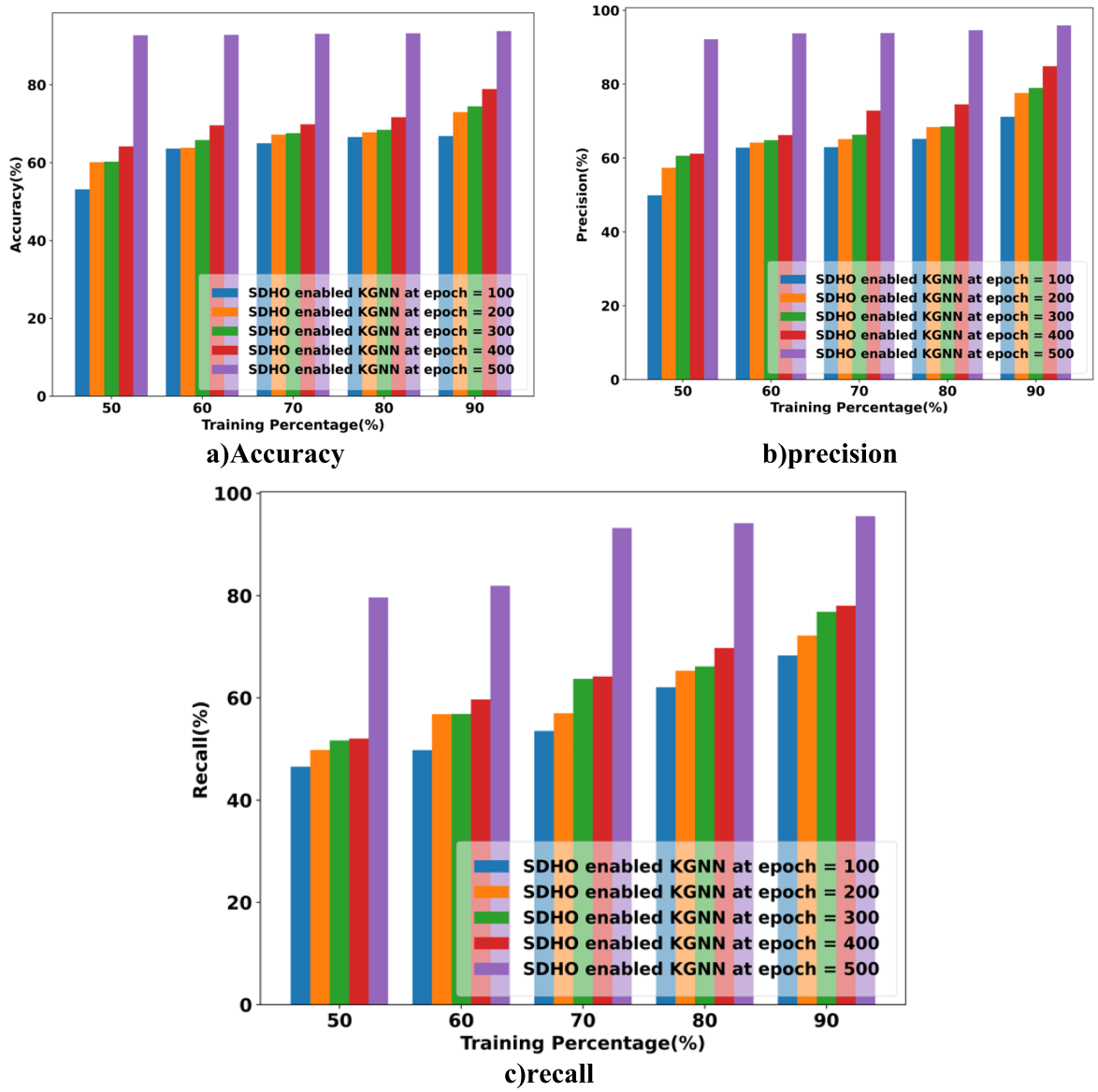


FIGURE 6 | Performance analysis with TP.

(GNN-BCO) [51], graph neural network with COA (GNN-COA) [44], SDHO-enabled KGNN.

4.6.1 | Comparative Evaluation With Training Percentage Analysis

Figure 8 visually illustrates a comparison of the SDHO-enabled KGNN accuracy, precision, and recall strategies with those of other methods. As shown in Figure 8a, the accuracy of SDHO-enabled KGNN at TP 90 is 93.8%, an improvement of 43.41% over the current NN, 27.61% over RNN, 27.49% over BILSTM, 23.77% over DCNN, 22.86% over GNN, 16.81% over GNN-BCO, and 12.53% over GNN-COA. This improvement is due to the use of SDHO, which enables the KGNN to make accurate predictions by tuning its parameters. A similar increase of 47.61% compared to NN, 27.62% compared to RNN, 26.49% compared to BILSTM, 23.78% compared to DCNN, 22.57% compared to GNN, 15.97% compared to GNN-BCO, and 4.87% compared to GNN-COA can be seen in Figure 8b, where the precision of the suggested technique at TP 90 is assessed at 95.91%. This shows that the model reduces the false positives and enhances the model's ability to precisely identify the relevant instances. Additionally, in Figure 8c, it is evident that the recall achieved by the SDHO-enabled KGNN reaches 95.53% at TP 90, showing that the model excels in identifying positive cases. This demonstrates a significant enhancement compared to NN by 33.18%, RNN by 31.56%, BILSTM by 37.9%, DCNN by 19.86%, GNN by 19.86%, GNN-BCO by 19.57%, and A7 by 14.39%. The enhancement in these outcomes indicates that the suggested approach surpasses the other conventional techniques, proving its effectiveness in identifying fraudulent calls.

4.6.2 | Comparative Analysis With *k*-Fold

The SDHO-enabled KGNN underwent a comparative analysis, and the results were measured using accuracy, precision, and recall, as illustrated in Figure 9. With a *k*-fold value of 10, the accuracy achieved by SDHO-enabled KGNN reached 94.25%,

21613915, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ett.70101 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

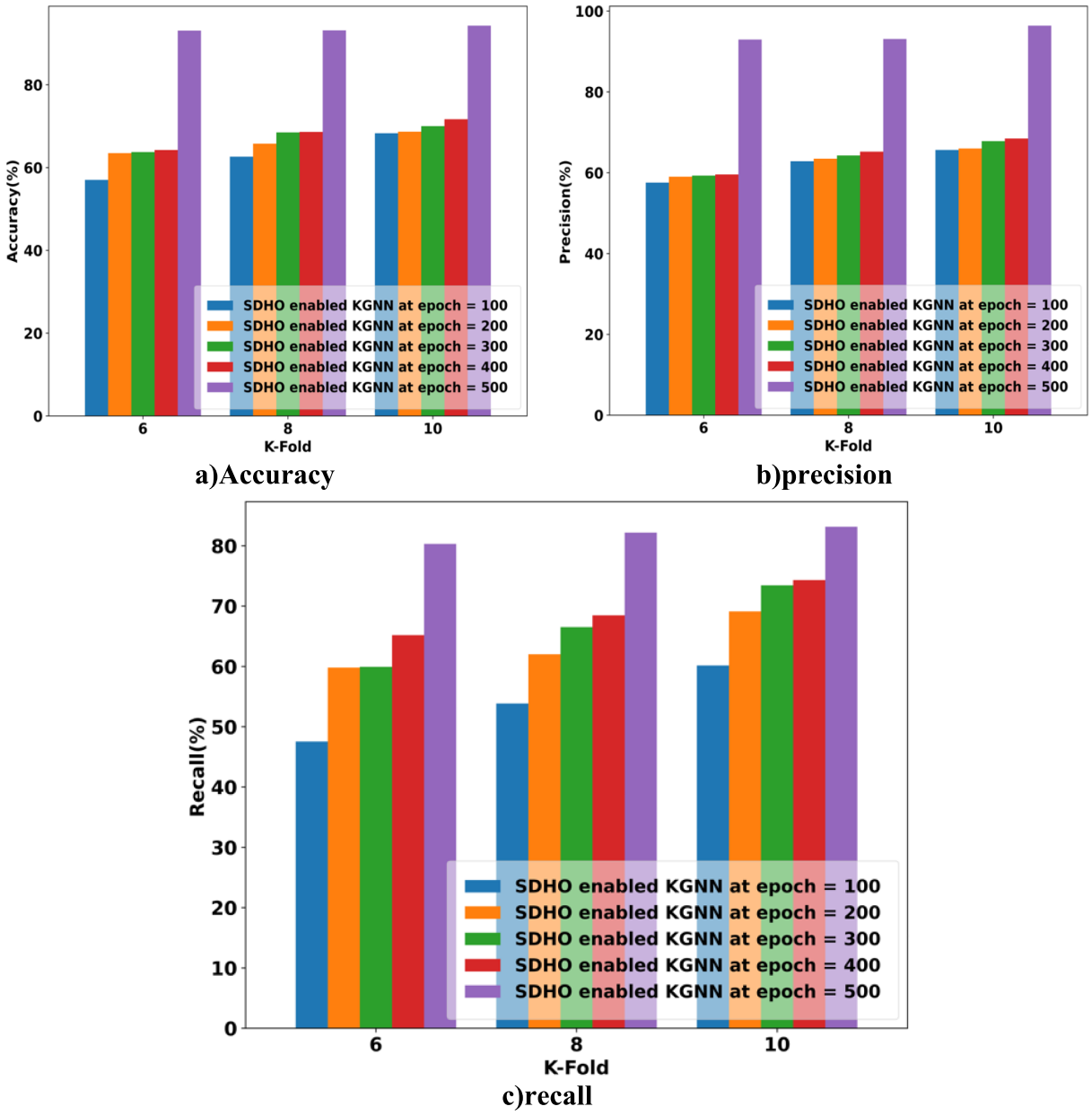


FIGURE 7 | Performance analysis with *k*-fold.

showing effective improvement of 41.19% over the existing NN, 37.26% over RNN, 32.51% over BILSTM, 30.51% over DCNN, 29.03% over GNN, 25.03% over GNN-BCO, and 22.32% over GNN-COA, as shown in Figure 9a. SDHO ensures that the model adapts well to diverse subsets of data during cross-validation. This adaptability allows the KNN to make accurate predictions across diverse folds. Similarly, the precision attained by SDHO-enabled KGNN reached 96.4% for *k*-fold 10, showing improvement of 44.96% over the existing NN, 37.02% over RNN, 33.57% over BILSTM, 33.50% over DCNN, 31.94% over GNN, 29.67% over GNN-BCO, and 27.91% over GNN-COA, as shown in Figure 9b. These improvements show that the model can effectively minimize false positives than other conventional methods. Additionally, for *k*-fold 10, the recall attained by SDHO-enabled KGNN reached 83.15%, showing improvements of 25.59%, 25.46%, 17.14%, 16.87%, 3.64%, 1.6%, and 0.86% over the existing methods NN, RNN, BILSTM, DCNN, GNN, GNN-BCO, and GNN-COA, respectively, as shown in Figure 9c. The improvement in these results shows that the developed method outperforms the other conventional methods, demonstrating its efficiency in detecting fraudulent calls.

4.7 | ROC Analysis

The ROC analysis of the proposed SDHO-enabled KGNN model with existing methods is shown in Figure 10. When the false positive rate is 0.6%, the true positive rates of the existing NN, RNN, BiLSTM, DCNN, GNN, GNN-BCO, GNN-COA methods, and the proposed SDHO-enabled KGNN model are 0.7046%, 0.8427%, 0.8759%, 0.9052%, 0.9501%, 0.9514%, and 0.9518%, respectively. Thus, it is clear that the proposed SDHO-enabled KGNN model achieved higher accuracy in fraudulent call detection when

21613915, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ett.70101 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

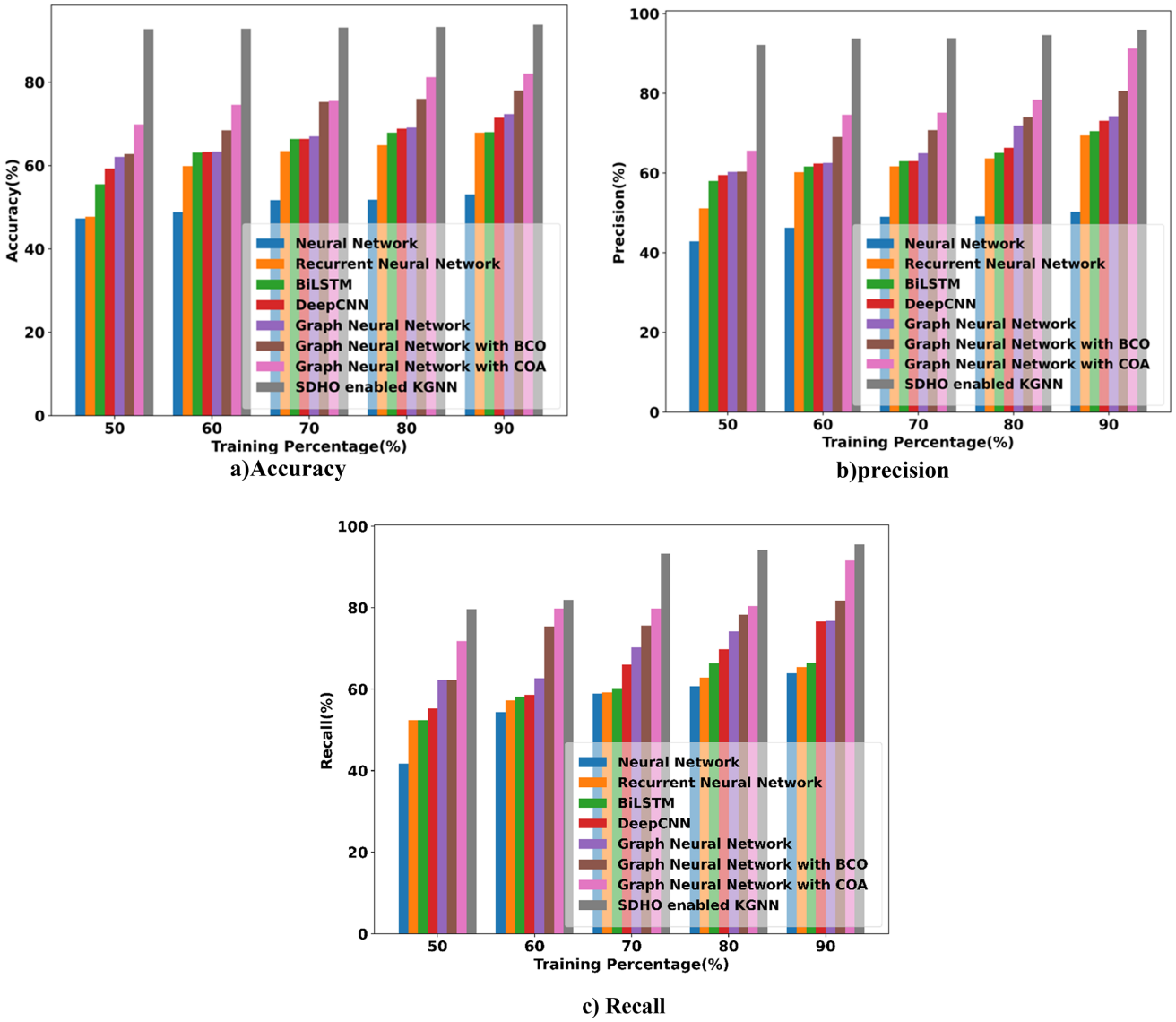


FIGURE 8 | Comparative analysis with TP.

compared to the existing NN, RNN, BiLSTM, DCNN, GNN, GNN-BCO, GNN-COA methods.

4.8 | AUC Analysis

AUC is utilized to evaluate the performance of the classifier, and it is typically calculated using the methods of numerical integration, like the trapezoidal rule, which is applied to the results of ROC analysis. The value of AUC always ranges from 0 to 1, and the variation is clearly shown in Figure 11.

4.9 | Model Loss and Accuracy Graph Analysis

Figure 12 shows the accuracy analysis of the proposed SDHO-enabled KGNN model with the epoch ranges from 0 to 100. At epoch 0, the accuracy of the proposed method is 0.0964, then it rises to 0.324 for epoch 20. At epoch 40, the accuracy gradually increases to 0.545. Further, at epoch 60, it reaches 0.691, 0.802 for epoch 80, and lastly, the accuracy attained for epoch 100 is 0.903.

Figure 13 shows the loss analysis of the proposed SDHO-enabled KGNN model with epoch ranges from 0 to 100. The loss attained by the proposed SDHO-enabled KGNN model are 0.903, 0.417, 0.268, 0.183, 0.129, and 0.096 for epochs 0, 20, 40, 60, 80, and 100, respectively.

4.10 | Comparative Discussion

In this section, the feasibility of fraudulent call detection through automatic feature learning techniques is analyzed to improve the security of telecommunications. Considering the security issues associated with the 5G and future 6G networks, this section will concentrate on the specific techniques related to fraudulent call detection in the context of 5G/6G networks. From the experimental validation, there are still some challenges that limit the applicability of the competent techniques, and the results are depicted

21613915, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ett.70101 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

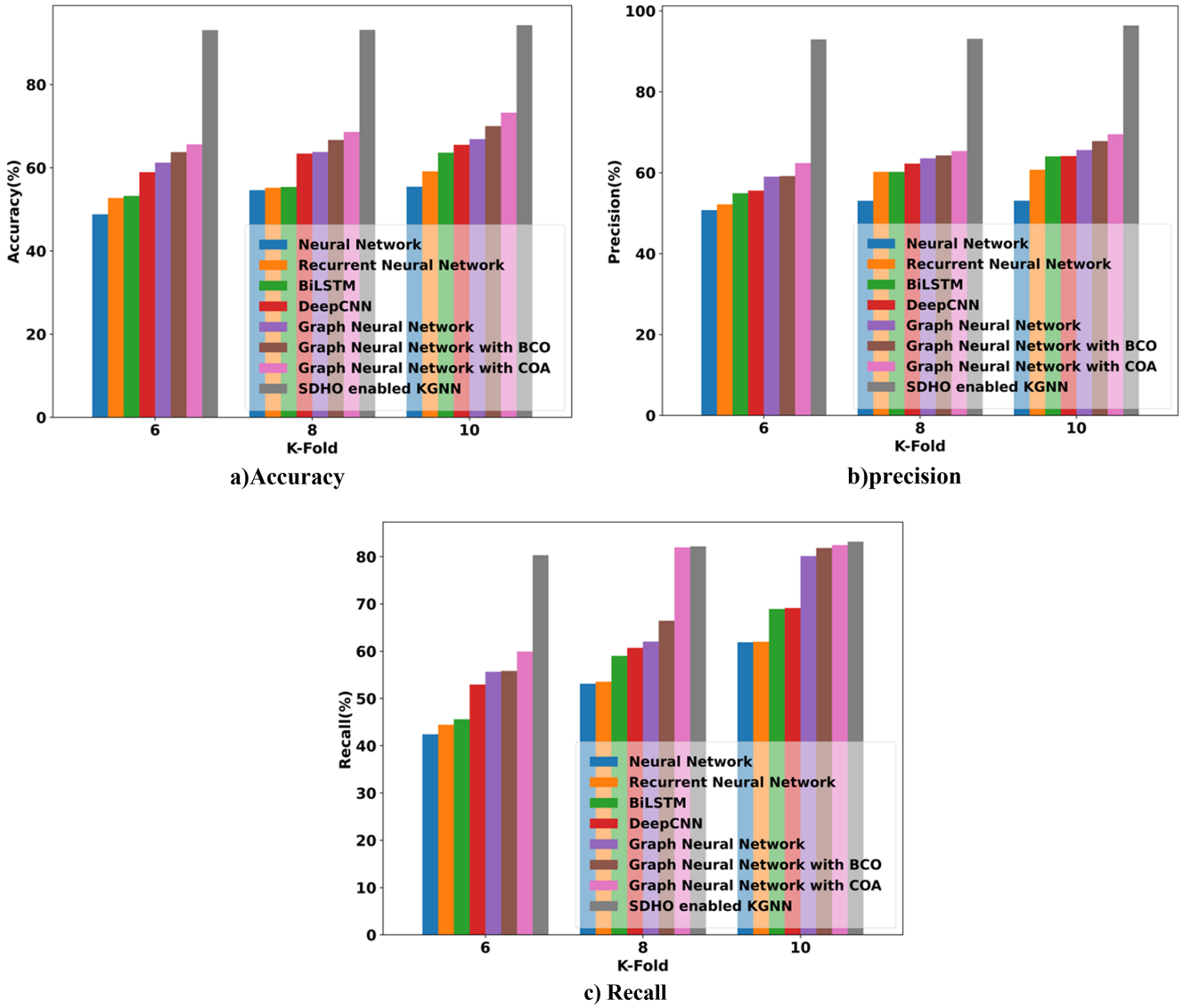


FIGURE 9 | Comparative analysis with *k*-fold.

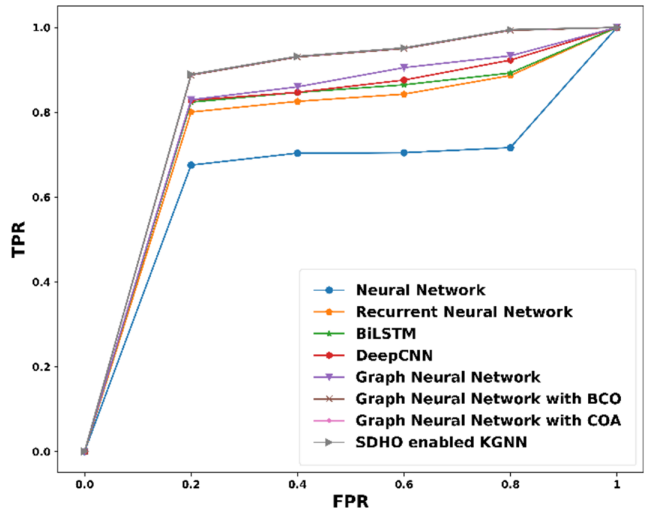


FIGURE 10 | ROC analysis.

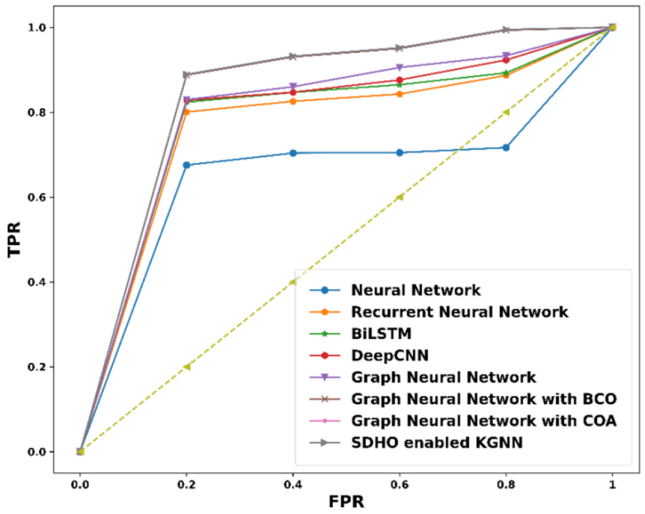


FIGURE 11 | AUC analysis.

21613915, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ett.70101 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

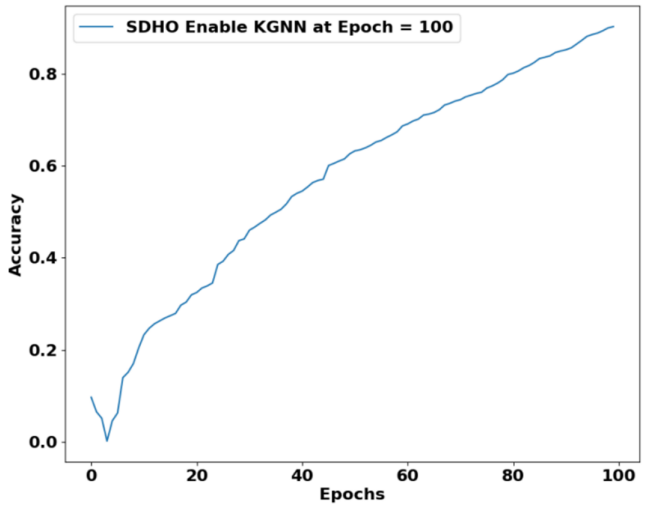


FIGURE 12 | Analysis of accuracy with respective epoch.

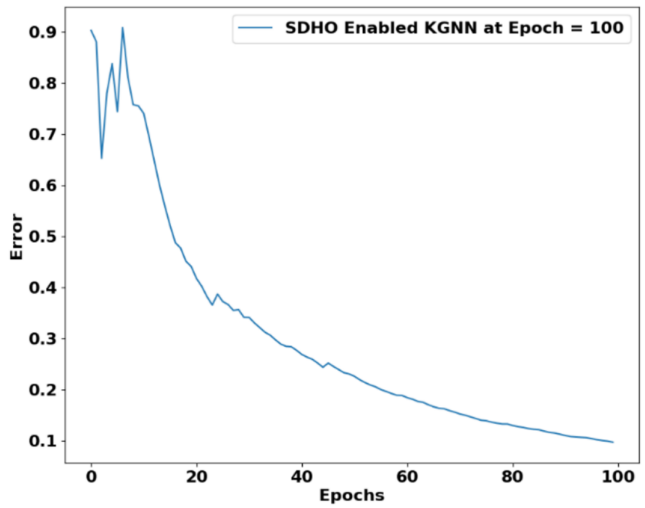


FIGURE 13 | Analysis of loss with respective epoch.

in this section. NN, RNN, BiLSTM, DCNN, GNN, GNN-BCO, and GNN-COA are the conventional approaches utilized for fraud call detection. These approaches contain various drawbacks; training deep learning models, such as DCNN and GNN, can be computationally intensive and may require access to powerful hardware or cloud resources. RNN, BiLSTM, and GNN can be complex and less interpretable. The performance of these models can be notably affected by the quality and accuracy of the data employed for training. Data that are noisy or incomplete can result in less than optimal outcomes. The SDHO-enabled KGNN model is developed to overcome these limitations. KGNNs offer significant advantages for fraud call detection by leveraging external knowledge to improve accuracy, reduce false positives, adapt to changing fraud patterns, and provide interpretability while handling large-scale and dynamic call networks. Remarkably, the proposed SDHO-enabled KGNN model outperforms the other existing techniques, showing its effectiveness in fraud call detection. The incorporation of KGNN helps to learn intricate patterns and relationships, which enhances its ability to detect fraud calls, and the use of SDHO increases the model's performance by tuning the classifier and reduces the computational complexity of the model. High accuracy across diverse training percentages and various *k*-fold ensures that the model performs well consistently demonstrating its ability to adapt and learn diverse amounts of labeled data. The result indicates that the developed method attains a high accuracy of 93.8% in detecting fraudulent calls by adapting to changing fraud patterns. Unlike other conventional methods, the proposed method balances between precision and recall ensuring that the model not only excels in high accuracy but also in identifying fraud calls and minimizes the false positives as well as false negatives. Consequently, the proposed approach learns implicit features that are more resilient to a fraudulent phone number's constantly changing dynamic and call behavior over time and improves the security of 5G/6G networks. Table 1 depicts the comparative analysis between the SDHO-enabled KGNN and other conventional methods with TP 90 and *k*-fold 10.

5 | Conclusion

In this research, a novel fraudulent call detection approach for 5G/6G networks is proposed with advanced strategies to improve the security and privacy of the user. In the realm of fraudulent call detection, the integration of KGNN as a classifier represents a cutting-edge and highly effective approach. KGNNs combine the power of graph-based modeling with external knowledge sources, yielding a range of compelling advantages. By incorporating domain-specific knowledge such as historical fraud patterns and network relationships, KGNNs significantly enhance accuracy and robustness that excel in distinguishing fraudulent calls from legitimate ones, thus minimizing false positives and reducing operational overhead. Moreover, KGNNs adapt to evolving fraud tactics by continuously updating their knowledge base, ensuring a proactive stance against emerging threats. Furthermore, the integration of SDHO-enabled KGNNs to excel in few-shot learning and anomaly detection, enabling the identification of new, previously unseen fraudulent patterns. Customizable to specific organizational needs, these models foster industry collaboration, creating a collective defense against fraud. Specifically, the proposed approach facilitates effective fraudulent call detection and is crucial in the 5G/6G era, as it enhances user privacy and security by terminating the requirement to transfer sensitive information and enhances network performance in terms of security. The experimental results demonstrate that the SDHO-enabled KGNN approach outperforms traditional methods and commonly employed techniques in fraud call detection. Specifically, it achieves an impressive overall accuracy, precision, and recall of 93.8%, 95.91%, and 95.53% with a TP rate of 90%, respectively. Future work could involve refining knowledge integration techniques, enhancing real-time processing capabilities, exploring federated learning for collaborative fraud detection across organizations, and addressing privacy concerns. Additionally, developing methods to handle adversarial attacks on the knowledge-enhanced GNN model is a crucial area for improvement. The proposed model will be extended with a hybrid deep learning mechanism by focusing on a different number of attributes in the data to improve performance that serves as the future direction of research.

21613915, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ett.70101 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 | Comparative discussion for SDHO-enabled KGNN.

Methods/metrics	Training percentage			*k*-fold		
	Accuracy (%)	Precision (%)	Recall (%)	Accuracy (%)	Precision (%)	Recall (%)
NN	53.07	50.24	63.82	55.42	53.05	61.87
RNN	67.9	69.42	65.38	59.12	60.7	61.98
BILSTM	68.01	70.5	66.51	63.61	64.03	68.89
Deep CNN	71.5	73.09	76.55	65.49	64.10	69.12
GNN	72.35	74.26	76.82	66.88	65.60	80.12
GNN-BCO	78.03	80.59	81.78	70	67.79	81.81
GNN-COA	82.05	91.24	91.55	73.22	69.48	82.43
SDHO-enabled KGNN	93.8	95.91	95.53	94.25	96.4	83.15

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

References

1. R. K. Mishra and R. Agarwal, "Impact of Digital Evolution on Various Facets of Computer Science and Information Technology," in *Digital Evolution: Advances in Computer Science and Information Technology* (Bhumi Publishing, 2024), 17.

2. J. Qu and H. Cheng, "Policing Telecommunication and Cyber Fraud: Perceptions and Experiences of Law Enforcement Officers in China," *Crime, Law and Social Change* 82 (2024): 1–23.

3. A. Blika, S. Palmos, G. Doukas, et al., "Federated Learning for Enhanced Cybersecurity and Trustworthiness in 5G and 6G Networks: A Comprehensive Survey," *IEEE Open Journal of the Communications Society* (2024): 1–42.

4. A. Alqabbani, K. Saleem, and A. S. Almazyad, "Digital Communication Forensics in 6G and Beyond Networks," *Applied Sciences* 13, no. 19 (2023): 10861.

5. S. N. Chen, F. Liu, C. X. Gao, and J. Li, "Gearbox Fault Diagnosis Classification With Empirical Mode Decomposition Based on Improved Long Short-Term Memory," in *2021 IEEE 6th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA)*, (IEEE, 2021), 568–575.

6. A. Shah, S. Varshney, and M. Mehrotra, "Threats on Online Social Network Platforms: Classification, Detection, and Prevention Techniques," in *Multimedia Tools and Applications* (Springer, 2024), 1–33.

7. Y. Jiang, G. Liu, J. Wu, and H. Lin, "Telecom Fraud Detection via Hawkes-Enhanced Sequence Model," *IEEE Transactions on Knowledge and Data Engineering* 35, no. 5 (2022): 5311–5324.

8. J. Mundia, E. K. Miriti, S. Mburu, A. M. Kahonge, and C. K. Chepken, "Assessing Mobile Network Fraud Threats and Prevention Strategies in Kenya," *East African Journal of Information Technology* 7, no. 1 (2024): 279–300.

9. J. Wang, C. Jiang, L. Zhou, and Z. Wang, "Assessing Financial Distress of SMEs Through Event Propagation: An Adaptive Interpretable Graph Contrastive Learning Model," *Decision Support Systems* 180 (2024): 114195.

10. V. Harinam and B. Ariel, "Agent-Based Modelling for Criminal Network Interventions," in *Law Enforcement Strategies for Disrupting Cryptomarkets: A Practical Guide to Network Structure, Trust Dynamics, and Agent-Based Modelling Approaches* (Springer Nature Switzerland, 2024), 127–171.

11. O. E. Akinbowale, H. E. Klingelhöfer, and M. F. Zerihun, "Investigating the Level of Effectiveness of the Anti-Fraud Technologies Employed by the South African Banking Industry for Cyberfraud Mitigation," *Journal of Financial Crime* 31, no. 1 (2024): 201–225.

12. J. Xing, M. Yu, S. Wang, Y. Zhang, and Y. Ding, "Automated Fraudulent Phone Call Recognition Through Deep Learning," *Wireless Communications and Mobile Computing* 2020 (2020): 1–9.

13. A. Bakumenko and A. Elragal, "Detecting Anomalies in Financial Data Using Machine Learning Algorithms," *System* 10, no. 5 (2022): 130.

14. O. A. Wahab, A. Mourad, H. Otrok, and T. Taleb, "Federated Machine Learning: Survey, Multi-Level Classification, Desirable Criteria and Future Directions in Communication and Networking Systems," *IEEE Communications Surveys & Tutorials* 23, no. 2 (2021): 1342–1397.

15. C. Lu, S. Lin, X. Liu, and H. Shi, "Telecom Fraud Identification Based on ADASYN and Random Forest," in *2020 5th International Conference on Computer and Communication Systems (ICCCS)*, (IEEE, 2020), 447–452.

16. T. Berhane, T. Melese, A. Walelign, and A. Mohammed, "A Hybrid Convolutional Neural Network and Support Vector Machine-Based Credit Card Fraud Detection Model," *Mathematical Problems in Engineering* 2023, no. 1 (2023): 8134627.

17. S. M. Gowri, G. S. Ramana, M. S. Ranjani, and T. Tharani, "Detection of Telephony Spam and Scams Using Recurrent Neural Network (RNN) Algorithm," in *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, (IEEE, 2021), 1284–1288.

18. S. Ji, J. Li, Q. Yuan, and J. Lu, "Multi-Range Gated Graph Neural Network for Telecommunication Fraud Detection," in *2020 International Joint Conference on Neural Networks (IJCNN)*, (IEEE, 2020), 1–6.

19. R. M. Aziz, M. F. Baluch, S. Patel, and A. H. Ganie, "LGBM: A Machine Learning Approach for Ethereum Fraud Detection," *International Journal of Information Technology* 14, no. 7 (2022): 3321–3331.

20. G. S. Thejas, S. Dheeshjith, S. S. Iyengar, N. R. Sunitha, and P. Badrinath, "A Hybrid and Effective Learning Approach for Click Fraud Detection," *Machine Learning With Applications* 3 (2021): 100016.

21. M. N. Ashtiani and B. Raahemi, "Intelligent Fraud Detection in Financial Statements Using Machine Learning and Data Mining: A Systematic Literature Review," *IEEE Access* 10 (2021): 72504–72525.

22. Z. Zhen and J. Gao, "CDR2IMG: A Bridge From Text to Image in Telecommunication Fraud Detection," *Computer Systems Science and Engineering* 47, no. 1 (2023): 955–973.

21613915, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ett.70101 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

23. Y. Zeng and J. Tang, "Rlc-Gnn: An Improved Deep Architecture for Spatial-Based Graph Neural Network With Application to Fraud Detection," *Applied Sciences* 11, no. 12 (2021): 5656.

24. X. Ma, J. Wu, S. Xue, et al., "A Comprehensive Survey on Graph Anomaly Detection With Deep Learning," *IEEE Transactions on Knowledge and Data Engineering* 35, no. 12 (2021): 12012–12038.

25. X. Hu, H. Chen, J. Zhang, et al., "GAT-COBO: Cost-Sensitive Graph Neural Network for Telecom Fraud Detection," *IEEE Transactions on Big Data* 10, no. 4 (2024): 528–542.

26. A. Wahid, M. Msahli, A. Bifet, and G. Memmi, "NFA: A Neural Factorization Autoencoder Based Online Telephony Fraud Detection," *Digital Communications and Networks* 10, no. 1 (2024): 158–167.

27. D. Ali and K. Omer, "Efficient Density and Cluster Based Incremental Outlier Detection in Data Streams," *Information Sciences* 607 (2022): 901–920.

28. Y. Wang and Y. Li, "Outlier Detection Based on Weighted Neighbourhood Information Network for Mixed-Valued Datasets," *Information Sciences* 564 (2021): 396–415.

29. L. Ren, Y. Zang, R. Hu, et al., "Do Not Ignore Heterogeneity and Heterophily: Multi-Network Collaborative Telecom Fraud Detection," *Expert Systems With Applications* 257 (2024): 124974.

30. R. Li, H. Chen, S. Liu, K. Wang, B. Wang, and X. Hu, "TFD-IIS-CRMCB: Telecom Fraud Detection for Incomplete Information Systems Based on Correlated Relation and Maximal Consistent Block," *Entropy* 25, no. 1 (2023): 112.

31. S. S. Sammen, M. Ehteram, Z. Sheikh Khozani, and L. M. Sidek, "Binary Coati Optimization Algorithm-Multi-Kernel Least Square Support Vector Machine-Extreme Learning Machine Model (BCOA-MKLSSVM-ELM): A New Hybrid Machine Learning Model for Predicting Reservoir Water Level," *Water* 15, no. 8 (2023): 1593.

32. S. Sangeetha, A. A. Victoire, and P. Shanthakumar, "Bayesian Fuzzy Clustering-Based Routing Protocol in Heterogeneous Wireless Sensor Network Using Fractional Border Collie Optimization," *International Journal of Communication Systems* 35, no. 18 (2022): e5341.

33. X. Zheng, C. Feng, Z. Yin, J. Zhang, and H. Shen, "Research on Fraud Detection Method Based on Heterogeneous Graph Representation Learning," *Electronics* 12, no. 14 (2023): 3070.

34. J. C. Daka and M. Nyirenda, "Smart Mobile Telecommunication Network Fraud Detection System Using Call Traffic Pattern Analysis and Artificial Neural Network," *American Journal of Intelligent Systems* 12, no. 2 (2022): 43–50.

35. J. Yan, "Intelligent Algorithms for Identification and Defense of Telecommunication Network Fraudulent Call Information Under Legal System," *International Journal of Network Security* 25, no. 2 (2023): 277–284.

36. P. Gao, Z. Li, D. Zhou, and L. Zhang, "Reinforced Cost-Sensitive Graph Network for Detecting Fraud Leaders in Telecom Fraud," *IEEE Access* 12 (2024): 173638–173646, https://doi.org/10.1109/ACCESS.2024.3448260.

37. F. R. Alzaabi and A. Mehmood, "A Review of Recent Advances, Challenges, and Opportunities in Malicious Insider Threat Detection Using Machine Learning Methods," *IEEE Access* 12 (2024): 30907–30927.

38. Y. Lin, S. Du, Y. Zhang, K. Duan, Q. Huang, and P. An, "A Recommendation Strategy Integrating Higher-Order Feature Interactions With Knowledge Graphs," *IEEE Access* 10 (2022): 119290–119300.

39. "Telecom-CDR-Dataset," accessed September, 2023, https://github.com/jamesrawlins1000/Telecom-CDR-Dataset-/tree/master.

40. X. Hu, H. Chen, S. Liu, H. Jiang, G. Chu, and R. Li, "BTG: A Bridge to Graph Machine Learning in Telecommunications Fraud Detection," *Future Generation Computer Systems* 137 (2022): 274–287.

41. Y. Wu, H. N. Dai, and H. Tang, "Graph Neural Networks for Anomaly Detection in Industrial Internet of Things," *IEEE Internet of Things Journal* 9, no. 12 (2021): 9214–9231.

42. Y. Chen, J. Jiang, S. Sun, B. He, and M. Chen, "RUSH: Real-Time Burst Subgraph Detection in Dynamic Graphs," *Proceedings of the VLDB Endowment* 17, no. 11 (2024): 3657–3665.

43. J. Hou, Y. Lei, Z. Peng, W. Lu, F. Zhang, and X. Du, "Efficient Anomaly Detection in Property Graphs," in *International Conference on Database Systems for Advanced Applications*, (Cham: Springer Nature Switzerland, 2023), (120–136).

44. X. Hu, H. Chen, H. Chen, X. Li, J. Zhang, and S. Liu, "Mining Mobile Network Fraudsters With Augmented Graph Neural Networks," *Entropy* 25, no. 1 (2023): 150.

45. M. Dehghani, Z. Montazeri, E. Trojovská, and P. Trojovský, "Coati Optimization Algorithm: A New Bio-Inspired Metaheuristic Algorithm for Solving Optimization Problems," *Knowledge-Based Systems* 259 (2023): 110011.

46. S. Osama, A. A. Ali, and H. Shaban, "A Hybrid of Information Gain and a Coati Optimization Algorithm for Gene Selection in Microarray Gene Expression Data Classification," *Kafrelsheikh Journal of Information Sciences* 4, no. 1 (2023): 1–16.

47. T. Dutta, S. Bhattacharyya, S. Dey, and J. Platos, "Border Collie Optimization," *IEEE Access* 8 (2020): 109177–109197.

48. T. Karthikeyan, M. Govindarajan, and V. Vijayakumar, "An Effective Fraud Detection Using Competitive Swarm Optimization Based Deep Neural Network," *Measurement: Sensors* 27 (2023): 100793.

49. A. Hamrouni, H. Ghazzai, T. Alelyani, and Y. Massoud, "Low-Complexity Recruitment for Collaborative Mobile Crowdsourcing Using Graph Neural Networks," *IEEE Internet of Things Journal* 9, no. 1 (2021): 813–829.

50. Y. V. R. S. Viswanadham and K. Jayavel, "A Framework for Data Privacy Preserving in Supply Chain Management Using Hybrid Meta-Heuristic Algorithm With Ethereum Blockchain Technology," *Electronics* 12, no. 6 (2023): 1404.

51. P. D. Kusuma and A. Dinimaharawati, "Extended Stochastic Coati Optimizer," *International Journal of Intelligent Engineering and Systems* 16, no. 3 (2023): 482–494.

21613915, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ett.70101 by Delhi Technological University, Wiley Online Library on [01/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Sensitivity Analysis of Dielectric Modulated Dual Gate Gallium Nitride Junctionless Accumulation Mode Nanowire FET Based Biosensor

Anupama
Department of Electrical Engineering
Delhi Technological University
Delhi, India
anupama.4789@gmail.com

Sonam Rewari
Department of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
sonam.rewari@dtu.ac.in

Neeta Pandey
Department of Electronics and Communication Engineering
Delhi Technological University
Delhi, India
n66pandey@rediffmail.com

2025 10th International Conference on Signal Processing and Communication (ICSC) | 979-8-3315-1189-0/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICSC64553.2025.10968002

***Abstract*— This study investigates the Dielectric Modulated Dual Metal Gate Gallium Nitride Junctionless Accumulation Mode Nanowire FET (DM-DG-GaN-JAM-NWFET) as a potential biosensor for applications in the biomedical field. To assess its sensing capabilities, neutral and charged biomolecules are analyzed, focusing on the effects of different dielectric constants and charge densities on the device's electrostatic properties. Key parameters, including threshold voltage (V_{TH}) and on-current (I_{on}) are examined to evaluate the sensitivity of device. Additionally, the performance of the Dual Metal Gate Gallium Nitride Junctionless Accumulation Mode (DG-GaN-JAM) Nanowire FET is compared with that of the Single Metal Gate Gallium Nitride Junctionless Accumulation Mode (SG-GaN-JAM) Nanowire FET. The results highlight that the DM-DG-GaN-JAM-NWFET biosensor is highly effective for biosensing applications.**

Keywords*— *biomolecules, dual metal, gallium nitride, junctionless accumulation mode, nanowire

I. INTRODUCTION

With advancements in technology, the size of MOSFETs has significantly decreased. This scaling was primarily aimed at improving device response times and enhancing gate controllability. However, it has also led to pronounced Short Channel Effects (SCEs), which become critical at the nanoscale. These effects include reduced velocity saturation, gate control over the channel, Drain Induced Barrier Lowering (DIBL), and the hot electron effect [1]-[3]. To mitigate these effects, numerous innovative designs have been developed, such as FinFETs, Gate-All-Around (GAA) MOSFETs, and multi-gate FET architectures [4]-[5]. Excessive series resistance and thermal issues in GAA structures often arise from abrupt p-n interfaces at the source-channel and channel-drain. To address these challenges, J.E. Lilienfeld proposed the Junctionless (JL) Transistor in the 1920s [6], later successfully fabricated by Colinge et al. [7]. This design eliminates p-n junctions and doping concentration gradients, simplifying fabrication. However, its highly doped channel can lead to mobility degradation [8]. To resolve this, the Junctionless Accumulation Mode (JAM) FET was proposed [9], which incorporates a lightly doped channel and heavily doped drain and source to enhance mobility, improve conductivity, and minimize parasitic resistance compared to JL FET. FET biosensors have made remarkable progress in electrochemical biosensing. Biomolecules influence the electrical behavior of sensors by modifying the gate terminals. Recently, there has been growing interest in dielectric modulated FET, which utilize the dielectric constant of biomolecules. This structure detects both uncharged and charged biomolecules [5,10]. High current sensitivity is needed to detect biomolecules at low concentrations. Various multigate FET have been explored for detecting both neutral and charged biomolecules, incorporating different types of nanocavities. Dual metal gate (DG) and GAA FETs offer improved frequency performance and lower power consumption. A key challenge in CMOS downscaling is the thinning of the SiO_2 gate dielectric, which has now approached its limit. This thin dielectric causes direct electron tunneling, leading to leakage currents that degrade performance of circuit. High-k dielectrics are employed to address this issue, but they can exacerbate SCEs. These effects can be mitigated by optimizing gate stack structures and controlling the Equivalent Oxide Thickness (EOT).

Advancements in technology drive the adoption of new materials to address the limitations of existing ones. Gallium Nitride (GaN) has gained attention as a potential material for MOS devices because of its advantages: a wide bandgap enabling higher breakdown voltage, strong piezoelectric effects for enhanced carrier concentration, high electron drift velocity for faster applications, and low intrinsic carrier density for high-temperature stability. These properties position GaN as a superior channel material compared to silicon [11,13]. Additionally, GaN-based biosensors are rapidly advancing due to their exceptional sensitivity, fast response, compact size, and potential for multi-array integration [13-14].

This research introduces the combination of Gallium Nitride (GaN) with Junctionless Accumulation Mode Nanowire Field Effect Transistor (JAM-NWFET) technology to enhance the biosensors sensitivity, enabling the detection of different biomolecules such as proteins, enzymes, DNA etc. To meet the growing demand for ultrasensitive biosensing applications, the Dielectric Modulated Dual Metal Gate Gallium Nitride Junctionless Accumulation Mode Nanowire FET (DM-DG-GaN-JAM-NWFET) is introduced. To further improve device performance, gate stack and Gate Material Engineering, such as dual metal gate, are employed. Additionally, a comparison of dual metal gate (DG) in GaN-JAM-NWFET

979-8-3315-1189-0/25/$31.00 ©2025 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:25:37 UTC from IEEE Xplore. Restrictions apply.

with single metal gate (SG) GaN-JAM-NWFET shows that the dual metal gate improves carrier velocity, gate transport efficiency, and diminishes the hot carrier effect. Therefore, the results demonstrate that the DM-DG-GaN-JAM-NWFET biosensor is more apt for ultrasensitive biosensing applications compared to its single metal gate counterpart.

II. DEVICE STRUCTURE

Fig. 1 describes the 2D structure of the Dielectric Modulated Dual Metal Gate Gallium Nitride Junctionless Accumulation Mode Nanowire FET (DM-DG-GaN-JAM-NWFET), designed for use as a biosensor. The parameters of the proposed structure are summarized in Table 1. The source (N_D^+), channel (N_D), and drain (N_D^+) regions are fabricated using Gallium Nitride (GaN). N-type GaN is utilized for the channel due to its superior electron mobility, which enables a high drain current [13]. A 2 nm layer of Aluminium Oxide (Al_2O_3) is formed over the channel, while a Hafnium Oxide (HfO_2) is placed at the center of the nanogap. Al_2O_3 is a highly effective dielectric material employed with GaN to enhance performance. During the fabrication process, the sacrificial material between the lower gate metal and the upper gate oxide layer is etched, forming a nanocavity. This biosensor can be realized as dielectric-modulated (DM) FET. When there are no biomolecules, the surrounding air (K=1) fills the nanogap. The nanogap thickness is 2 nm, the same as the insulating layer. Biomolecules like Biotin (K=2.63), APTES (K=3.57), and DNA (K=8), is placed into the nanogap [5]. Dual Metal Gate (DG), achieved through Gate Material Engineering (GME), divide the gate into two halves: Gate1 (L_{G1}) and Gate2 (L_{G2}), each with different work functions (ϕ_{m1} and ϕ_{m2}). The higher work function, ϕ_{m1}, is positioned near the source, while the lower work function, ϕ_{m2}, is located near the drain. Device simulations were performed using Silvaco Atlas TCAD, which includes models for Shockley-Read-Hall (SRH) recombination, field-dependent mobility (FLDMOB), concentration-dependent mobility (CONMOB), band-to-band tunneling (BBT.STD), Boltzmann statistics, and the Newton-Gummel numerical method.

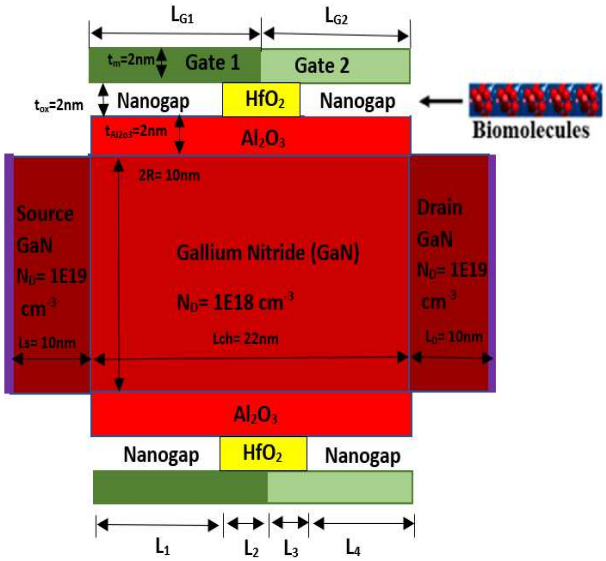


Fig. 1. DM-DG-GaN-JAM-NWFET structure 2-D view

TABLE I. DEVICE PARAMETERS

Various Parameters	Symbols	DM-DG-GaN-JAM-NWFET	DM-SG-GaN-JAM-NWFET
Length of Channel	L_{ch}	22 nm	22 nm
Source/Drain Length	$L_{S/D}$	10 nm	10 nm
Length of Gate1 and Gate2	$L_{G1} = L_{G2}$	11 nm	11 nm
Nano-gap Cavity Length	L_1=L_2	8 nm	8 nm
Length of Region II and Region III	L_3=L_4	3 nm	3 nm
Channel Doping	N_D	10^{18} cm^{-3}	10^{18} cm^{-3}
Source/Drain Doping	N_D^+	10^{19} cm^{-3}	10^{19} cm^{-3}
Gate Work Function	φ_{m1}, φ_{m2}	5.14 eV, 4.36 eV	5.1 eV
Oxide Thickness	t_{Al2O3}	2 nm	2 nm
Hafnium Oxide Thickness	t_{ox}	2 nm	2 nm
Nano-gap Cavity Thickness	t_{ox}	2 nm	2 nm
Channel Radius	R	5 nm	5 nm

III. SIMULATIONS AND RESULTS

This section explores the detection of biomolecules like Biotin, APTES, and DNA, having different dielectric constants, using DM-DG-GaN-JAM and DM-SG-GaN-JAM NWFETs. Figure 2 illustrates the energy band diagrams of the DM-DG-GaN-JAM NWFET and DM-SG-GaN-JAM NWFET for air (K=1) and DNA (K=8) biomolecules. Variation in biomolecules dielectric constant within the cavity have a significant effect on energy band at the channel-drain junction, leading to noticeable changes in device conduction. The presence of DNA molecules increases the Conduction Band Energy (CBE) and Valence Band Energy (VBE) compared to air (K=1). Additionally, the DM-DG-GaN-JAM NWFET exhibits higher CBE and VBE values compared to the SG-GaN-JAM NWFET for both biomolecules. This is attributed to the reduced influence of high-speed electron movement, which mitigates hot carrier effects [12, 14].

Fig. 3 illustrates the drain current variation for various biomolecule dielectric constants in both NWFETs. The threshold voltage of both devices experiences a significant shift when biomolecules are absent in the cavity, as indicated in the figure.

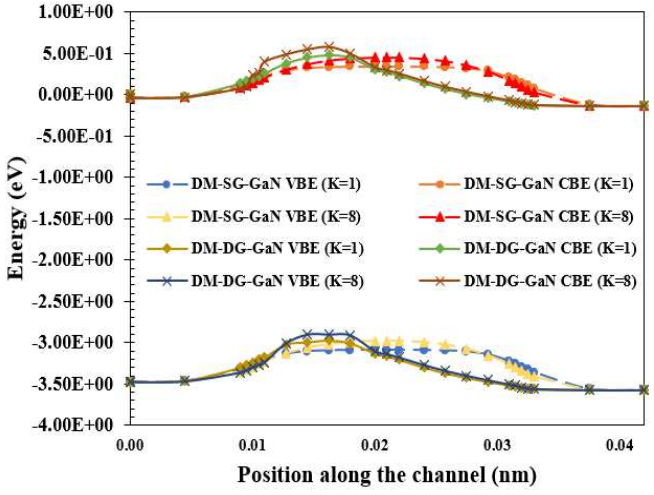


Fig. 2. Energy band diagrams of both devices for air(K=1) and DNA (K=8) biomolecules

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:25:37 UTC from IEEE Xplore. Restrictions apply.

As biomolecules with different dielectric constants fill the nanogaps, the threshold voltage shifts back toward the initial value observed in the absence of biomolecules [5,14]. It is evident from Fig. 3 that the DM-DG-GaN-JAM NWFET exhibits a greater change compared to the DM-SG-GaN-JAM NWFET for sensing Biotin, APTES, and DNA. Due to dual-metal gate (DG) structure in the GaN-JAM NWFET, the impact of high-speed electrons is reduced, thereby decreasing hot carrier effects, unlike in single-metal gate (SG) structure [5,14]. As a result, the DM-DG-GaN-JAM NWFET shows a more significant change in drain current, enhancing biomolecule detection.

Fig. 4 illustrates logarithmic drain current versus gate voltage, indicating subthreshold (leakage) current. In subthreshold region, leakage current reduces from Biotin to DNA. The greater dielectric constant of DNA, in comparison to Biotin, improves gate control, causing the threshold voltage to shift upward. This makes it more difficult to turn-on the channel, thus reducing the subthreshold leakage current [5,14]. Additionally, the DM-DG-GaN-JAM NWFET shows lower subthreshold leakage current compared to the DM-SG-GaN-JAM NWFET. The dual metal gate (DG) in the GaN-JAM NWFET mitigates the effects of high-speed electrons, minimizing short-channel effects and making it more suitable for biomolecule detection than the single-gate (SG) FETs.

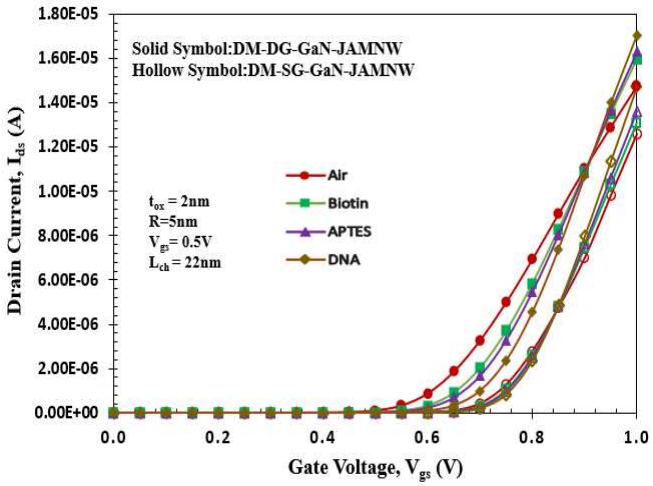


Fig. 3. Drain current (I_{ds}) for various biomolecule dielectric constant in both NWFETs.

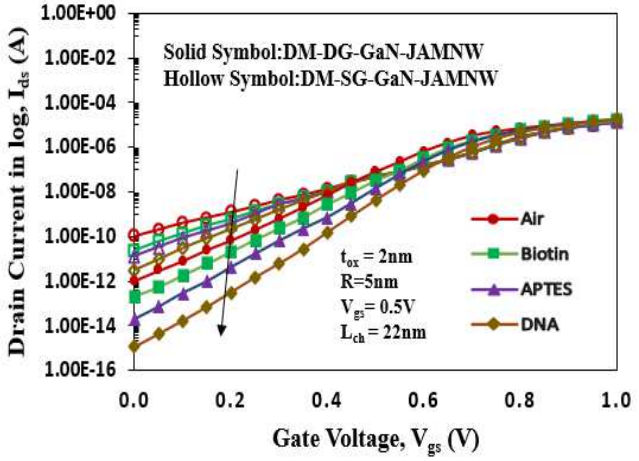


Fig. 4. Drain current (I_{ds}) in log scale for various biomolecule dielectric constant in both NWFETs.

Fig. 5 illustrates the transconductance (G_m) variation for different biomolecule dielectric constant. Transconductance is a key factor in assessing the performance of GaN-based biosensors. A higher transconductance indicates lower power consumption, which is crucial for portable biosensors [5]. As depicted in the figure, the DM-DG-GaN-JAM NWFET has higher transconductance with larger drain current compared to the DM-SG-GaN-JAM NWFET, owing to its improved channel control and reduced SCEs [5].

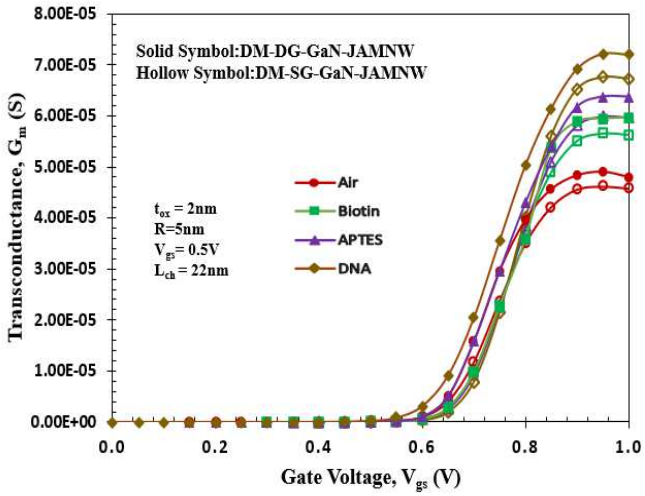


Fig. 5. Transconductance variation for different biomolecule of both devices

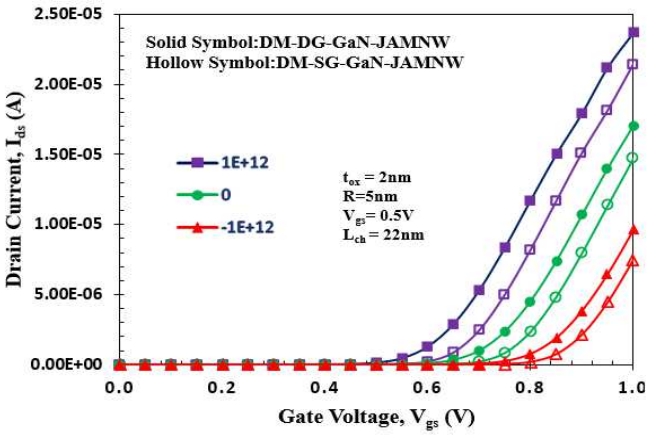


Fig. 6. Drain current (I_{ds}) variation for different DNA biomolecule concentrations in both devices

Fig. 6 demonstrates the simulation of I_{ds} versus V_{gs} for both DM-DG-GaN-JAM NWFET and DM-SG-GaN-JAM NWFET biosensors, with biomolecules of DNA (K = 8) carrying both positive and negative charge densities (Q_f) in the nanogaps. As the positive charge concentration in the nanogap increases, the curves shift toward larger drain currents, resulting in increased channel conductance and a lower threshold voltage (V_{TH}). Conversely, an increase in negative charge density results in a shift toward lower drain currents, with a corresponding increase in V_{TH} [5]. It is evident that the DM-DG-GaN-JAM NWFET outperforms the DM-SG-GaN-JAM NWFET because of its more gate control of the channel, with a higher drain current.

Sensitivity is a critical aspect of sensor design. For FET biosensors, it is typically evaluated by examining the relative change in an electrical characteristic that shows the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:25:37 UTC from IEEE Xplore. Restrictions apply.

presence of biomolecules. This characteristic may include factors like ON-state current (I_{on}), Subthreshold Swing (SS), I_{on}/I_{off}, and threshold voltage (V_{TH}). Mathematically, sensitivity (S) is expressed as [5]:

$$S_{V_{TH}} = \frac{V_{TH,bio} - V_{TH,air}}{V_{TH,air}} \quad (1)$$

$$S_{I_{on}} = \frac{I_{on,bio} - I_{on,air}}{I_{on,air}} \quad (2)$$

$$S_{SS} = \frac{SS_{bio} - SS_{air}}{SS_{air}} \quad (3)$$

$$S_{I_{on}/I_{off}} = \frac{I_{on}/I_{offbio} - I_{on}/I_{offair}}{I_{on}/I_{offair}} \quad (4)$$

The biomolecule sensitivity is assessed by modifying the dielectric constant and observing the corresponding change in threshold voltage. A larger shift in threshold voltage leads to higher sensitivity (S_{VTH}). Fig. 7 clearly demonstrates that DM-DG-GaN-JAM NWFET is more sensitive than DM-SG-GaN-JAM NWFET in sensing Biotin, APTES, and DNA biomolecules. This enhanced sensitivity is attributed to the proposed dual metal gate configuration, which enhances the V_{TH} sensitivity of the biosensor [5]. As the dielectric constant rises, the sensitivity improves, showing a linear relationship between sensitivity and threshold voltage as indicated in Fig. 7 graph.

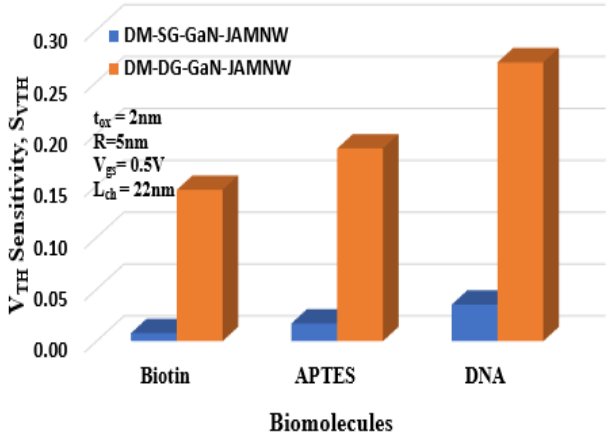


Fig. 7. V_{TH} sensitivity for various biomolecules

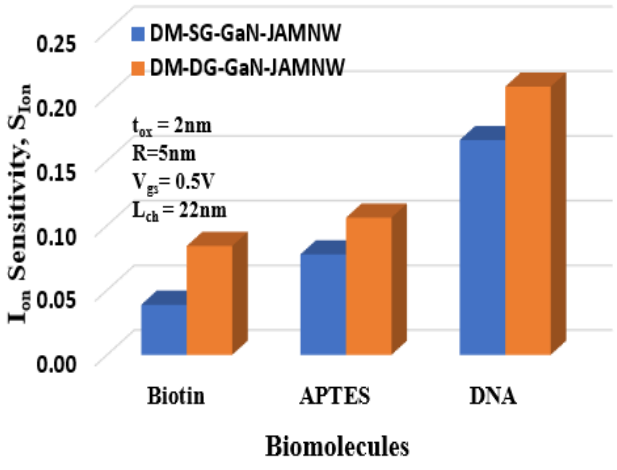


Fig. 8. I_{on} sensitivity for various biomolecules

Fig. 8 illustrates ON current sensitivity for both structures, with the DM-DG-GaN-JAM NWFET showing significantly higher sensitivity than the DM-SG-GaN-JAM NWFET. A higher dielectric constant enhances the interaction between the biomolecule and the gate region, thereby increasing the ON current to its peak value [5].

Fig. 9 and Fig. 10 present the sensitivity of I_{on}/I_{off} ($S_{Ion/Ioff}$) and Subthreshold Swing (S_{ss}) for the two structures under comparison. The DM-DG-GaN-JAM NWFET demonstrates notably superior sensitivity in both metrics compared to the DM-SG-GaN-JAM NWFET. This improvement is due to the higher dielectric constant, which enhances the interaction between the biomolecule and the channel, leading in greater ON current. Additionally, the reduction in the OFF current is facilitated by the dual-gate, which mitigates short-channel effects (SCEs) [5,10].

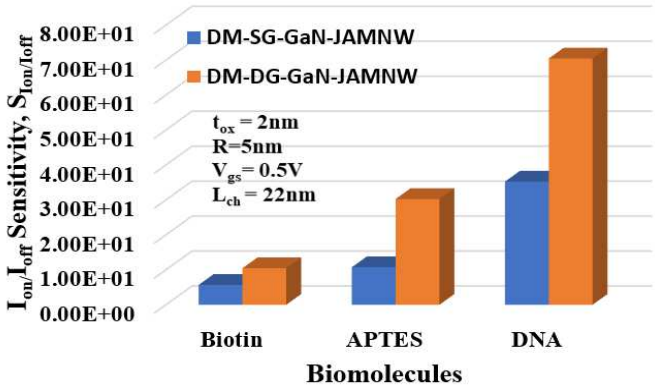


Fig. 9. I_{on}/I_{off} sensitivity for various biomolecules dielectric constant

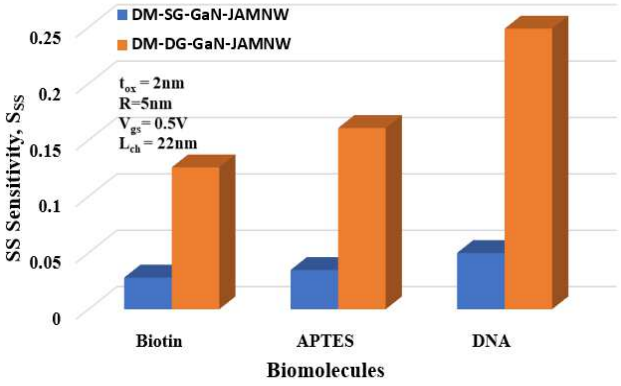


Fig. 10. Subthreshold Swing sensitivity for various biomolecules

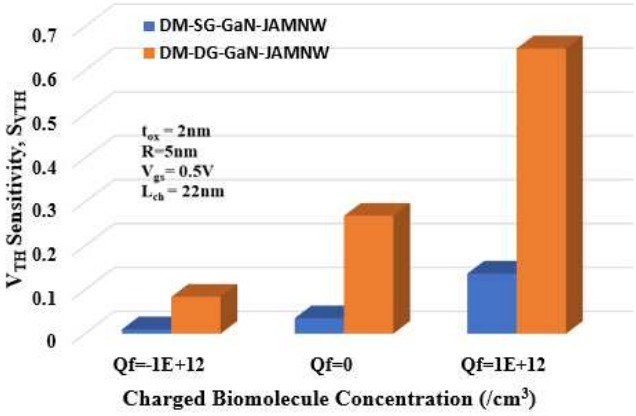


Fig. 11. V_{TH} sensitivity variation with charge density for DNA biomolecule

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:25:37 UTC from IEEE Xplore. Restrictions apply.

TABLE II. COMPARISON PARAMETERS FOR BOTH DEVICES

Devices	DM-SG-GaN-JAM-NWFET				DM-DG-GaN-JAM-NWFET			
Bio-molecules	Air	Biotin	APTES	DNA	Air	Biotin	APTES	DNA
$I_{d_{sat}}$(A)	1.26E-05	1.31E-05	1.36E-05	1.47E-05	1.47E-05	1.60E-05	1.63E-05	1.70E-05
g_m (S)	4.58E-05	5.63E-05	5.97E-05	6.73E-05	4.80E-05	5.97E-05	6.38E-05	7.20E-05
I_{sub}(A)	1.03E-10	2.26E-11	1.26E-11	3.05E-12	1.05E-12	2.08E-13	1.96E-14	1.06E-15
S_{VTH}	-	0.01	0.02	0.04	-	0.15	0.19	0.27
S_{Ion}	-	0.04	0.08	0.17	-	0.08	0.11	0.21
$S_{Ion/Ioff}$	-	5.71E+00	1.08E+01	3.53E+01	-	1.05E+01	3.03E+01	7.05E+01
S_{SS}	-	0.028	0.035	0.050	-	0.126	0.161	0.249

TABLE III. S_{VTH} SENSIVITY COMPARISON OF THE PROPOSED DM-DG-GAN-JAM NWFET FOR APTES BIOMOLECULES

Device specifications	Ref. (S. Sharma et al. 2020) [13]	Ref. (S. Yadav et al. 2024) [10]	DM-SG-GaN-JAM-NWFET	Proposed DM-DG-GaN-JAM-NWFET
Cavity thickness	4 nm	2 nm	2 nm	2 nm
Cavity length (nm)	30 nm	50 nm	16 nm	16 nm
Channel length	50 nm	100 nm	22 nm	22 nm
Dielectric constant of biomolecules	APTES (K=3.57)	APTES (K=3.57)	APTES (K=3.57)	APTES (K=3.57)
S_{VTH} Sensivity	0.05879	0.0427	0.02	0.19

Fig. 11 illustrates the variation in V_{TH} sensitivity with charge density for DNA biomolecules. The DM-DG-GaN-JAM NWFET exhibits superior sensitivity compared to other devices in detecting various concentrations of charged biomolecules. This enhanced sensitivity can be attributed to the dual-metal gate structure, which mitigates the influence of high-speed electron movement, thereby enhancing biomolecule detection efficiency [5,14].

A detailed comparison of various parameters is enlisted in Table II. Table III presents a comparison of the proposed DM-DG-GaN-JAM NWFET with recently published devices that detect the same biomolecule, APTES (K=3.57), ensuring a fair comparison for detection. The comparative analysis highlights that the DM-DG-GaN-JAM NWFET excels in S_{VTH} sensitivity, outperforming over existing structures.

IV. CONCLUSION

Dielectric Modulated Dual Metal Gate Gallium Nitride Junctionless Accumulation Mode Nanowire FET (DM-DG-GaN-JAM-NWFET) has been introduced as a biosensor due to its enhanced detecting capabilities. Sensitivity and drain current variation have been used as key biosensing factors for detecting biomolecules in the nanogap. This research also includes a comparison between the DM-DG-GaN-JAM and DM-SG-GaN-JAM NWFETs based biosensors. The DM-DG-GaN-JAM-NWFET exhibits a larger variation in drain current compared to the DM-SG-GaN-JAM-NWFET. Additionally, the threshold voltage sensitivity for DNA detection is higher in the DM-DG-GaN-JAM-NWFET based sensor. Also, the ON-current sensitivity is enhanced in the DM-DG-GaN-JAM-NWFET. Hence, the analysis demonstrates that the DM-DG-GaN-JAM-NWFET biosensor is more effective for ultrasensitive biosensing applications.

REFERENCES

[1] B Yu, H. Lu, M. Liu and Y. Taur, "Explicit Continuous Model for Double Gate and Surrounding Gate MOSFET," IEEE Tran. On Electron Device, Vol.54, No.10, pp.2715-2722, Oct., 2007.

[2] P. Kumar, D. Joy and B.K. Jeblin, "Nanoscale tri-gate MOSFET for Ultra low power applications using high-k dielectrics,", 5th IEEE International Conference on Nanoelectronics (INEC), 2-4, pp.12-19, Jan., 2013

[3] D. Jimenez, J. J. Saenz, B. Iniguez, J. Sune, L. F.Marsal and J. Pallares, "Modeling of Nanoscale Gate-All-Around MOSFETs," IEEE, Electron device letters, Vol.25, No.5, pp.314-316, May, 2004.

[4] P. Kumar, M. Vashisht, N. Gupta, and R. Gupta, "Subthreshold Current Modeling of Stacked Dielectric Triple Material Cylindrical Gate All Around (SD-TM-CGAA) Junctionless MOSFET for Low Power Applications," Silicon, vol. 14, pp. 6261–6269, 2021.

[5] S. Yadav, S. Rewari and R. Pandey, "Sensitivity analysis of Dielectric Modulated Dual Gate Ferroelectric Junctionless Transistor based Biosensor," 2022 IEEE International Conference of Electron Devices Society Kolkata Chapter (EDKCON), Kolkata, India, 2022, pp. 270-273.

[6] Method and apparatus for controlling electric currents, by J. E. Lilienfeld (1926).

[7] R. Doria, M. Pavanello, C. Lee, I. Ferain, N. Dehdashti-Akhavan, R. Yan et al., "Analog Operation and Harmonic Distortion Temperature Dependence of nMOS Junctionless Transistors," ECS Transactions, vol. 31, 2010, pp. 13-20.

[8] N. Trivedi, M. Kumar, S. Haldar, S. Deswal, M. Gupta, R. Gupta, Analytical modeling of Junctionless Accumulation Mode Cylindrical Surrounding Gate MOSFET (JAM-CSG), International Journal of Numerical Modelling: Electronic Networks, Devices And Fields, vol. 29, 2016, pp.1036-1043.

[9] T. Kim, D. Kim, Y. Yoon, J. Moon, B. Hwang, D. Moon et al., "First Demonstration of Junctionless Accumulation-Mode Bulk FinFETs with Robust Junction Isolation," IEEE Electron Device Letters, vol. 34, 2013, pp. 1479-1481.

[10] S. Yadav, S. Rewari, and R. Pandey, "Gate Engineered Ferroelectric Junctionless BioFET for Label-Free Detection of Biomolecules," *Journal of Electronic Materials*, vol. 53, pp. 683–692, 2024.

[11] Anupama, S. Rewari, N. Pandey, "Numerical simulation and characterization of high-power gallium nitride based Junctionless Accumulation Mode Nanowire FET (GaN-JAM-NWFET) for small signal high frequency terahertz applications," AEU – International Journal of Electronics and Communications, vol.174, 2024.

[12] S. Yadav and S. Rewari, "Dual metal dual layer GAA NW–FET (DMDL–GAA–NW–FET) biosensor for label-free SARS-CoV-2 detection," *Microsystems Technology*, vol. 30, pp. 565–582, 2024.

[13] S. Sharma, V. Nath, S. S. Deswal, and R. S. Gupta, "Analytical modelling and sensitivity analysis of Gallium Nitride-Gate Material and dielectric engineered-Schottky nano-wire FET (GaN-GME-DE-SNW-FET) based label-free biosensor," Microelectronics Journal, vol. 129, 2022.

[14] A. Goel, S. Rewari, S. Verma and R. S. Gupta, "Dielectric Modulated Triple Metal Gate All Around MOSFET (TMGAA)for DNA Bio-Molecule Detection," 2018 IEEE Electron Devices Kolkata Conference (EDKCON), Kolkata, India, 2018, pp. 337-340, doi: 10.1109/EDKCON.2018.8770406.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:25:37 UTC from IEEE Xplore. Restrictions apply.

Journal of
Micromechanics and
Microengineering

PURPOSE-LED
PUBLISHING™

PAPER

Silicon slab waveguide based XOR and XNOR all optical logic gates

To cite this article: Yash Yadav and Yogita Kalra 2025 *J. Micromech. Microeng.* **35** 055002

View the article online for updates and enhancements.

You may also like

- A comprehensive analysis of ultra low power GNRFET based 20T hybrid full adder for computing applications
Sneha Arora and Suman Lata Tripathi
- Reconfigurable dynamic all-optical chaotic logic operations in an optically injected VCSEL
Dong-Zhou Zhong, , Ge-Liang Xu et al.
- All-optical XNOR gate based on 2D photonic-crystal ring resonators
Tamer A. Moniem



This content was downloaded from IP address 14.139.251.98 on 01/05/2025 at 08:03

IOP Publishing Journal of Micromechanics and Microengineering
J. Micromech. Microeng. **35** (2025) 055002 (7pp) https://doi.org/10.1088/1361-6439/adcc3e

Silicon slab waveguide based XOR and XNOR all optical logic gates

Yash Yadav and Yogita Kalra*

Advanced Simulation Laboratory, TIFAC-CORE in Fiber Optics and Optical Communication, Department of Applied Physics, Delhi Technological University, Delhi, India

E-mail: dryogitakalra@dtu.ac.in

Received 11 December 2024, revised 7 April 2025
Accepted for publication 14 April 2025
Published 22 April 2025

Abstract

This paper focuses on the design and implementation of silicon slab waveguides (SSWs) based all-optical XOR and XNOR logic gates. Utilizing silicon's high refractive index at optical communication wavelengths (1.55 μm), these gates exhibit strong optical confinement and efficient propagation in SSWs. The proposed design is optimized for both transverse electric and transverse magnetic modes. By strategically introducing air holes at key intersections of the waveguide, phase variations are induced to achieve the desired output logic for XOR and XNOR logic operations. These gates demonstrate femtosecond-scale propagation delays, significantly outperforming traditional logic gate designs, with low propagation losses and compact design.

Keywords: logic gates, silicon slab waveguide, XOR gate, XNOR gates, propagation losses

1. Introduction

Traditional electronic logic gates are approaching their physical and performance limits. Optical logic gates have emerged as a promising solution, offering higher speeds and lower power consumption. Recent advancements in optical logic gate implementation [1] have explored a diverse range of approaches, including semiconductor optical amplifiers (SOAs), four-wave mixing, microring resonators, two-photon absorption (TPA), spatial filtering, and photonic crystal structures [2–7]. Various optical gate designs have been explored, utilizing technologies such as SOAs, nonlinear fibers, and photonic crystals. Among these, silicon photonics has garnered significant attention for its potential to revolutionize data processing and signal transmission [8]. Silicon's high refractive index and compatibility with CMOS fabrication processes make it an ideal platform for integrating optical components into scalable and low-cost photonic circuits.

Silicon slab waveguides (SSWs), enhanced with photonic crystal structures, promise to transform data transmission, signal processing, and computing capabilities. Recent studies on SSW-based logic operations [1, 9–11] have primarily aimed at improving waveguide performance by minimizing propagation losses and optimizing dispersion characteristics, thereby facilitating functionalities such as wavelength conversion and optical amplification. Recent progress in silicon waveguide technology has substantially impacted optical communication and information processing, driven largely by developments in silicon photonics [12]. Among these, SSW-based logic gates have gained prominence due to their unique advantages [1, 13]. Silicon's high refractive index contrast enables strong light confinement and compact designs, making it ideal for miniaturized photonic circuits. This compact design is compatible with standard CMOS fabrication processes [14, 15], making silicon waveguides an appealing solution for scalable and These waveguides can also integrate nonlinear optical effects, enabling functionalities of all optical logic gates with high-speed operation and low energy requirements [16, 17].

The proposed XOR and XNOR gates serve as essential building blocks for universal logic circuits, with their implementation in an all-optical framework enabling interference between electromagnetic fields from two input waveguides at the waveguide junction. A significant advantage of this switching architecture is the high signal contrast between the on and off states at the output port, while also eliminating back-reflections into the input ports. Due to the instantaneous nature of the interference, the switching process within the

* Author to whom any correspondence should be addressed.

© 2025 IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

waveguide junctions is inherently ultrafast, making it well-suited for the development of low-power, high-speed optical computing systems.

The realization of all-optical SSW-based universal logic gate operations leverages linear interference effects in dielectric crossed waveguide structures, where photons serve as information carriers, significantly enhancing data bandwidth. This linear approach minimizes excessive heating within the structures, as only low electromagnetic intensities are required to drive the switching and logic elements [18]. The switching mechanism relies on the interference of a signal (Gaussian source) originating from two input ports and one control port at the waveguide junction. When the incoming signal and the control signal are in phase, constructive interference occurs, yielding a logic output of 1. Conversely, when these signals are out of phase, destructive interference takes place, resulting in a logic output of 0. Due to the low field intensities and the potential for a high degree of integration, linear interference effects present a promising approach, which will be explored further in this study [19, 20].

Further all the earlier reported works on logic gates operate on the non-linear phenomena such as four-wave mixing and TPA. Complex designs of logic gates based on resonant modes in microring cavities have been reported. Also, logic gates have also been designed on the photonic crystal platform operating on the principle of photonic band gap guidance. However, this paper focuses on the design and implementation of all-optical XOR and XNOR logic gates designed on SSW operating on linear interference effect in a simple dielectric crossed waveguide structure, which can be easily fabricated with the existing technology [21]. The novelty of the present work lies in its simple design relying on the linear interference phenomena with low power requirement. We leverage silicon's intrinsic properties for high-speed optical computing. This configuration enables tighter mode confinement, allowing the structure to support both transverse electric (TE) and transverse magnetic (TM) modes [22] with minimal propagation loss and high efficiency [23]. The finite-difference time-domain software, MIT electromagnetic equation propagation, is utilized for the simulation, optimization, and implementation of NAND and NOR logic gates for both polarizations. The simulation parameters include a spatial resolution with a step size of 20, a perfectly matched layer as the boundary condition, and an iteration unit of 200 [24].

2. Design of XOR and XNOR all optical logic gates

The design of (16 * 16) μm^2 2D device using *Y* shaped SSW on based XOR and XNOR optical logic gates is proposed leveraging the distinct propagation characteristics of TE and TM modes to achieve all-optical logic operations. At the core of the structure there is a *Y* shaped SSW [25, 26] SSW with two input ports (A and B) with an angle of 120° and one output port of dimension 7 μm long and width of 0.241 μm for TE mode and 0.345 μm for TM mode. A reference waveguide has been created which is indicated as reference R, used to create phase difference between input signals resulting into interfere either constructively or destructively for both polarizations, depending on the phase angle of the light beam launched into the waveguide. The phase angle of the input light beam can be varied, as demonstrated by Birr *et al* [27]. However, phase angle has to introduce manually in the reference waveguide. This waveguide functions as the primary medium for light propagation to realize XOR and XNOR logic operations [28].

To enhance the performance and mode control for TE and TM mode respectively, air holes are strategically placed within the structure as shown in figures 1(a) and (b). Figures 2(a) and (b) show the variation of transmission intensity vs. radius of the air hole for TE and TM mode respectively. For TE mode maximum transmission intensity is obtained at the radius of radius of 0.15 μm at center of waveguide. For TM mode the optimized radius of air hole is obtained as 0.05 μm placed at 0.5 μm below center of waveguide.

The waveguides are engineered for single-mode operation, with width d optimized at 0.241 μm for TE mode and 0.345 μm for TM mode, ensuring efficient mode confinement and minimal propagation loss. The single-mode condition is governed by equation (1),

$$v_c = \frac{c}{2d\sqrt{n_1^2 - n_2^2}} \tag{1}$$

d is the width of the SSW, n_{Si} is the refractive index of silicon ($n_{\mathrm{Si}=3.47}$), n_{air} is the refractive index of air, and c is the speed of light in a vacuum.

The performance of SSW based XOR and XNOR logic gates is evaluated by realization of the truth table and calculation of power loss. Power loss, key to efficiency is calculated as

$$\mathrm{loss} = 10\log_{10}\left(\frac{P_{\mathrm{out}}}{P_{\mathrm{in}}}\right), \tag{2}$$

where P_{in} and P_{out} are input and output powers, respectively [29]. Higher output power indicates greater gate efficiency. The contrast ratio (CR) [30] is defined as,

$$\mathrm{CR\,(dB)} = 10\ln\left(\frac{P_{\mathrm{out}}|ON_{\mathrm{min}}}{P_{\mathrm{out}}|OFF_{\mathrm{max}}}\right) \tag{3}$$

where $P_{\mathrm{out}} \mid ON_{\mathrm{min}}$ and $P_{\mathrm{out}} \mid OFF_{\mathrm{max}}$ represent the minimum and maximum output powers for logic levels '1' and '0.' Modulation depth (MD) [31] is calculated using equation (4),

$$\mathrm{MD} = \left(\frac{T|ON_{\mathrm{max}} - T|OFF_{\mathrm{min}}}{T|ON_{\mathrm{max}}}\right) \times 100\% \tag{4}$$

where $T \mid ON_{\mathrm{max}}$ and $T \mid OFF_{\mathrm{min}}$ are the maximum and minimum transmission values for logic levels '1' and '0.'

Data rate as determined using the Nyquist equation is given by [32],

$$\mathrm{Data\ rate} = 2\mathrm{B}\log_2(M) \tag{5}$$

where $M = 4$ which indicates the total number of signal levels (00, 01, 10, and 11), while B (optical bandwidth) is given by equation (6),

$$B = \left(\frac{c}{\lambda^2}\right)\Delta\lambda. \tag{6}$$

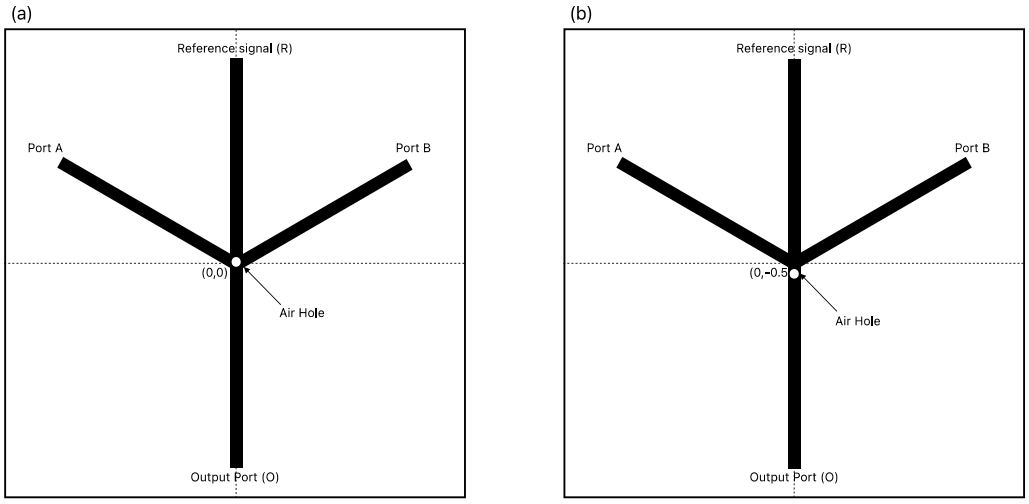


Figure 1. The schematic structure of the XOR and XNOR logic gates created using silicon slab waveguides (SSWs) (a) for TE mode, (b) for TM mode of propagation.

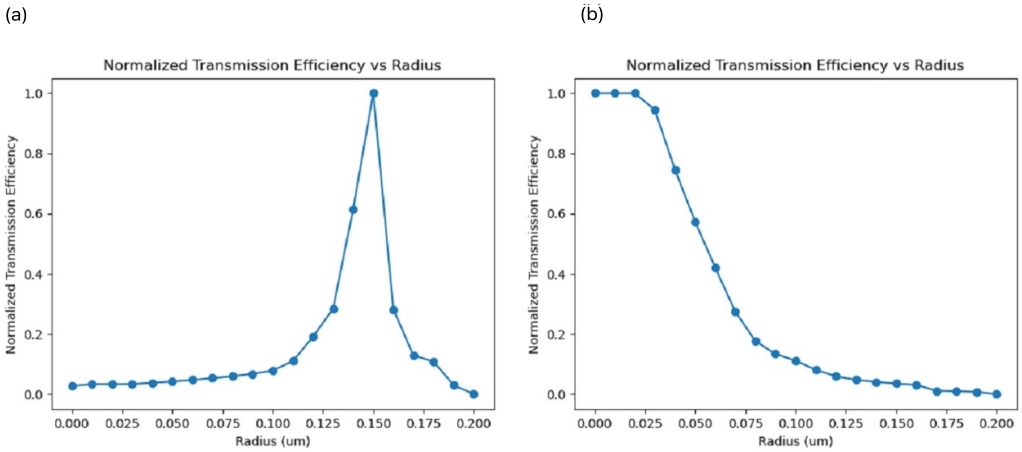


Figure 2. Variation of transmission intensity with radius of air hole (a) for TE mode, (b) TM mode.

with $\lambda = 1.55$ μm and $\Delta\lambda = 10$ μm, $B = 1.25$ THz yielding a data rate of 5 Tb s^{-1}.

The output-to-input power ratio determines the logic states: values above 0.5 indicate logic '1' and below 0.5 indicate logic '0'. This precise differentiation between logic states ensures accurate signal detection.

3. Realization of all optical XOR and XNOR logic gate

3.1. XOR logic gate

The intensity distribution for different input combinations is depicted in figure 3 for TE mode and TM modes. This figures provide a visual representation of how the optical power is distributed within the waveguide for each input state. For both TE mode and TM mode, the intensity is concentrated within the central waveguide when the output is '1', indicating efficient transmission. For the input combination (1, 1), the intensity drops significantly, showing that the XOR gate is successfully suppressing the output signal.

Table 1 shows the truth table of the realized all optical XOR logic gate, where output reflects the expected logic behavior. Table 1 also shows the calculated losses for both TE and TM modes. These losses are indicative of how efficiently the gate operates in both modes, which are crucial for determining its application in photonic circuits.

3.2. XNOR logic gate

The intensity distribution for the XNOR gate under various input conditions is shown in figure 4 for TE mode and TM mode respectively. For both TE mode and TM mode the intensity is concentrated within the central waveguide for the (0, 0) input, indicating strong signal transmission for the logic '1' output. For input configurations that yield a '0' output (e.g., 0, 1 or 1, 0), the intensity reduces significantly.

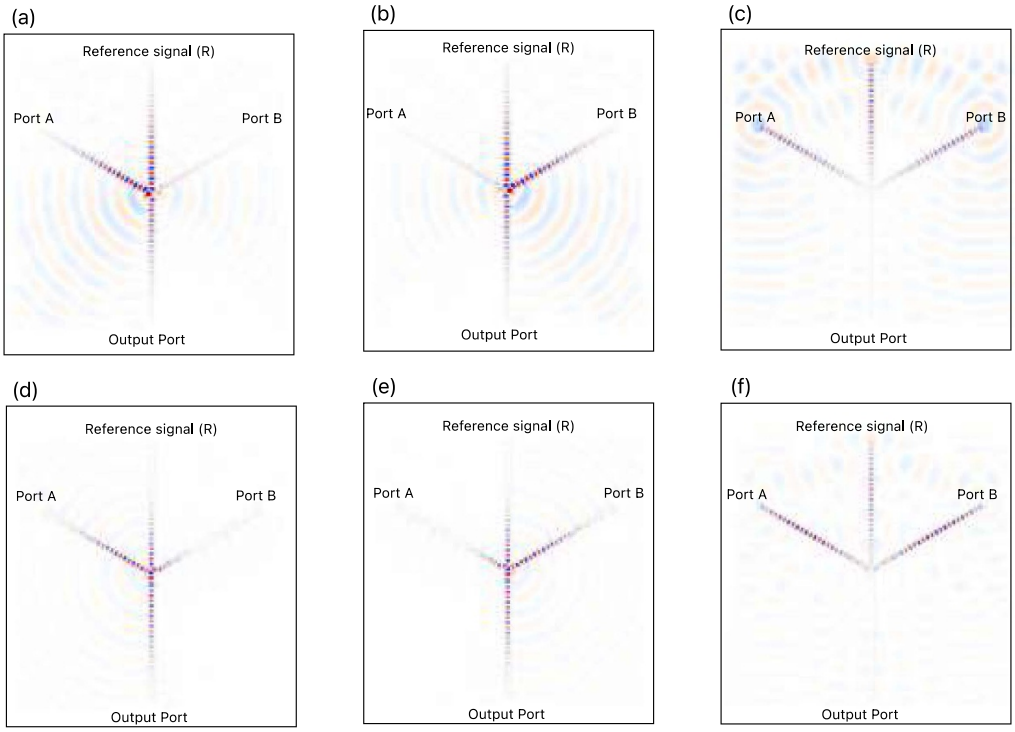


Figure 3. The intensity distributions of the XOR logic gate for TE mode of propagation (a) $A = 1$, $B = 0$, $R = 1$ ($\phi = 0°$), $O = 1$, (b) $A = 0$, $B = 1$, $R = 1$ ($\phi = 0°$), $O = 1$, (c) $A = 1$, $B = 1$, $R = 1$ ($\phi = 180°$), $O = 0$ and for TM mode of propagation (d) $A = 1$, $B = 0$, $R = 1$ ($\phi = 0°$), $O = 1$, (e) $A = 0$, $B = 1$, $R = 1$ ($\phi = 0°$), $O = 1$, (f) $A = 1$, $B = 1$, $R = 1$ ($\phi = 180°$), $O = 0$.

Table 1. Truth table and losses for XOR logic gate in transverse electric (TE) and transverse magnetic (TM) modes of propagation.

Logical operations				TE mode		TM mode	
Logical input at port A	Logical input at port B	Logical output	Logical input at reference signal (R) (ϕ = phase difference)	Normalized output power	Loss (dB)	Normalized output power	Loss (dB)
0	0	0	0	0	0	0	0
0	1	1	1 ($\phi = 0°$)	0.8430 P_{in}	0.7416	0.9965 P_{in}	0.0154
1	0	1	1 ($\phi = 0°$)	0.8529 P_{in}	0.6913	0.9965 P_{in}	0.0154
1	1	0	1 ($\phi = 180°$)	0.2609 P_{in}	5.8346	0.1843 P_{in}	7.3432

Table 2 summarizes the truth table along with losses for the realize XNOR gate for both polarizations

4. Analysis of XOR and XNOR optical logic gate

Various parameters such as maximum transmission, CR and MD for the designed XOR and XNOR gates have been calculated and summarized in table 3.

Figure 5(a), illustrates the time evolution curve of the proposed structure for TE mode, showing that the time taken for the output power to increase from 0%–90% of the average steady-state output power, P_{av}, is $ct = 0.21$ μm or $t_1 = 0.419$ fs. The response time consists of two components: the transmission delay ($t_{11} = 0.33$ fs) and the time required for the output power to increase from 0% to 90% ($t_{12} = 0.24$ fs). Given that the system operates on linear materials, the fall time from P_{av} to 10% of P_{av} can approximately be taken as $t_{12} = 0.24$ fs. Consequently, the logic gate can produce a narrow pulse with a width of $2t_{12} = 0.48$ fs, then the response time is 0.81 fs. Similarly, figure 5(b) depicts the time evolution curve of the proposed structure for TM mode where $t_1 = 0.70$ fs, $t_{11} = 0.37$ fs, $t_{12} = 0.33$ fs, $2t_{12} = 0.66$ fs leading to the response time of 1.03 fs.

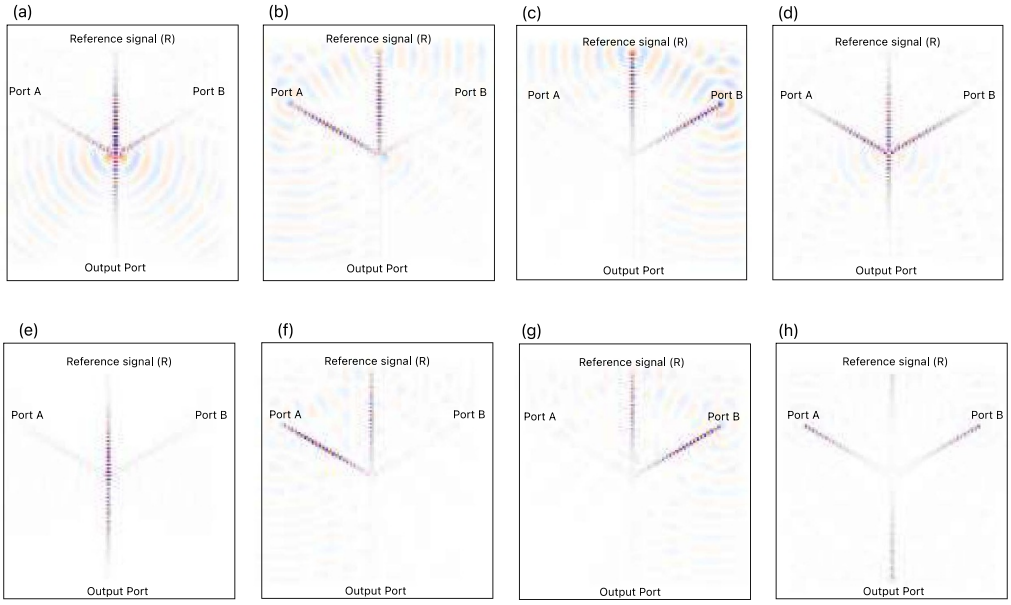


Figure 4. The intensity distributions of the XNOR logic gate for the TE mode of propagation (a) $A = 0$, $B = 0$, $R = 1$ ($\phi = 0^\circ$), $O = 1$, (b) $A = 1$, $B = 0$, $R = 1$ ($\phi = 180^\circ$), $O = 0$, (c) $A = 0$, $B = 1$, $R = 1$ ($\phi = 180^\circ$), $O = 0$, (d) $A = 1$, $B = 1$, $R = 1$ ($\phi = 0^\circ$), $O = 1$and for TM mode of propagation (e) $A = 0$, $B = 0$, $R = 1$($\phi = 0^\circ$), $O = 1$, (f) $A = 1$, $B = 0$, $R = 1$ ($\phi = 180^\circ$), $O = 0$, (g) $A = 0$, $B = 1$, $R = 1$ ($\phi = 180^\circ$), $O = 0$, (h) $A = 1$, $B = 1$, $R = 1$ ($\phi = 0^\circ$), $O = 1$.

Table 2. Truth table and losses for XNOR logic gate in transverse electric (TE) and transverse magnetic (TM) modes of propagation.

Logical operations				TE mode		TM mode	
Logical input at port A	Logical input at port B	Logical output	Logical input at reference signal (R) (ϕ = Phase difference)	Normalize output power	Loss (dB)	Normalized output power	Loss (dB)
0	0	1	1 (ϕ= 0°)	0.9026 P_{in}	0.4452	0.5574 P_{in}	2.5386
0	1	0	1 ($\phi = 180^\circ$)	0.1501 P_{in}	8.2369	0.1204 P_{in}	9.1950
1	0	0	1 ($\phi = 180^\circ$)	0.1518 P_{in}	8.1862	0.1204 P_{in}	9.1950
1	1	1	1 ($\phi = 0^\circ$)	0.9831 P_{in}	0.0740	0.9354 P_{in}	0.2901

Table 3. Number of all-optical logic gates, type of materials used, maximum transmission, threshold transmission, maximum contrast ratio, maximum modulation depth.

Logic gate	Polarization	Maximum transmission	Threshold transmission	Maximum Contrast ratio	Maximum modulation depth
XOR gate	TE mode	85.29%	0.50	17.83 dB	69.41%
XNOR gate		98.31%			
XOR gate	TM mode	99.65%	0.50	11.73 dB	81.50%
XNOR gate		93.54%			

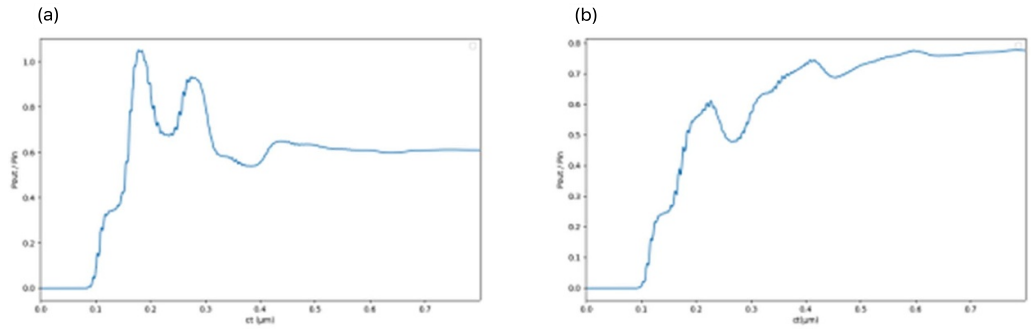


Figure 5. Time evolution curve of the output power for all optical XOR and XNOR logic gate design for (a) TE and (b) TM mode of propagation.

Table 4. Comparison of proposed work with the previous existing methods.

Design and methods	Logic gates	Wavelength (μm)	Contrast ratio (db)	Response time	References
Optical interferometric metal slot waveguide	NOT, XOR, AND, OR, NOR, NAND, XNOR	0.6328	6–16	NA	[34]
Y shaped Photonic crystal waveguides	AND, XOR, OR, NOT, NAND, NOR, XNOR	1.55	5.42–9.59	1.024 ps	[5]
Ultra-compact photonic gates	AND, OR, NOT, NAND	1.3	0.5–5.79	NA	[35]
1D MIM structures	OR, XOR, NOT, AND, XNOR, NAND, NOR	1.55	15	NA	[36]
Ψ-shaped for linear optical logic gates	AND	1.55	6	NA	[9]
T-shaped defect	NAND, NOR, XNOR, OR	1.55	8.66	NA	[37]
Z-shaped silicon waveguides	XOR, AND, OR, NOT, NOR, XNOR, NAND	1.55	29–33.26	NA	[17]
Dielectric-metal-dielectric design	OR, XOR, NOT, AND, XNOR, NAND, NOR	0.9–1.33	5.37–22	NA	[38]
SSW based universal logic gates	XOR, XNOR	1.55	17.83–11.73	0.81–1.03 fs	This work

The proposed all-optical logic gate design operates at very low power levels, making it a promising candidate for the development of all-optical integrated circuits [33]. A tolerance study, crucial from a fabrication perspective, was conducted by varying the air hole radius by $\pm 5\%$. The results indicate that this variation leads to an approximately 3% change in transmitted intensity for the TE mode and $\sim$0.8% for the TM mode.

The proposed all-optical SSW-based universal logic gates are evaluated against previous works based on key performance criteria, including the number of logic functions achieved, operating wavelength, CR, and response time, as summarized in table 4. The data in this table demonstrates that the proposed gates successfully perform the required logic operations while achieving a significantly higher CR and operating at the commercially viable wavelength of 1550 nm.

5. Conclusion

This paper demonstrated the design of all-optical XOR and XNOR logic gates utilizing SSWs for high-speed, low-power optical computing. These gates are optimized for fundamental logic operations, with a fan-in of 2 and fan-out of 1, showcasing their ability to perform essential computing functions. The gates maintain minimal propagation losses while ensuring precise logic operations, making them highly suitable for integration into larger photonic computing systems. The ability to control and leverage both TE and TM modes provides additional flexibility for optimizing performance across a range of computing applications.

Data availability statement

No new data were created or analysed in this study.

Acknowledgment

Authors gratefully acknowledge the initiatives and support from the TIFAC-Centre of Relevance and Excellence in Fiber Optics and Optical Communication at Delhi College of Engineering, now DTU, Delhi, under the Mission REACH program of Technology Vision-2020, Government of India.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ORCID iD

Yogita Kalra https://orcid.org/0000-0003-1589-4628

References

[1] Kotb A, Zoiros K E and Li W 2023 Silicon-on-silica waveguides-based all-optical logic gates at 1.55 μm *Phys. Scr.* **98** 035517

[2] Kotb A, Zoiros K E and Li W 2022 320 Gb/s all-optical logic operations based on two-photon absorption in carrier reservoir semiconductor optical amplifiers *Optik* **265** 169494

[3] Hou J, Chen L, Dong W and Zhang X 2016 40 Gb/s reconfigurable optical logic gates based on FWM in silicon waveguide *Opt. Express* **24** 2701

[4] Wu W, Sun Q, Wang G, Zhang L and Zhao W 2021 Tunable all-optical AND logic gates via four-wave mixing based on graphene-on-silicon slot microring resonators *Opt. Laser Technol.* **138** 106926

[5] Rani P, Kalra Y and Sinha R K 2015 Design of all optical logic gates in photonic crystal waveguides *Optik* **126** 950–5

[6] Li Z, Liu Y, Zhang S, Ju H, De Waardt H, Khoe G D, Dorren H J S and Lenstra D 2005 All-optical logic gates using semiconductor optical amplifier assisted by optical filter *Electron. Lett.* **41** 51–52

[7] Xu Q, Lipson M, Almeida V R, Barrios C A, Panepucci R R and Lipson M 2007 All-optical logic based on silicon micro-ring resonators *Opt. Express* **15** 924

[8] Zhang H, Wang B and Wang S 2025 A novel method of high-speed all-optical logic gate based on metalens *Opt. Commun.* **574** 131132

[9] Kita S, Nozaki K, Takata K, Shinya A and Notomi M 2020 Ultrashort low-loss Ψ gates for linear optical logic on Si photonics platform *Commun. Phys.* **3** 33

[10] Wang J and Long Y 2018 On-chip silicon photonic signaling and processing: a review *Sci. Bull.* **63** 1267–310

[11] Aggarwal A, Mittal A and Kalra Y 2024 Design of silicon slab waveguides based all-optical logic gates *Microw. Opt. Technol. Lett.* **66** e33981

[12] Lan Z, Chen M L N, Gao F, Zhang S and Sha W E I 2022 A brief review of topological photonics in one, two, and three dimensions *Rev. Phys.* **9** 100076

[13] Merlo J M, Wu X, Kempa K and Naughton M J 2021 All-optical logic gates based on anomalous Floquet photonic topological insulator structures *J. Opt.* **23** 065001

[14] Tsybeskov L, Lockwood D J and Ichikawa M 2009 Silicon photonics: CMOS going optical [Scanning the Issue] *Proc. IEEE* **97** 1161–5

[15] Moon S J and Lee S S 2005 A novel fabrication method of a microneedle array using inclined deep x-ray exposure *J. Micromech. Microeng.* **15** 903–11

[16] Jindal P, Abou Houran M, Goyal D and Choudhary A 2023 A review of different techniques used to design photonic crystal-based logic gates *Optik* **280** 170794

[17] Kotb A, Zoiros K E, Hatziefremidis A and Guo C 2023 Optical logic gates based on Z-shaped silicon waveguides at 1.55 μm *Micromachines* **14** 1266

[18] Qian L and Caulfield H J 2006 What can we do with a linear optical logic gate? *Inf. Sci.* **176** 3379–92

[19] Caulfield H 2004 The logic of optics and the optics of logic *Inf. Sci.* **162** 21–33

[20] Zhang Y, Zhang Y and Li B 2007 Optical switches and logic gates based on self-collimated beams in two-dimensional photonic crystals *Opt. Express* **15** 9287

[21] Korkmaz S 2024 Investigation of metal-insulator-metal waveguides with elliptical-nanodisk resonator for high contrast ratio all-optical logic gates *J. Opt.* **26** 115005

[22] Xu K 2021 Silicon electro-optic micro-modulator fabricated in standard CMOS technology as components for all silicon monolithic integrated optoelectronic systems * *J. Micromech. Microeng.* **31** 054001

[23] Vivien L and Pavesi L 2016 *Handbook of Silicon Photonics* ed L Vivien and L Pavesi (CRC Press)

[24] Oskooi A F, Roundy D, Ibanescu M, Bermel P, Joannopoulos J D and Johnson S G 2010 Meep: a flexible free-software package for electromagnetic simulations by the FDTD method *Comput. Phys. Commun.* **181** 687–702

[25] Rani P, Kalra Y and Sinha R K 2016 Design and analysis of polarization independent all-optical logic gates in silicon-on-insulator photonic crystal *Opt. Commun.* **374** 148–55

[26] Rani P, Kalra Y and Sinha R K 2013 Realization of AND gate in Y shaped photonic crystal waveguide *Opt. Commun.* **298–299** 227–31

[27] Birr T, Zywietz U, Chhantyal P, Chichkov B N and Reinhardt C 2015 Ultrafast surface plasmon-polariton logic gates and half-adder *Opt. Express* **23** 31755

[28] Korkmaz S 2024 Realization of all-optical logic gates using MIM waveguides and a rectangular ring resonator *Microelectron. Eng.* **294** 112259

[29] Anagha E G and Jeyachitra R K 2022 Review on all-optical logic gates: design techniques and classifications—heading toward high-speed optical integrated circuits *Opt. Eng.* **61** 060902

[30] Tresna W P and Maruyama T 2019 Propagation loss on a Si-Slab Waveguide: simulation revisited *J. Phys. Appl.* **2** 76

[31] Al-Sabea Z S, Ibrahim A A and Abdulnabi S H 2022 Plasmonic logic gates at optimum optical communications wavelength *Adv. Electromagnetics* **11** 10–21

[32] Hossain M, Rakshit J K, Bhatnagar A and Chattopadhyay T 2023 Silicon microring resonator based all-optical 3-input majority gate and its applications *Optik* **282** 170891

[33] Wu C J, Liu C P and Ouyang Z 2012 Compact and low-power optical logic NOT gate based on photonic crystal waveguides without optical amplifiers and nonlinear materials *Appl. Opt.* **51** 680

[34] Pan D, Wei H and Xu H 2013 Optical interferometric logic gates based on metal slot waveguide network realizing whole fundamental logic operations *Opt. Express* **21** 9556

[35] Neşeli B, Yilmaz Y A, Kurt H and Turduev M 2022 Inverse design of ultra-compact photonic gates for all-optical logic operations *J. Appl. Phys.* **55** 215107

[36] Bian Y and Gong Q 2014 Compact all-optical interferometric logic gates based on one-dimensional metal-insulator-metal structures *Opt. Commun.* **313** 27–35

[37] Priya N H, Swarnakar S, Krishna S V and Kumar S 2021 Design and analysis of a photonic crystal-based all-optical 3-input OR gate for high-speed optical processing *Opt. Quantum Electron.* **53** 720

[38] Al-Musawi H K, Al-Janabi A K, Al-abassi S A W, Abusiba N A H A and Al-Fatlawi N A H Q 2020 Plasmonic logic gates based on dielectric-metal-dielectric design with two optical communication bands *Optik* **223** 165416

SobelMNet: A Lightweight CNN based Approach for Median Filtering Detection in compressed images

Anukriti Kaushal[1], Rajeev Kumar[1], Sanjay Kumar[1]
[1]*Dept. of Computer Science & Engineering, Delhi Technological University, Delhi, India*
anukritikaushal@dtu.ac.in, rajeevkumar@dtu.ac.in, sanjay.kumar@dtu.ac.in

***Abstract*—Digital image forgery is the process of manipulating an image to deceive or mislead observers with real or manipulated content. Median filtering is widely used to smooth images and obscure traces of tampering, making its detection critical for image forensics. However, identifying median filtering becomes more complex when additional operations, such as compression, resampling, or noise addition, are applied. To address this issue, we propose a lightweight convolutional neural network (CNN) model named SobelMNet, specifically designed for detecting median filtering in compressed images. The proposed model utilises a Sobel filter-based preprocessing step to enhance the residual differences between the original and manipulated images. These residuals, which capture subtle features indicative of median filtering, are analysed by CNN for classification. Further, the proposed model is evaluated on grayscale low-resolution images generated from the Dresden dataset for both binary and multiclass classification tasks. The model achieved a remarkable detection accuracy of 99.43% in median filter detection and outperformed state-of-the-art methods in various scenarios, including combinations of median filtering with Gaussian blur, resampling, and additive white Gaussian noise (AWGN) with an average accuracy of 98.32%. Finally, its lightweight architecture ensures computational efficiency, making it practical for real-world forensic applications.**

***Index Terms*—Median Filter, CNN, Sobel Filter, Operation Chain**

I. Introduction

Digital images play a vital role in modern communication, entertainment, and archiving, spanning areas such as social media, academic visuals, and medical imaging [1]. However, with the rise of advanced editing tools, image manipulation has become widespread and increasingly complex [2]. This has significant implications, such as spreading misinformation and forging legal documents, highlighting the need for reliable forensic methods to verify image authenticity and detect alterations [3].

Image manipulation ranges from simple contrast adjustments to complex filtering techniques. Among these, the median filter is widely used due to its ability to reduce noise while preserving edges [4]. This nonlinear filter replaces each pixel's value with the median of surrounding pixels, making it effective for noise removal [5]. While helpful for preprocessing, the median filter can also obscure evidence of tampering, making its detection critical in image forensics.

Detecting median filtering is essential for tracing an image's processing history and identifying potential alterations [6]. Researchers have explored various methods, from statistical feature extraction to advanced deep learning models [7], to uncover artefacts specific to median filtering. [8].

Deep learning has advanced image forensics, enabling the detection of subtle changes that are often invisible to the human eye. Models such as support vector machine (SVM), CNNs and generative adversarial networks (GANs) have been adapted to identify processing artefacts, including those from median filtering. These models are highly effective due to their ability to extract detailed features and complex analysis in forensic classification models. However, challenges remain, particularly when dealing with low-resolution or heavily compressed images [9]–[11].

II. Realted Work

Earlier statistical methods were utilised for image manipulation detection. Zhang et al. [12] introduce a statistical method for detecting median filtering manipulations in digital images using a second-order Local Ternary Pattern (LTP). The LTP captures texture changes caused by median filtering by encoding local derivative direction variations. Kernel Principal Component Analysis is applied to reduce feature dimensionality. Some commonly used statistical methods for detecting median filtering involve histogram residuals, which effectively capture distinctive patterns in images. These patterns, resulting from the median filter operation, allow for the identification of filtered images by analysing variations in the histogram's distribution [13], [14]. Gao and Gao [15] find differences in the pixel values of the image in four directions using pixel-pair histograms and autoregressive moving average coefficients, which outperform compressed JPEG images and low-resolution images. Similarly, Gao et al. [16] compute the first-order and second-order difference images using joint conditional probability density functions. Additionally, principal component analysis is employed for dimensionality reduction, making the model efficient for filter detection in JPEG-compressed and low-resolution images. Kang et al. [17], [18] give an autoregressive model that uses the statistical property of median filter for detection and localisation median filter in an image. The above models outperform work for image forensics but compromise if there is a chain of operations.

The introduction of machine learning and deep learning models has significantly enhanced the ability to detect com-

ISBN: 979-11-88428-13-7 Copyright 2025, Global IT Research Institute (GIRI). All right reserved.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 05:52:39 UTC from IEEE Xplore. Restrictions apply.

plex operations applied to images. Jin et al. [19] present a median filtering detection method for RGB images using dark channel residual extraction, multi-scale fusion, and GANs. The dark channel residuals reduce content interference, while multi-scale fusion captures traces from different filter sizes. GANs improve robustness and replace SVM classification. The method outperforms existing techniques, especially with small blocks and JPEG compression, and aims to extend to grayscale images. A CNN-based method introduces MFNet [20] for detecting median filtering in small or compressed images. It uses nearest neighbour interpolation to up-sample images, enhancing the difference between original and filtered versions. MFNet incorporates two mlpconv layers to improve nonlinear classification, followed by three convolutional layers for feature extraction. The method shows significant improvements, especially for images as small as 16×16, and offers potential for broader application in digital image forensics. Ahmed and Islam [21] discuss an approach that leverages the streaking effect caused by median filtering, using SVMs with 10-fold cross-validation to distinguish between median-filtered and unfiltered images. This method also performs well on compressed and low-resolution images. Another study [22] combines texture features extracted via GLCM with streak area analysis, employing an 18-dimensional feature set processed through neural networks to identify median filtering under various manipulations like JPEG compression and Gaussian blur. An advanced CNN-based framework introduces the MFFNet [23], integrating high-pass filters and Squeeze-and-Excitation modules, excelling in detecting manipulations in compressed, low-resolution images. Despite advancements, limitations include challenges in handling operator chains, which are multiple operations on images, including median filtering, and scaling methods to large datasets. These methods demonstrate a shift from traditional SVMs to robust deep learning architectures for feature learning and classification. A deep learning-based approach is introduced for detecting median filtering in the DCT domain. The method employs an Adaptive Filtering Layer to focus on relevant frequency ranges and a multi-scale convolutional block with a maxout function for effective feature extraction. The approach shows superior performance over existing methods, particularly for low-resolution and compressed images [24].

A limitation in existing methods is the decrease in accuracy when detecting median filtering, especially when it is followed by operations such as Gaussian blur, downsampling, or noise addition. Additionally, some techniques are highly dependent on high-resolution images, making them less effective for low-resolution or resized images.

III. Proposed Work: CNN-Based Model

Proposed SobelMNet, a two-dimensional residual-based 2D CNN, specifically designed for detecting traces of median filter in digital images. The key features of the proposed approach are as follows:

1) The Sobel filter-based residual is introduced as the preprocessing layer in the model. This filter kernel modifies the image pixels and generates residuals that amplify the classifier's ability to distinguish between different classes.
2) After the Sobel residual processing, we extract features that emphasise the statistical and structural differences in the image. These features are designed to make it easier for the neural network to identify whether an image has been median-filtered.
3) A lightweight neural network architecture is used for analyzing the extracted features and classifying the images. The network is efficient and can handle both binary tasks and multiclass classification.
4) The model is evaluated on images with low resolution and compression artifacts to ensure it works effectively even under challenging conditions.
5) To further validate the model's robustness, we tested it on images processed with median filtering followed by other operations like Gaussian blur, downsampling, additive white Gaussian noise (AWGN), and compression. This shows that the model can handle more complex real-world scenarios.

The architecture consists of three primary blocks designed to handle image feature extraction and classification efficiently: Preprocessing block, feature extraction block, and classification block as discussed in subsections.

A. Preprocessing Layer

This block applies a Sobel filter to enhance the edges in the image. The Sobel filter helps detect subtle changes introduced by median filtering, which tends to blur edges. By emphasizing these changes, it makes the subsequent feature extraction phase more effective. The input is a grayscale image, typically of size 64×64 pixels.

The Sobel filter [25] is a well-known edge-detection operator commonly used in image processing. It is especially effective for highlighting changes in intensity or gradients within an image. The filter works by using two convolution kernels: one to measure the gradient in the horizontal direction (Sobel-X) and another for the vertical direction (Sobel-Y). The gradient magnitude is then computed to identify edges, which represent significant transitions in pixel intensity. Let the input image be denoted as I, with dimensions $H \times W$.

Step 1: Sobel Filter Kernels The Sobel filters for the x- and y-directions are defined as:

$$K_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad K_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}.$$

Step 2: Convolution Operation Convolve the input image I with K_x and K_y to compute the gradients in the horizontal and vertical directions, respectively:

$$G_x = I * K_x, \quad G_y = I * K_y,$$

where $*$ denotes the convolution operation, and G_x, G_y are of size $H \times W$.

ISBN: 979-11-88428-13-7
Copyright 2025. Global IT Research Institute (GIRI). All right reserved.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 05:52:39 UTC from IEEE Xplore. Restrictions apply.

Step 3: Sobel Residual Calculation Combine the gradients using the Euclidean norm to compute the residual:

$$R_{\text{Sobel}}(i,j) = \sqrt{G_x(i,j)^2 + G_y(i,j)^2},$$

where i, j denote the pixel coordinates. The output R_{Sobel} is an edge-enhanced image of the same dimensions as the input ($H \times W$).

The Sobel filter can reveal subtle changes caused by enhanced edges and highlight residuals that are left behind by median filtering, such as the streaking effect left by median filtering. When applied to noisy images, it makes these residuals easier to observe. This enhanced information is used in machine learning models or neural networks to help them more accurately identify manipulated images, especially when distinguishing between original and filtered versions.

B. Feature Extraction Layer

The architecture model, shown in Figure 1, consists of four convolutional layers designed to progressively extract hierarchical features. Each block includes a convolutional layer followed by batch normalization and a ReLU activation function. A kernel of size 3×3 is applied in each convolutional layer. In the figure, FM represents feature maps, with the dimensions provided at each stage. Max-pooling layers are incorporated between convolutional layers to reduce spatial dimensions while preserving essential features, and dropout layers are included to minimise the risk of overfitting.

C. Classification Block

The features obtained from the previous block are converted into a one-dimensional vector and then passed through fully connected layers for advanced decision-making. The model is designed to classify multiple types of operations applied to an image, including a single median filter or a median filter followed by operations such as Gaussian blur, resampling, and AWGN for compressed and low resolution image datasets. The model classifies images by extracting hierarchical features through convolutional layers and fully connected layers, The final layer generates output class probabilities, enabling binary or multi-class classification depending on the number of classes in a training dataset. During training, the Adam optimiser adjusts the model's parameters, while CrossEntropyLoss is used to compute the loss.

IV. Experiment and Result

A. Dataset Description

To evaluate the performance of the proposed model in detecting image manipulations, we created a custom image dataset based on the high-quality Dresden dataset [24]. A total of 10,000 images were cropped to 64x64 pixels and converted to grayscale to standardize the input data. Pre-compression and post-compression were applied with a quality factor (QF) of 80, before and after performing single or multiple operations. This process simulates the degradation typically caused by compression algorithms. After the initial compression, median filtering with 3x3 and 5x5 kernel sizes was applied to simulate manipulation. The dataset for median filtering includes an equal distribution of 50% images processed with a 3x3 kernel and 50% with a 5x5 kernel. As shown in Table I, the images were compressed again to replicate multiple stages of image processing. The final dataset includes 30,000 training images, 6,000 validation images, and 10,000 test images for each class. This diverse set of images allows for robust model training and evaluation. The approach ensures that the model can detect manipulations effectively across various image quality levels and processing stages.

For binary classification, the following two classes were defined: The sequence of operations for an image (I) undergoing compression with QF 80, followed by median filtering (MF) with kernel sizes of 3x3 or 5x5, and then compression again with QF 80 can be represented as:

$$\text{I} \xrightarrow{\text{Compression}} \xrightarrow{\text{Median Filter}} \xrightarrow{\text{Compression}} \text{MF(I)}$$

In multi-class classification, the sequence of operations the image I undergoes is pre-compression with QF 80, then an initial median filtering, followed by another operation O as given in table I, and finally, post-compression with QF 80. This can be expressed as:

$$I \xrightarrow{\text{Compression}} \xrightarrow{\text{Median Filter}} \xrightarrow{O} \xrightarrow{\text{Compression}} O(\text{MF}(I))$$

TABLE I: Image manipulation operations and parameters applied to the dataset

Operation	Parameters
Median Filtering (MF)	Window sizes: size=3, size=5
Gaussian Blurring (GB)	Kernel sizes: (3, 3), (5, 5)
Resampling (RS)	Scaling factors: 1.2, 1.5
Additive Gaussian Noise (AWGN)	Noise variance: var=0.5^2, var=2^2
JPEG Compression	Quality factors: QF=80%

B. Performance Evaluation of Median Filter Detection Model

In this subsection, we discuss the results obtained from the model's training and evaluation. The training and validation performances are presented, followed by a detailed discussion of the confusion matrix and the model's ability to classify manipulated and original images.

1) Single Manipulation Detection: For binary classification, the model is trained, and the best model is saved, with the overall training accuracy of 96.83% and validation accuracy of 98.89%. The validation accuracy of the class median filter is 99.43% and the original accuracy is 97.75%. These high accuracies suggest that the model can effectively distinguish between manipulated and original images in the dataset.

We compare True positive rate (TPR) of SobelMNet with other existing methods designed for median filtering detection under similar parameters. Table II presents the result of the proposed method in comparison with MedFCN [22] and Streak analysis [21]. The accuracy of the median filter detection approach has been demonstrated to be competitive with state-of-the-art techniques, but further improvements in distinguishing original images are still required.

ISBN: 979-11-88428-13-7
Copyright 2025. Global IT Research Institute (GIRI). All right reserved.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 05:52:39 UTC from IEEE Xplore. Restrictions apply.

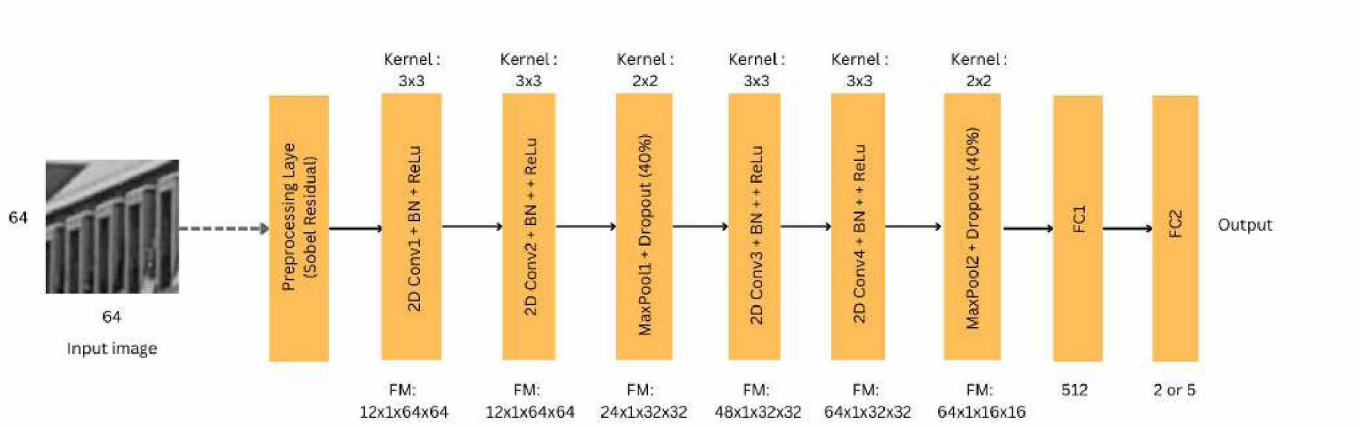


Figure 1: SobelMNet architecture

TABLE II: Comparison of SobelMNet with Different Methods

Method	True Positive Rate (TPR)
Proposed Method	99.43%
MedFCN [22]	98.50%
Streak [21]	97.80%

2) Multiple Manipulation Detection: This section presents the evaluation of the proposed model's performance on the median filter chain dataset. The results include the validation accuracy, class-specific performance, and the confusion matrix. Additionally, the trends in training and validation accuracy and loss are shown in Figure 3.

The model achieved a best validation accuracy of 98.32%, showcasing its effectiveness in identifying manipulated and original images. Table III presents the confusion matrix in percentage, providing a comparative representation of the model's classification performance across five classes. Class-specific accuracies demonstrate the robustness of the model across different image manipulations.

TABLE III: Confusion Matrix for Median Filter Chain Classification with True and Predicted Class labels.

True Class	Predicted Class				
	MF	**MF+AWGN**	**MF+GB**	**MF+RS**	**ORG**
MF	88.33	0.00	5.18	0.06	0.55
MF+AWGN	0.00	100.00	0.00	0.00	0.00
MF+GB	0.25	0.00	93.97	0.06	0.04
MF+RS	0.03	0.00	0.63	93.72	0.00
ORG	1.24	0.00	0.42	0.01	99.41

Figure 3 illustrated the trends in training and validation accuracy and loss. The training accuracy and validation accuracy consistently improved over epochs, demonstrating effective learning and generalization. The training loss decreased steadily, while the validation loss remained low, confirming that the model was not overfitting to the training data. The proposed model's performance on the median filter chain dataset demonstrates its high accuracy and robustness in distinguishing various manipulated and original images. The results underscore the model's potential for real-world applications in image forensics.

C. Result and Analysis

The proposed model SobelMNet achieves 99.43% accuracy in detecting median filtering on low-resolution, compressed images. It outperforms state-of-the-art methods like MedFCN and streak analysis in True Positive Rate (TPR) given in table II. Additionally, it accurately identifies complex operations such as median filtering combined with Gaussian blur, resampling, or AWGN, achieving an average accuracy of 98.32% and the highest accuracy of 100% when the image contains a median filter and AWGN, as shown in Table III.

V. Conclusion and Future Work

The paper presented SobelMNet, a lightweight model that utilized a Sobel filter residual to detect differences between an original image and an image processed with a median filter. Designed with fewer parameters, the model was highly efficient and effectively tested whether an image had been manipulated with a median filter or other median filter chain operations. The results demonstrated that SobelMNet outperformed existing models, achieving an impressive accuracy of 99.43% on compressed and low-resolution images.Moreover, the model offered computational efficiency while maintaining high performance, making it suitable for real-world scenarios involving low-resolution images, such as those found in social media, legal evidence, and archival data. SobelMNet also demonstrated robust performance across diverse images using the Dresden dataset. In future work, this model could be extended and applied to detect other types of image manipulations. SobelMNET's future research can focus enhancing SobelMNet's generalization to detect image manipulations, deepfake generation, copy move forgery, and inpainting can

ISBN: 979-11-88428-13-7
Copyright 2025, Global IT Research Institute (GIRI). All right reserved.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 05:52:39 UTC from IEEE Xplore. Restrictions apply.

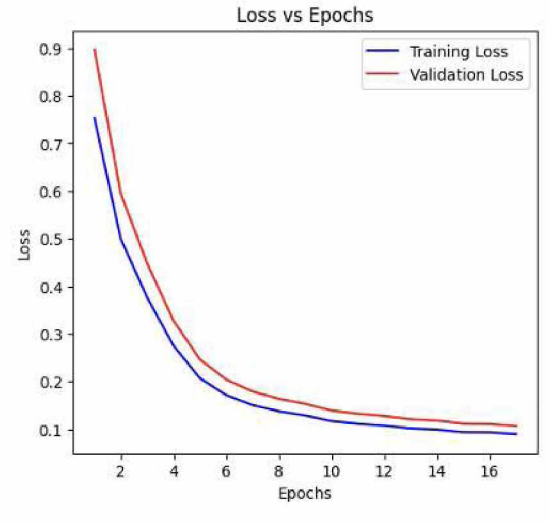


Figure 2: Loss plot for Training and Validation

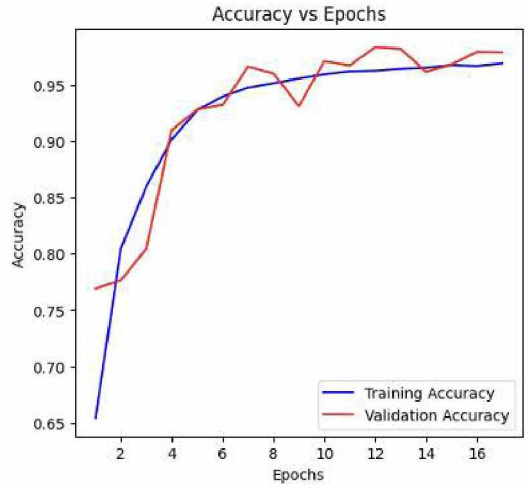


Figure 3: Accuracy plot for Training and Validation

Figure 4: Training and Validation Accuracy and Loss Curves for Median Filter Chain Detection

improve forensic scenarios. Evaluating SobelMNet on different datasets, including high definition and color images, can confirm its stability. Improving the SobelMNet's defense towards adversarial attacks, which exploit subtle distortions to mislead detection, is crucial.

REFERENCES

[1] Hany Farid. Exposing digital forgeries from jpeg ghosts. *IEEE transactions on information forensics and security*, 4(1):154–160, 2009.

[2] Alessandro Piva. An overview on image forensics. *International Scholarly Research Notices*, 2013(1):496701, 2013.

[3] Jessica Fridrich. Digital image forensics. *IEEE Signal Processing Magazine*, 26(2):26–37, 2009.

[4] Thomas Huang, GJTGY Yang, and Greory Tang. A fast two-dimensional median filtering algorithm. *IEEE transactions on acoustics, speech, and signal processing*, 27(1):13–18, 1979.

[5] Jaakko Astola and Pauli Kuosmanen. *Fundamentals of nonlinear digital filtering*. CRC press, 2020.

[6] Thomas Gloe and Rainer Böhme. The dresden image database'for benchmarking digital image forensics. In *Proceedings of the 2010 ACM symposium on applied computing*, pages 1584–1590, 2010.

[7] Mauro Barni and Andrea Costanzo. A fuzzy approach to deal with uncertainty in image forensics. *Signal Processing: Image Communication*, 27(9):998–1010, 2012.

[8] Jiansheng Chen, Xiangui Kang, Ye Liu, and Z Jane Wang. Median filtering forensics based on convolutional neural networks. *IEEE Signal Processing Letters*, 22(11):1849–1853, 2015.

[9] Yuan Rao and Jiangqun Ni. A deep learning approach to detection of splicing and copy-move forgeries in images. In *2016 IEEE international workshop on information forensics and security (WIFS)*, pages 1–6. IEEE, 2016.

[10] Belhassen Bayar and Matthew C Stamm. A deep learning approach to universal image manipulation detection using a new convolutional layer. In *Proceedings of the 4th ACM workshop on information hiding and multimedia security*, pages 5–10, 2016.

[11] Belhassen Bayar and Matthew C Stamm. On the robustness of constrained convolutional neural networks to jpeg post-compression for image resampling detection. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2152–2156. IEEE, 2017.

[12] Yujin Zhang, Shenghong Li, Shilin Wang, and Yun Qing Shi. Revealing the traces of median filtering using high-order local ternary patterns. *IEEE Signal Processing Letters*, 21(3):275–279, 2014.

[13] Xinlu Gui, Xiaolong Li, Wenfa Qi, and Bin Yang. Blind median filtering detection based on histogram features. In *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2014 Asia-Pacific*, pages 1–4. IEEE, 2014.

[14] Abhinav Gupta and Divya Singhal. Analytical global median filtering forensics based on moment histograms. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 14(2):1–23, 2018.

[15] Hang Gao and Tiegang Gao. Detection of median filtering based on arma model and pixel-pair histogram feature of difference image. *Multimedia Tools and Applications*, 79(17):12551–12567, 2020.

[16] Hang Gao, Mengting Hu, Tiegang Gao, and Renhong Cheng. Robust detection of median filtering based on combined features of difference image. *Signal Processing: Image Communication*, 72:126–133, 2019.

[17] Xiangui Kang, Matthew C Stamm, Anjie Peng, and KJ Ray Liu. Robust median filtering forensics using an autoregressive model. *IEEE Transactions on Information Forensics and Security*, 8(9):1456–1468, 2013.

[18] Xiangui Kang, Matthew C Stamm, Anjie Peng, and KJ Ray Liu. Robust median filtering forensics based on the autoregressive model of median filtered residual. In *Proceedings of The 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference*, pages 1–9. IEEE, 2012.

[19] Xiao Jin, Peiguang Jing, and Yuting Su. Amfnet: an adversarial network for median filtering detection. *IEEE Access*, 6:50459–50467, 2018.

[20] Hongshen Tang, Rongrong Ni, Yao Zhao, and Xiaolong Li. Median filtering detection of small-size image based on cnn. *Journal of Visual Communication and Image Representation*, 51:162–168, 2018.

[21] Sajjad Ahmed and Saiful Islam. Median filter detection through streak area analysis. *Digital Investigation*, 26:100–106, 2018.

[22] Sajjad Ahmed and Saiful Islam. Robust median filtering forensics using texture feature and deep fully connected network. *International Journal of Information Technology*, 16(1):601–610, 2024.

[23] Kaijun Wu, Wanli Dong, Yunfei Cao, Xue Wang, and Qi Zhao. An improved method of median filtering forensics for enhanced image security detection. In *2021 International Conference on Networking and Network Applications (NaNA)*, pages 308–312. IEEE, 2021.

[24] Jun Zhang, Yixin Liao, Xinshan Zhu, Hongquan Wang, and Jie Ding. A deep learning approach in the discrete cosine transform domain to median filtering forensics. *IEEE Signal Processing Letters*, 27:276–280, 2020.

[25] Irwin Sobel, Gary Feldman, et al. A 3x3 isotropic gradient operator for image processing. *a talk at the Stanford Artificial Project in*, 1968:271–272, 1968.

ISBN: 979-11-88428-13-7
Copyright 2025, Global IT Research Institute (GIRI). All right reserved.
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on April 08,2025 at 05:52:39 UTC from IEEE Xplore. Restrictions apply.

Indian Journal of Advanced Physics (IJAP)
ISSN: 2582-8983 (Online), Volume-5 Issue-1, April 2025

Space Physics of the Universe and the Evolution of π Space Quantum

Nishant Sahdev, Chinmoy Bhattacharya

Abstract: ***The mathematical constant π, defined as the ratio of a circle's circumference to its diameter, has remained an important base of mathematics and science, yet its origins and connections to fundamental universal properties remain elusive. Traditionally viewed through a mathematical lens, this study proposes a novel perspective, linking π to the space-time of the universe. Building on Lord Kelvin's mathematical formulation involving the square root of π, we explore its physical significance as a measure of the interplay between surface and bulk properties of matter in the cosmos. We demonstrate that π governs critical relationships in spatial pressure dynamics, linking "direct space," "inverse space," and "hybridised space" in equilibrium. Furthermore, π is shown to underlie the transition between bulk and surface phenomena, influencing fundamental processes such as harmonic motion. These findings suggest that π is not merely a mathematical abstraction but a universal parameter deeply embedded in the structure and dynamics of space-time. This perspective opens new avenues for understanding π in relation to the physical universe, offering insights into its non-converging nature and its role as a driving force in cosmic processes.***

Keywords: Mathematical Constant π, Space-Time Fabric, Surface and Bulk Phenomena, Spatial Pressure Dynamics, Universal Parameters.

Abbreviations:
EFE: Einstein Field Equations
PFF: Push Forward Forces
TSQ: Theory of Space Quantization
PBF: Pull Back Forces

I. INTRODUCTION

The space of the universe is composed of space points or the point spaces [10]. The physical variables of the space (in the form of quantum) [12], example - entropy, force, energy, EM - wave, time, mass, etc., and others are formed from the translation, rotation, and vibration of the said space points.

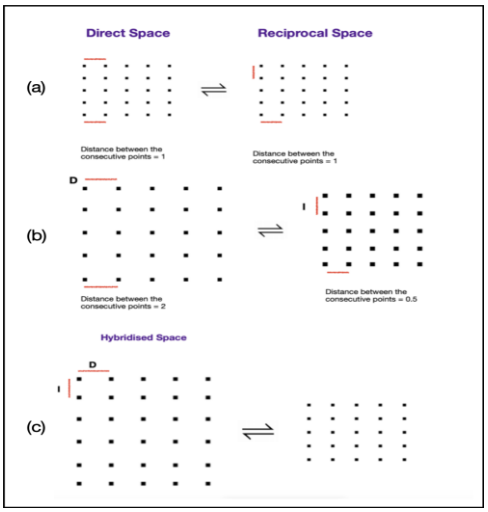
There does exist an equilibrium between 'Direct linear space' and 'Reciprocal or inverse linear space'. In the following Figure 1, the said three spaces have been shown diagrammatically. The concept of the 'direct space - inverse or reciprocal space' is:

Manuscript received on 10 January 2025 | First Revised Manuscript received on 12 January 2025 | Second Revised Manuscript received on 18 March 2025 | Manuscript Accepted on 15 April 2025 | Manuscript published on 30 April 2025.
Correspondence Author(s)
Nishant Sahdev, University of North Carolina at Chapel Hill, NC 27599, United States of America. Email: nishantsahdev.onco@gmail.com, ORCID ID: 0009-0007-2249-1006
Chinmoy Bhattacharya*, Austin Paints & Chemicals Private Limited, 3 Ambika Mukherjee Road, Belghoria, Kolkata, West Bengal, India. Email: chinmoy00123@gmail.com, ORCID ID: 0000-0002-1962-0758

© The Authors. Published by Lattice Science Publication (LSP). This is an open access article under the CC-BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

I. The product of the distance between two consecutive points in the direct space (D) and the corresponding distance between two points in the inverse space (I) is always unity
II. Whenever there is any event taking place in either of the two said spaces, the inverse event will take place in the other space. For example following figure 1 (b), the distance between the two consecutive points of direct space is doubled, the distance of the two consecutive points in the inverse space will be halved such that the product of the two new distances retained to be unity.
III. The hybridised space of the universe is formed from the 'direct space' and the 'reciprocal space' in the manner shown in Figure 1(c) below.

So each side (H) of a 'unit cell' of the hybridised space is a geometric mean of D and I: $H = (D \times I)^{1/2.}$



[Fig.1: (a, b & c) Geometry of Direct Space, Inverse Space & Hybridised Space] [10]

This is to understand the fact that the space of the universe must exist in the form of 'hybridised' space, as explained above. Otherwise, the space would have either expanded infinitely or contracted to such an extent that everything would have completely collapsed.

We learn a great deal from the literature about various aspects of the universe, including phenomena such as 'space ripples,' 'gravitational waves,' and 'distorted spacetime,' among others [14]. Notably, the parameter π is intricately related to these phenomena, as seen in the Einstein Field Equations (EFE) of the general theory of relativity. However, there is limited scope in the current literature for students and the scientific community to gain a clear understanding of the physical significance of the parameter π [15]. Until the parameter π is thoroughly understood, the cloud of ambiguity surrounding the

© Copyright: All rights reserved.

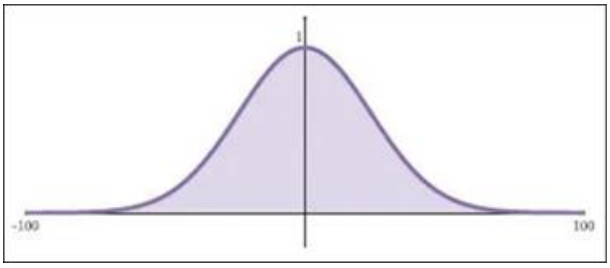
theories of relativity, cosmology, and space physics will continue to loom overhead.

If the space points in the 'direct space' and the 'reciprocal space,' or the 'direct space' and the 'reciprocal space' themselves, continued to remain in equilibrium with each other forever, no energy [4], no mass, no matter would have formed at all, and no evolution would have taken place—neither of plants, trees, rivers, oceans, mountains, the animal kingdom, nor the human race of the universe..

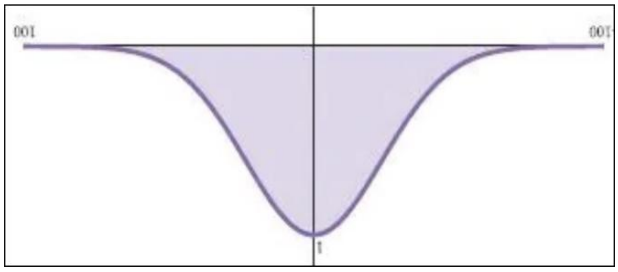
For matter to form, the space points must come closer to each other and interact. In the linear 'direct space' and the associated 'reciprocal space,' the mathematical probability of the space points coming closer to each other is very low because the said points are static. For matter (mass and energy) to form, the space points must move from the bulk interior of the space to the surface, interact with each other, and then flow back to the bulk of the space. A dynamic equilibrium must exist in the space as follows.

Rate at which the space points move from the bulk to the surface = Rate at which the space points move from the surface to the bulk of the space

So, to enable the above equilibrium to occur, space possesses its own impulses, which are short-lived 'push forward forces (PFF)' and 'pull back forces (PBF)'. As shown in Figures 2 and 3, these forces are represented as exponential functions. They exhibit peak maxima or minima, are of a non-converging type, and extend infinitely in both directions.



[Fig.2: The Short-Lived Force Impulse of the Direct Space of the Universe in the form of Square Root of Pi] [12]



[Fig.3: The Short-Lived Inverse Force Impulse of the Reciprocal or Indirect Space of the Universe in the form of Inverse of Square Root of Pi] [12]

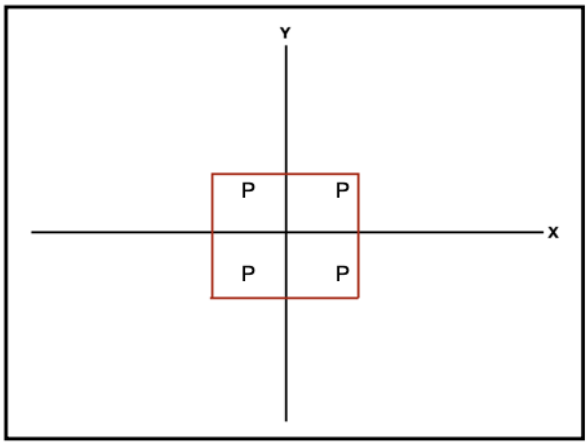
The area under the curves of the Figure 2 and 3 is being represented by Lord Kelvin equations as given below [12]:

$$\int_{-\infty}^{\infty} e - x2\, dx = \sqrt{\pi}. \quad \ldots \quad (1)$$

The area under the curve of Figure 3 is obtained by inverting the LHS and RHS of the above equation (1), which gives us the expression $(1/\pi)^{1/2}$. The value of the above integral, or the area under the curve, is 1.772... (which is the sum of all the smallest possible unit squares under the curve in Figure 2, and the said unit squares are made up of the points in direct space) [13]. Though the above equation of Lord Kelvin is very important and significant for the proper understanding of the quantum nature of the universe, it must be noted that this equation has been highly neglected since its inception in the 19th century. Neither mathematicians nor physicists have fully realized the profound underlying philosophy of the said equation, and as a result, it has not earned the recognition it deserves. In this article, it will be straightforwardly proven that the principles of classical physics, quantum physics, quantum mechanics, the theories of relativity, the recently discovered Theory of Space Quantization (TSQ) [2], and modern quantum computing are all based on π, and this π is grounded in this single equation of Lord Kelvin. The π parameter is an integral part of all the aforementioned theories, as well as an essential element of our very breath of life. The origin of this π can only be traced back to the space and the equation of Kelvin.

As has been mentioned earlier in this article [9], any event occurring in the direct space enforces an inverse event in the reciprocal space [10]. So this 'PFF short lived impulse $(\pi)^{1/2}$' of the direct space is reciprocated as a 'PBF short lived impulse $(1/\pi)^{1/2}$' of the reciprocal space [3]. Now, in direct space, the superposition (overlapping or hybridisation) of the four numbers of 'PFFs' occurs, and a π quantum is evolved from it. Similarly, the overlapping of the four numbers of 'PBFs' in reciprocal space produces an 'anti-π or inverse π space quantum.

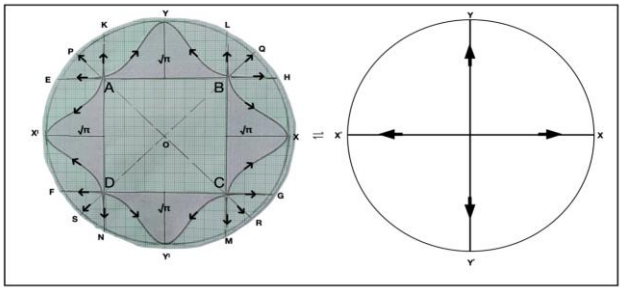
- Note that to obtain the square of any variable, say P, representing a straight line segment, the usual practice is to place the variable P in both longitudinal and lateral positions first, and then place them at equal distances in the diametrically opposite direction. The enclosed area is computed as shown in Figure 4 below, and the same principle has been followed to find the topology of π and (1/π), respectively.



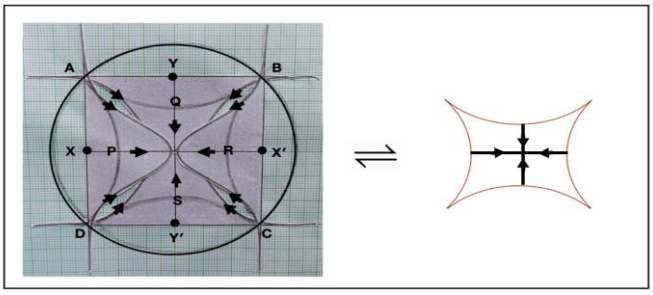
[Fig.4: Process of Finding out the Square of a Variable P] [3].

Figures 5 and 6 below show the mechanism of formation of the π quantum and anti π quantum respectively.

© Copyright: All rights reserved.



[Fig.5: Evolution of π Quantum From √π From the Space Time in the from of the Smallest Circle of the Universe] [1]



[Fig.6: Evolution of Anti π Quantum (2-D Saddle)] [5]

In Figure 5, the Following Points are to be Noted:

1. At the junction points A, B, C & D of the four numbers of √π areas, the two perpendicular forces are acting in each of the points (for point A, AK & AE, for point B, BL & BH, for point C, CG and CM, for point D, DF, DN and the resultants of the said forces are AP, DS, RC & DS respectively. So due to this resultant forces acting on points A, B, C & D, the points A, B, C & D are driven outwards and the new positioning of the said 4 points become P, Q, R & S respectively.
2. The sixteen numbers of points starting from X, H, Q, L, Y, K, P, E, X′, F, S, N, Y′, M, R, G, in fact form the circumference of a circle and whose area would be π.
3. This said π, or the circle formed, is the smallest possible circle in the universe, with an area of π and a radius of unity (OP = OQ = OR = OS = 1.0). It should be noted that this unit length represents the smallest possible length in the universe.
4. Now, in Figure 5, O is the midpoint of YY′. As the length of YY′ increases, the distances between points P & Q and R & S will also increase, causing the circumference of the circle to increase. Therefore, the circumference of the circle is directly proportional to the diameter YY′, and the constant of proportionality is indeed π. The ratio of the circumference to the diameter of a circle is constant, but the reason behind this constancy, or the mechanism responsible for it, has yet to be discovered. In this article, the long-standing mystery is revealed.
5. In Figure 6, it is shown that four (1/√π) areas come into close proximity with each other to form a (1/π) area or a 2D saddle APDSCRB. In fact, partial overlaps of the (1/√π) curves occur, and as shown in Figure 6, a significant pullback situation develops, causing the points X, X′, Y, and Y′ to be pulled inward and shifted to the points P, Q, R, and S, respectively, thus forming the saddle APDSCRB. The area of the 2D saddle [8] thus formed is (1/π).
6. Einstein postulated that somewhere in the universe, there does exist a point of singularity [14]. However, it was a qualitative statement, which originated from the high level of his intuitiveness [12]. Here, it is shown in the 'inverse space' that point O is the point of singularity, as all the forces are directed towards point O only (as shown by the arrows in Figure 6) [5]. In the recently discovered Theory of Space Quantization (TSQ) it has been shown that from the said point of singularity [11], the entire cosmological phenomena of the universe [7] have evolved.

A. Different Interpretations of π and (1/π)

- *Interpretation 1:* Regarding Figure 5, it should be noted that the space points along the diameter XX′ of the circle are in a state of certainty, but as soon as they move to the circumference of the circle, the positional uncertainty increases. Now, XX′ can be considered the bulk, and the circumference is the surface. Therefore, π can be interpreted as an index of:

'Push of the space points from bulk certainty to surface uncertainty' or 'the forward push forces of the universe' in the form of the quantum of the smallest possible volume of direct space.'

For the same reasoning as above but in the inverse sense, (1/π) is an index of:

'Pull of the space points from surface uncertainty to bulk certainty' or 'the backward pull forces of the universe' in the form of the quantum of the smallest possible volume of inverse space.'

- *Interpretation 2:* The forward push force, by which the space points from the bulk of direct space are driven to the surface of direct space (in the form of π or multiples of π), multiplied by the pull-back force, by which the space points from the surface of inverse space are driven toward the bulk of inverse space (in the form of the reciprocal of π or its multiples), is unity.

- *Interpretation 3:* The 'pressure,' P, of the hybridised space of the universe in a state of equilibrium is unity, since. P = [(push forward force, π) x (pull back force, 1/π) = 1. The pressure is a dimensionless parameter indeed. In the 'time-space' of the universe the above said equilibrium does persist but sometimes in a day the π –(1/π) equilibrium shifts to the right hand side or to the left hand side and as a result the pressure fluctuates (either P>1 or P< 1) and such fluctuations in pressure leads to the 'tides' in the rivers and oceans and as a result the water rises to high level when P>1 and the level falls when P<1.

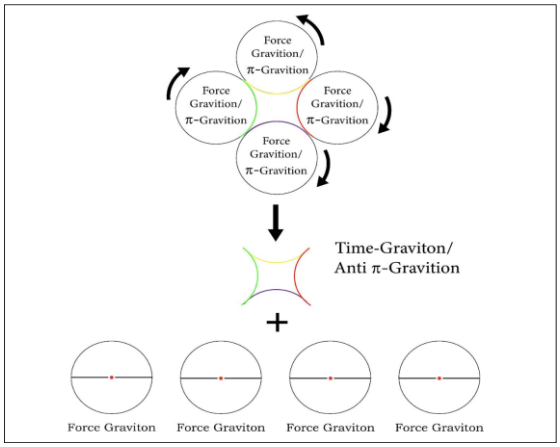
- *Interpretation 4:* π is an expanding factor of linear space that is transformed into non-linear space or space with curvature. As shown in Figure 5, the points of the linear space (A, B, C, and D) are mapped to points on the curvature of the circle. For example, if we consider that throughout the length XX′ in Figure 5, there are n points in space and the radius is r, then the average distance between the space points is (r/n) [1]. When these points are shifted and placed on the circumference of the circle [14],

the average distance between the points becomes (πr/n), which is greater than (r/n)

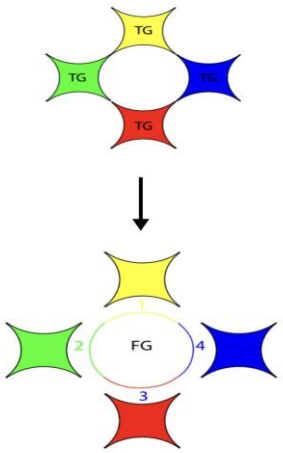
© Copyright: All rights reserved.

[16]. Therefore, another definition of π in relation to the space points would be: π = (Average inter space points distance in the curved direct space)/ (Average inter space points distance in the linear direct space) (1/π) = (Average inter space points distance in the curved direct space)/ (Average inter space points distance in the curved indirect space) So π is the factor of the enlargement or expansion of direct linear space and (1/π) is the factor of the squeezing of the linear direct space [17].

- *Interpretation 5:* π and 1/π, being the universal space quanta of the universe, are interconvertible with each other. In Figure 7 below, from the topological orientations of four π quanta in space, one anti-π quantum is formed [6]. The corollary is also true: from the topological orientations of four anti-π quanta in space, one π quantum is formed [7].



[Figure 7 (a) Formation of π Quantum from Anti π Quantum] [7]



[Fig.7(b): Formation of π Quantum from Anti π Quantum in the Space] [7]

B. From the Very Ancient Days People Have Been Educated Mainly the Following Properties of π

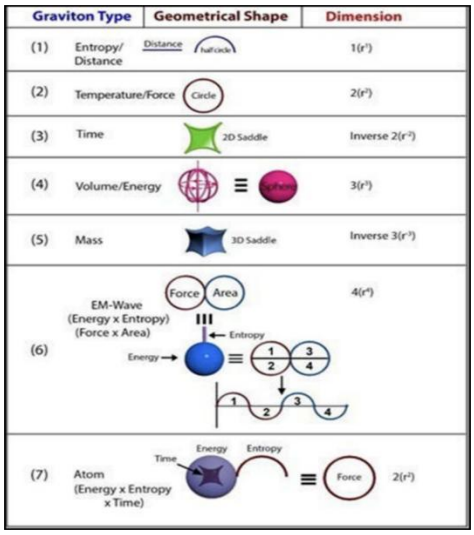
1. The parameter π is a universal constant in the fields of science, principally in mathematics, physics, and geometry.
2. The parameter π is irrational, and its value is 3.14159265359…; it is also a non-terminating, non-repeating decimal.
3. It is the ratio of the perimeter to the diameter of a circle, which always yields a non-converging value of 3.14159265359…
4. The half-circumference of a circle subtends an angle of π radians at the center of the circle (such that π radians = 180 degrees).

Regarding the π parameter, numerous discussions have been held in the past, many research papers published, and countless seminars organised. Various equations for π have been proposed over time, in all corners of the globe, to understand its origin and its mysterious non-converging character. This is, in fact, an endless exercise, and discussions are still ongoing to this day, but the mystery remains elusive.

The parameter π is a closely related, real-life constant of the universe, omnipresent and reminiscent of the eternal proverb, 'God is omnipresent.' The global scientific community has been fascinated by this god-like parameter, π, and honors it by observing a particular day each year as International π Day—March 14th.

The current research article reveals all the mysteries of π. The non-converging nature of π is deeply rooted in the geometry of the square root of π, as shown in Figure 2. The area under the curve in Figure 2 asymptotically extends to infinity, imparting the form of an irrational number. Since the square of any irrational number is always irrational, π, being the square of the square root of π, is also irrational. Similarly, the reciprocal of π (1/π) is also irrational, since the reciprocal of the square root of π is irrational as well (the reciprocal of any irrational number is always irrational). The circle shown in Figure 5 represents the smallest possible circle in the universe, with its radius considered to be unity. Similarly, the 2D saddle shown in Figure 6 represents the smallest 2D saddle in the universe, with its radius also considered to be unity.

[The π is the most fundamental quantum or packet (π quantum) of the following push forward quantum of the direct space of the universe (as shown in Figure 8)] [9].



[Fig.8: Geometrical Profile of the Space Quantum of the Universe] [9]

C. π Circles and Energy Continuum

The space π circles formed from π cannot have radiuses (r) of any value of desire. The values of r have to be full integers like 1, 2, 3, 4, 5, 6 etc..since the lengths of the space (of all the direct space, inverse space and the hybridised space) are

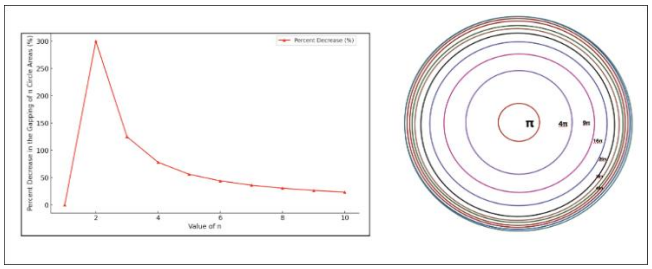
© Copyright: All rights reserved.

quantized and the lengths cannot be a fraction. Niels Bohr proposed that the electrons cannot rotate around the nucleus in any orbit of their own choice but only in the selected orbits they can rotate and those are the quantized orbits. As a matter of fact what he had talked about is 'space quantization' but how and why space is quantized in the integral form of π (as described in the previous paragraph) could not be demonstrated by him. The mystery of space quantization is revealed and described in this article only for the first time in science through π 'space quantum'.

The following table shows how the energy levels increase from the smallest possible circle with an area of π ($\pi(1)^2$, where r is the radius of the smallest circle) to $\pi(2)^2$, $\pi(3)^2$, $\pi(4)^2$, $\pi(5)^2$, or 4π, 9π, 16π, 25π, and so on, up to $n^2\pi$, where n is an integer with values 1, 2, 3, 4, 5, etc.

Table-I: Percent Change of Gaping of the π Circles with Increase of n

Value of n	Value of n2(π)	Gap of the area of the circle over the previous value of n	Percent increase of grapping
1	*	-	
2	4π	3π	300%
3	9π	5π	125%
4	16π	7π	78%
5	25π	9π	56%
6	36π	11π	44%
7	49π	13π	36%
8	64π	15π	30.6%
9	81π	17π	26.5%
10	100π	19π	0.234



[Fig.9: Graphical Presentation of the Percentage Change of Gapping of the Area of the π Circle with Increase of n] [14]

From the data presented in Table1 and from the Figure 9 (Percentage decrease of gaping of the areas of the circles against the value of n) it is being clearly revealed that the gaping of the circles The energy decreases exponentially, and ultimately, the gap disappears. In the language of quantum physics, the energy levels of the electrons around the nucleus transition from quantized to a continuous state, or the energy levels become continuous from a state of 'discontinuous,' 'discrete,' or 'quantized.' The physical significance of the phenomenon is that higher energy levels cannot remain confined within a certain area or volume, as this would have led to the death of living organisms due to the high energy density. The spread of the higher energy levels must be greater so that their intensity is reduced at any point in space. This is why the 'energy density of space remains constant' [1]. This is also referred to as the safeguard of nature against the 'UV catastrophe' [14].

D. Simple Harmonic Motion and π

In Figure 5, the length OP is considered to be unity, representing the radius of the smallest circle. The space point O is at the bulk, and the point P is on the surface/circumference. Due to the angular force developed from point A, as shown in the figure, the space point O moves to position P, and point P reaches position S. Since the circumference of the circle is 2π (which is divided into 4 equal segments: PS, SR, RQ, and RP), the length of each segment is (π/2). Thus, the unit length of OP (the distance from the surface space point P to the bulk space point O) is transposed to a length PS [where length = (π/2), representing the distance between the two surface points P and S], and it becomes (π/2) times its original length. This effect is referred to as 'the surface-to-bulk certainty being transposed to surface-to-surface uncertainty of the space points.'

For the 2D saddle, as shown in Figure 6, by the same reasoning, this can be described as 'surface-to-bulk uncertainty being transposed to bulk-to-bulk certainty' [for the segments AQ and QO or DS and SO]. The effect of π is to increasingly randomise the space points, while the effect of (1/π) is to take the space points to an ordered arrangement, ultimately approaching 'singularity.'

In simple harmonic motion (the case of a horizontal pendulum), a π space quantum is responsible for the smallest possible oscillation of the pendulum. The first swing of the pendulum from its mean position is driven by the first angular force (say AP, as shown in Figure 5), which moves the pendulum bob from point P (the mean position) to position S (say extreme left). The next angular impulse of force, DS, brings the pendulum bob back to the mean position, where point S moves to point R. In the subsequent angular impulse, force CR moves the pendulum bob to the extreme right position, and point R reaches point Q. The final angular impulse brings the bob back to the mean position, and point Q returns to point P. This completes one full 'time period of oscillation' of the pendulum.

II. CONCLUSION

The following conclusions are being drawn:

1. There exist short-lived force impulses in direct space, in the form of the square root of π, which are Gaussian in nature. From the topological combinations of the four numbers of such impulses, the universal π parameter is formed in the shape of a circle.
2. The mechanism for the formation of circular quantum forces in the universe is revealed in the exercise conducted in this article, which was previously elusive.
3. There do exist short-lived inverse force impulses in the indirect or inverse space in the form of the square root of (1/π), which are also Gaussian in nature. From the topological combinations of the four numbers of such impulses, the universal (1/π) parameter is formed in the shape of a circle.
4. The mechanism of the formation of inverse circular quantum forces in the universe is revealed in the exercise conducted in this article, which was also elusive earlier.
5. The simple harmonic motion of a horizontal pendulum makes use of the four numbers of short-lived force impulses in direct space (in the form of the

© Copyright: All rights reserved.

square root of π) to complete one full cycle of evolution (T), which gives rise to the formation of the π space quantum. In other words, one π space quantum (which already exists in space) disintegrates and generates a full cycle of the pendulum's revolution. At the same time, due to this full oscillation, a π space quantum is evolved and subsequently put back into space. The principle of energy conservation is maintained in this manner. The moral of this phenomenon is that whatever quantum is borrowed from space to perform work (such as entropy quantum, force quantum, energy quantum, etc.) must be repaid to space. Similarly, if space borrows a quantum from a system, it repays it.

6. The physics of simple harmonic motion (such as the swinging of pendulums), as taught to students at the school and higher levels, must be updated immediately, as the current material being taught is largely abstract.

- The following concepts of conventional physics regarding the horizontal pendulum are contradictory:
- The velocity of the pendulum bob is maximum at the mean position.
- The accelerations of the pendulum bob (towards the mean position) are maximum at the extreme left and extreme right positions, respectively.

The contradictory nature of the above two statements lies in the fact that for each swing of the pendulum (a. mean position to extreme left, b. extreme left to mean position, c. mean position to extreme right, and d. from extreme right back to the mean position), the angular forces of exactly the same magnitude are acting. Hence, in each swing, the pendulum sweeps out the same volume of space and moves with the same velocity throughout its movement. This suggests that the simple harmonic motion of a pendulum is quantized, not continuous. Therefore, applying Newton's laws of motion (which are valid for continuous motion) and using equations such as Force = mass x acceleration to derive the kinetic equations of motion of a pendulum are not justified. This new approach needs to be discussed.

7. The non-converging character of π originates from the topology or geometry of the infinitely spread, short-lived force pulses of the universe in the form of the square root of π, and this non-converging character is retained in the value of π as well.

DECLARATION STATEMENT

After aggregating input from all authors, I must verify the accuracy of the following information as the article's author.

- **Conflicts of Interest/ Competing Interests:** Based on my understanding, this article has no conflicts of interest.
- **Funding Support:** This article has not been funded by any organizations or agencies. This independence ensures that the research is conducted with objectivity and without any external influence.
- **Ethical Approval and Consent to Participate:** The content of this article does not necessitate ethical approval or consent to participate with supporting documentation.
- **Data Access Statement and Material Availability:** The adequate resources of this article are publicly accessible.
- **Authors Contributions:** Each author has individually contributed to the article. Great Idea given by Dr. Chinmoy Bhattacharya and synopsis, conceptualization, methodology, writing and critical thinking done by Nishant Sahdev.

REFERENCES

1. Bhattacharya C (2020) Novel Quantum Gravity interpretation of Chemical Equilibrium, Free Energy, Dark Energy and Dark Matter of the Universe. Advances in Theoretical & computational physics 3(3): 86-95. DOI: https://dx.doi.org/10.33140/ATCP
2. Bhattacharya C (2020) Novel Quantum Gravity Approach to Evaluate The Dimensionalities and the Geometrical Profiles of the Chemical Reactions. International Journal of Scientific & Engineering Research 11 (4):373-391. https://www.ijser.org/researchpaper/NOVEL-QUANTUM-GRAVITY-APPROACH-TO-EVALUATE-THE-DIMENSIONALITIES-AND-THE-GEOMETRICAL-PROFILES-OF-THE-CHEMICAL-REACTIONS.pdf
3. Bhattacharya C (2020) Cosmology and Unified Quantum Gravity Theory of the Universe. Advances in Theoretical & Computational Physics. 3(3): 114-211. DOI: https://dx.doi.org/10.33140/ATCP
4. Bhattacharya C (2020) Unified Quantum Gravity Theory Driven Concepts of the Classical Laws of Physics, the Dark Energy, the General Theory of Relativity and the 'Zero-Energy Universe'. Advances in Theoretical & Computational Physics 3(4): 265-286. DOI: https://dx.doi.org/10.33140/ATCP
5. Bhattacharya C (2021) Novel Quantum Gravity Model of the Physics of Operability of Galvanic Cells and Electrical Power Generation. Advances in Theoretical & Computational Physics 4(1): 7-13. DOI: https://dx.doi.org/10.33140/ATCP
6. Bhattacharya C (2023) TTQG (Topological Theory of Quantum Gravity) Driven New Theory of colour Physics. Advances in Theoretical & Computational Physics 6(1): 1-9. DOI: https://dx.doi.org/10.33140/ATCP
7. Bhattacharya C (2023) Reformulating the Basics of Conventional Newtonian Physics, Quantum Physics and the Einstein Theories of Relativities Based on the newly discovered Topological Theory of Quantum Gravity (TTQG). Physical Science and Biophysics J 7(1): 1-23. https://medwinpublishers.com/PSBJ/reformulating-the-basics-of-conventional-newtonian-physics-quantum-physics-and-the-einstein-theories-of-relativities-based-on-the-newly-discovered-topological-theory-of-quantum-gravity-ttqg.pdf
8. Bhattacharya C (2023) The Newly Discovered Topological Theory of Quantum Gravity (TTQG) - a Multi block Compatibilizer Cum Modifier of the Existing Theories of Physics, Cosmology, Quantum Mechanics and Quantum Computing. Global Journal of Science Frontier Research 23(a4): 13-58. https://journalofscience.org/index.php/GJSFR/article/view/102674
9. Bhattacharya C (2024) Theory of Space Quantization and the Quantum Level Understanding of the True Logics of the Principles of Thermodynamics and Heat Engines. Global Journal of Science Frontier Research: A Physics and Space Science, (LCC): QC174.17. https://globaljournals.org/GJSFR_Volume24/2-Theory-of-Space-Quantization.pdf
10. Bhattacharya C (2024) Theory of Space quantization (TSQ) Driven Version of the Conventional Theories of Physics, Relativity, and Gravitation - A Concise Report. Open Access Journal of Astronomy (OAJA)" Medwin Publishers, DOI: 10.23880/oaja-16000134. https://medwinpublishers.com/article-description.php?artId=13353
11. Bhattacharya C (2024) Embodying the Abstract 'Time Variable' of the Universe through Thermodynamics and the Theory of Space Quantization (TSQ). Open Access Journal of Astronomy (OAJA), DOI: https://doi.org/10.23880/oaja-16000142.
12. Propp, J. (2021). The Square Root of Pi. *Math Horizons*, *28*(3), 14–17. https://www.jstor.org/stable/48664731
13. Lord Kelvin, & Tait, P. G. (1902). *Mathematical and physical papers* (Vol. 5). Cambridge University Press. https://archive.org/details/mathematicalphys05kelvuoft
14. Einstein, A. (1915). The Field Equations of Gravitation. In *The Collected Papers of Albert Einstein* (Vol. 6, Doc. 25). Princeton University Press. https://einsteinpapers.press.princeton.edu/vol6-trans/129
15. Einstein Albert (1916) "The foundation of the General Theory of Relativity Annalen

© Copyright: All rights reserved.

Der Physik 354 97:769. DOI: http://dx.doi.org/10.1002/andp.19163540702

16. Tandon, R. (2019). Bianchi Type V Universe and Bulk Viscous Models with Time Dependent Gravitational Constant and Cosmological Constant in General Relativity. In International Journal of Engineering and Advanced Technology (Vol. 9, Issue 1s3, pp. 251–256). DOI: https://doi.org/10.35940/ijeat.a1048.1291s319
17. More, R. (2024). Exploring Mathematical Models of Dark Energy: A Comprehensive Literature Review. In International Journal of Basic Sciences and Applied Computing (Vol. 10, Issue 10, pp. 1–4). DOI: https://doi.org/10.35940/ijbsac.k0522.10100624

AUTHOR'S PROFILE



Chinmoy Bhattacharya earned his PhD in Polymer Physics in 1988 and completed post-doctoral research on liquid crystal polymers at Laval University, Canada. Returning to India in 1991, he joined ICI India Ltd. and later founded his own paint company. He is a former chairman of the Indian Paint Association and a leading figure in India's paint industry. Bhattacharya's research spans quantum gravity, cosmology, and dark matter, culminating in a unified quantum gravity theory and a novel space quantization model. He has numerous publications in prestigious journals, including his work on a new initiator for free radical polymerization, published in *Polymer Chemistry* (Royal Society of Chemistry, DOI: 10.1039/COPY00180E). As a guest faculty member at the University of Calcutta, he teaches Colour Physics, Polymer Physics, and Rheology of Coatings to postgraduate students, continuing to inspire the scientific community.



Nishant Sahdev is a distinguished researcher in Theoretical Physics and currently serves as a Research Affiliate at the University of North Carolina at Chapel Hill, United States. With extensive national and international experience at prestigious institutions, Nishant's academic journey spans Delhi Technological University, New Delhi, India, and the University of Debrecen, Hungary. His work is based on quantum gravity, cosmology, and dark matter, contributing to the development of a unified quantum gravity theory and an innovative space quantization model. Nishant has published extensively in leading scientific journals and has further honed his expertise in advanced mathematics during research training at King's College London, England. Dedicated to expanding the frontiers of scientific knowledge, Nishant's current work focuses on pioneering theories and equations that challenge conventional understanding. Beyond his academic pursuits, he is also a passionate romance writer and a columnist for national newspapers, seamlessly blending science with creative expression.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of the Lattice Science Publication (LSP)/ journal and/ or the editor(s). The Lattice Science Publication (LSP)/ journal and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

© Copyright: All rights reserved.

scientific reports

Check for updates

OPEN

SS-EMERGE - self-supervised enhancement for multidimension emotion recognition using GNNs for EEG

Chirag Ahuja[1,2]✉ & Divyashikha Sethia[2]

Self-supervised learning (SSL) is a potent method for leveraging unlabelled data. Nonetheless, EEG signals, characterised by their low signal-to-noise ratio and high-frequency attributes, often do not surpass fully-supervised techniques in cross-subject tasks such as Emotion Recognition. Therefore, this study introduces a hybrid SSL framework: Self-Supervised Enhancement for Multidimension Emotion Recognition using Graph Neural Networks (SS-EMERGE). This model enhances cross-subject EEG-based emotion recognition by incorporating Causal Convolutions for temporal feature extraction, Graph Attention Transformers (GAT) for spatial modelling, and Spectral Embedding for spectral domain analysis. The approach utilises meiosis-based contrastive learning for pretraining, followed by fine-tuning with minimal labelled data, thereby enriching dataset diversity and specificity. Evaluations on the widely-used Emotion recognition datasets, SEED and SEED-IV, reveal that SS-EMERGE achieves impressive Leave-One-Subject-Out (LOSO) accuracies of 92.35% and 81.51%, respectively. It also proposes a foundation model pre-trained on combined SEED and SEED-IV datasets, demonstrating performance comparable to individual models. These results emphasise the potential of SS-EMERGE in advancing EEG-based emotion recognition with high accuracy and minimal labelled data.

Keywords EEG, Self-Supervised Learning, GNN, Constrastive Learning, Cross-Subject Emotion Recognition

The rapid advancement in data acquisition technologies and the exponential growth in data volume have made it increasingly inadequate to rely solely on supervised datasets to gain insights into environmental phenomena or understand the human brain's complexities. Supervised learning paradigms require vast amounts of labelled data, often expensive,time-consuming, and limited scope. In fields such as Computer Vision (CV)[1,2], Natural Language Processing (NLP)[3,4], and Audio Processing[5,6], recent progress has underscored the significance of leveraging extensive data available on the internet to develop foundation models. These models capture fundamental principles governing these modalities or their intermodal relationships, enabling robust performance even with limited labelled data.

Understanding brain dynamics through Electroencephalogram (EEG) signals provides crucial insights into neurological activities. These non-invasive signals are crucial in applications like emotion recognition, stress estimation, and sleep staging[7–9]. In the context of EEG-based emotion recognition, evaluations can be conducted in several ways to validate the robustness and generalizability of models:

1. **Subject-Independent Evaluation (a.k.a. Cross-Subject or LOSO):** This method tests the model's ability to generalise across individuals. It involves training the model on data from a subset of subjects and testing it on data from entirely new subjects not included in the training set. This evaluation is crucial for applications where the model must perform well on new users without retraining.
2. **Subject-Dependent Evaluation:** This approach tests the model on the same subjects but on different sessions or instances than those used for training. It assesses how well the model performs when it has already been exposed to data from the test subjects during training. This type of evaluation might be less challenging than subject-independent tests since the model is already tuned to the characteristics of the specific subjects.

[1]Department of Computer Science Engineering, Delhi Technological University, Rohini, Delhi 110042, Delhi, India. [2]Department of Software Engineering, Delhi Technological University, Rohini, Delhi 110042, Delhi, India. ✉email: cahuja1992@gmail.com

3. **Subject-Mixed Evaluation (a.k.a. Trial-Level Random Split):** In this evaluation, all EEG trials across subjects are pooled together and randomly split into training and testing sets (e.g., 80-20), without ensuring subject separation. As a result, trials from the same individual may appear in both training and test sets. While this setup facilitates rapid benchmarking of model architectures, it does not accurately reflect cross-subject generalization and may lead to inflated performance due to subject-specific feature leakage.

These evaluation strategies are essential for developing EEG-based systems that are both effective and adaptable to various real-world scenarios. However, unlike more straightforward data types encountered in Computer Vision (CV) or Natural Language Processing (NLP), EEG data collection is fraught with unique challenges. These challenges include variability in sensor types, data collection errors, and individual differences in demographics and health, all of which contribute to high noise levels and low signal. These issues are particularly pronounced in emotion recognition, where the relevant signals can be subtle and frequently masked by artifacts or unrelated physiological events. Representation learning isolates and enhances the informative components of EEG signals crucial for distinguishing between emotional states. It mitigates the need for extensive labelled datasets for each emotional state by leveraging the intrinsic properties of the data.

Self-supervised learning (SSL), a prominent method within representation learning, excels in extracting rich representations from unlabeled EEG data[10–13]. By learning from the inherent variability among individuals, SSL facilitates the development of models that are accurate and capable of generalising across diverse subjects[14,15]. Once trained, these models can be fine-tuned for specific downstream tasks, such as emotion classification, with only minimal labelled data. It represents a significant paradigm shift in machine learning, especially for applications with scarce or expensive labelled data. The SSL techniques can be broadly categorised into two main types:

- *Generative Models:* These models aim to reconstruct their input data, thereby learning rich and detailed data representations. The process involves predicting missing parts of data or generating new data instances based on the observed data, which helps the model capture the underlying structure of the input distribution.
- *Contrastive Models:* These models learn by comparing and contrasting different data instances. They typically work by distinguishing between similar (positive) and dissimilar (negative) pairs of data instances, which sharpens their ability to discern critical features that define the data.

Both approaches have proven effective across various domains, including image processing, natural language understanding, and audio recognition, by leveraging large amounts of unlabeled data to learn valuable representations without explicit annotations. Specifically, in the context of EEG signals, a survey by Rafiei et al.[16] indicates that generative modelling-based SSL techniques excel with smaller datasets. In comparison, contrastive learning techniques benefit from larger datasets.

Recent advancements in Self-Supervised Learning (SSL) have significantly impacted emotion recognition using EEG signals[17–20]. Based on state-of-the-art research, the following summary illustrates diverse SSL techniques and their applications to EEG data analysis, as detailed in Table 1.

Background and related work

The Multi-view Spectral-Spatial-Temporal Masked Autoencoder (MV-SSTMA)[17]employs a generative SSL approach to reconstruct signals across spectral, spatial, and temporal domains using the Masked AutoEncoder (MAE) technique, achieving high accuracy on the SEED dataset[21]. However, it shows a significant decrease in performance in self-supervised settings compared to fully-supervised ones. Additionally, MV-SSTMA's[17] reliance on multi-head attention for spatial modelling assumes all channels are interconnected, potentially leading to overfitting, as it does not consider the actual electrode placements.

Graph-based Multi-Space Self-Supervised (GMSS)[19]utilises a spatial contrastive approach via Simple Contrastive Learning (SimCLR)[11]within a Graph Neural Network (GNN). It models spatial and spectral domains, showing decent cross-subject performance on SEED[21], but lacks robustness on the more challenging SEED-IV[22] dataset, indicating limitations in model adaptability across datasets.

Work	SSL Tech.	Model	Domains	SEED[21]		SEED-IV[22]		Limitations
				Fully SL	SSL	Fully SL	SSL	
MV-SSTMA[17]	Generative	Multi-Domain reconstruction (MAE)	Spectral, Spatial and Temporal	95.32	83.49	92.82	82.55	Doesn't explore the Contrastive Learning
GMSS[19]	Spatial Contrastive (SimCLR)	GNN	Spatial and Spectral	86.52	76.04	86.67	65.31	Cross-subject is decent across SEED[21]but poor on SEED-IV[22] shows it lacks across different datasets.
DS-AGC[20]	Contrastive (SimCLR)	GNN	Spatial and Spectral	85.00	82.18	65.79	61.32	SEED[21]numbers are good for cross-subject whereas SEED-IV[22] fully-supervised numbers show that the model is not robust.
SGMC[18]	Meiosis-based Contrastive Learning	ResNet(CNN)	Temporal	94.04	89.83	NA	NA	Only models in the temporal domain, without fully exploiting the potential of spatial interactions between channels. Also, it does not evaluate cross-subject evaluation.

Table 1. State of the art Emotion Recognition work done using SSL. SL: Supervised Learning, SSL: Self Supervised Learning, Tech.: Technique.

Differential Entropy-based Self-Attentive Graph Convolution (DS-AGC)[20], which incorporates spatial contrastive learning with Simple Contrastive Learning (SimCLR)[11]and a Graph Neural Network (GNN) architecture, demonstrates good cross-subject performance on the SEED dataset[21]. However, it struggles with robustness on the SEED-IV dataset[22] in both fully-supervised and self-supervised settings. It indicates that the model architecture may not be sufficiently adaptable or resilient, as it fails to perform effectively even in fully-supervised scenarios.

Self-Supervised Group Meiosis Contrastive Learning (SGMC)[18]implements a contrastive model using SimCLR[11], over a group of subjects coming from the same stimuli while performing meiosis augmentation and contrastive loss for self-supervised learning with a Convolutional Neural Network (CNN) architecture, focusing on the temporal domain. It shows outstanding results on SEED[21]but does not evaluate its performance on SEED-IV[22], nor fully utilises spatial interactions between EEG channels. Furthermore, it lacks evaluation across cross-subject variations. Enhancing contrastive learning with meiosis-based techniques improves model performance by creating more diverse negative samples.

The works summarised in Table 1 collectively highlight a recurrent theme: while SSL techniques have significantly advanced emotion recognition from EEG data, they still face challenges in maintaining robustness across diverse subjects and datasets. These limitations often stem from model architectures that focus predominantly on individual domains such as temporal, spatial, or spectralâ€"without effectively integrating them.

Causal Convolutions, as employed in Temporal Convolutional Networks (TCNs), address temporal dependencies by ensuring that each time step is influenced only by the current and past inputs, preserving the causality intrinsic to EEG signals[23–25]. This feature is particularly beneficial for emotion recognition, where the sequence of brain activity reflects emotional transitions. Furthermore, dilated convolutions allow TCNs to capture long-range temporal dependencies without excessive computational overhead, providing an efficient alternative to recurrent networks.

*Graph Attention Networks (GATs)*dynamically enhance spatial domain modelling by assigning varying importance to interactions between EEG channels. Unlike static graph methods, GATs utilise an attention mechanism that identifies the most relevant electrode relationships for a given emotional or cognitive state, making them highly adaptable to the variability present in EEG data[26,27]. This capability is crucial for capturing spatial dependencies, as interactions between brain regions vary across subjects and tasks. GATs prioritise key spatial patterns, improving robustness in subject-dependent scenarios.

*Spectral Embeddings*leverage frequency-domain features to encode EEG signals into a compact and meaningful vector space. Differential Entropy (DE), in particular, has been widely recognised as one of the most practical features for emotion recognition, as it captures variability across critical frequency bands (e.g., delta, theta, alpha, beta, and gamma)[7,17]. By deriving DE-based embeddings before modelling spatial and temporal aspects, SS-EMERGE ensures that frequency-domain information is preserved and seamlessly integrated into the overall representation.

The combination of these techniquesâ€"temporal modelling with Causal Convolutions, spatial relationship modelling with GATs, and frequency-domain encoding with Spectral Embeddingsâ€"offers a unified and holistic framework in SS-EMERGE. Each component contributes uniquely to addressing the inherent challenges of EEG-based emotion recognition. Temporal modelling captures the dynamic evolution of emotional states, spatial modelling identifies critical inter-channel relationships, and spectral embeddings provide robust frequency-domain features. Together, these techniques enable SS-EMERGE to achieve superior generalizability and performance across diverse datasets and subject variations, overcoming the limitations of methods that prioritise single-domain representations.

This work proposes a comprehensive approach to enhance EEG-based emotion recognition by integrating self-supervised learning with multidimension graph neural networks. The key innovations and contributions of this work are as follows:

1. Proposal of SS-EMERGE model for SSL training for Emotion Recognition with main contributions as follows:
   (a) *Temporal, Spatial, and Spectral Integration*: Combines causal convolutions for encoding temporal interactions, and using Graph Attention Transformer (GAT)[28]inspired by famous multivariate time series anomaly detection model, MTAD-GAT[27]for learning spatial channel interactions, and borrowing integration of spectral domain by using spectral embeddings from MV-SSTMA[17]. The multidimension integration ensures a holistic representation of EEG signals, capturing the intricate dynamics in different domains.
   (b) *Meiosis Augmentations Based Contrastive Learning*: Integration of the meiosis-based contrastive learning technique proposed by SGMC[18]for EEG. It performs contrastive learning on groups of stimuli, such as video samples in SEED[21]and SEED-IV[22], instead of individual subject samples. The augmentation technique cross-exchanges EEG segments within a video group, creating new augmented samples that facilitate contrastive learning and enhance the model's ability to generalise across different subjects and emotional states.
   (c) *Foundation Model for SEED and SEED-IV*: Proposes a foundation model for both SEED[21]and SEED-IV-[22]datasets that can be further fine-tuned on data collected through the same EEG sensor used by SEED-[21]and SEED-IV[22] to be used for the downstream task of Emotion Classification.
2. Evaluates Subject Dependent performance on SEED[21]and SEED-IV[22] datasets:

(a) Subject dependent SSL performance with respect to fully-supervised learning for SGMC[18] and SS-EMERGE.
(b) The effect of foundational pretraining while combining SEED[21]and SEED-IV[22] on cross-subject classification performance.

This work introduces a novel SS-EMERGE model that integrates SSL and GNN for advanced emotion recognition from EEG data. It addresses key challenges, including the need for extensive labelled datasets and handling high noise and variability in EEG signals. The SS-EMERGE model achieves 95.05% accuracy on SEED[21]and 83.67% on SEED-IV[22] with minimal labelled data for emotion recognition.

The structure of this paper is organised as follows: First, Section 2 defines the problem statement and details the proposed solution. Next, Section 4 showcases the experimental results obtained by applying the methodology described in Section 2. It is followed by a Section on Experimental Setup. Finally, Section 5 summarises the findings and discusses potential future research directions.

Methodology

This section details SS-EMERGE for emotion recognition using a comprehensive framework that integrates multiple domains of EEG data processing. The methodology includes a pretraining phase with Meiosis-based contrastive learning followed by a fine-tuning phase with labelled data. The system processes EEG data in spectral, spatial, and temporal domains, utilising specialised network components for each aspect.

Figure 1 describes the overall flow of SS-EMERGE that integrates a dual-phase approach to enhance the effectiveness of an EEG-based emotion recognition model. Phase-1 (pretraining) and Phase-2 (fine-tuning) followed by evaluation for the held-out subject, are briefly explained as follows:

1. **Phase-1 (pretraining without labels)**: This phase learns the generalised features from a diverse EEG dataset using self-supervised contrastive learning. This phase employs components such as meiosis-based constrative learning without using the emotion labels of the sample. By leveraging data from all subjects except one (N-1 subjects), this phase captures broad patterns and features inherent in the entire dataset, promoting generalisation across different contexts and conditions[29].
2. **Phase-2 (Fine-tuning with labels)**: The fine-tuning phase aims to adapt the pre-trained model to the downstream emotion classification task. This phase involves fine-tuning the pre-trained model with a classifier head using Cross-Entropy Loss on the same N-1 subjects. It ensures that the model is specifically tailored to recognise the emotional states from the EEG data of the given subjects.
3. **Prediction on the hold-out subject**: Finally, the prediction happens on all the sessions of the hold-out subject (the Nth subject) for evaluation.

Model architecture

The SS-EMERGE model architecture leverages a combination of advanced techniques to process EEG signals for emotion recognition effectively. This section explains each component in the model architecture, from input and initial processing, spectral, temporal, and spatial modelling to latent representation modelling. Figure 2

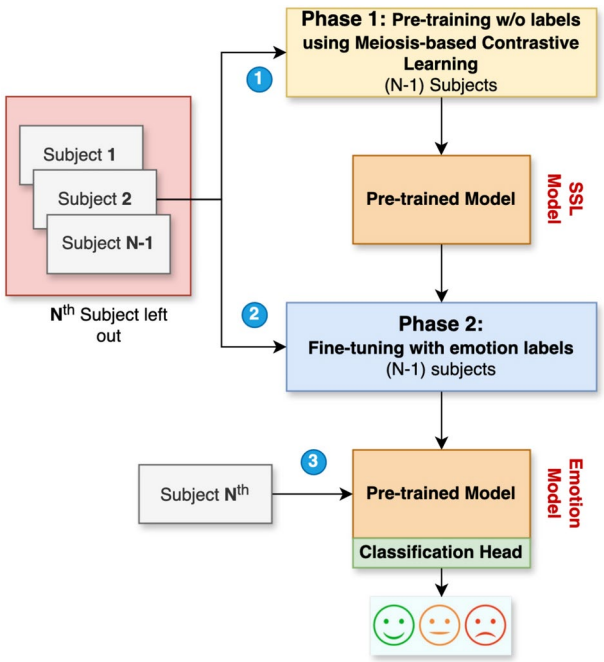


Fig. 1. SS-EMERGE Flow Diagram.

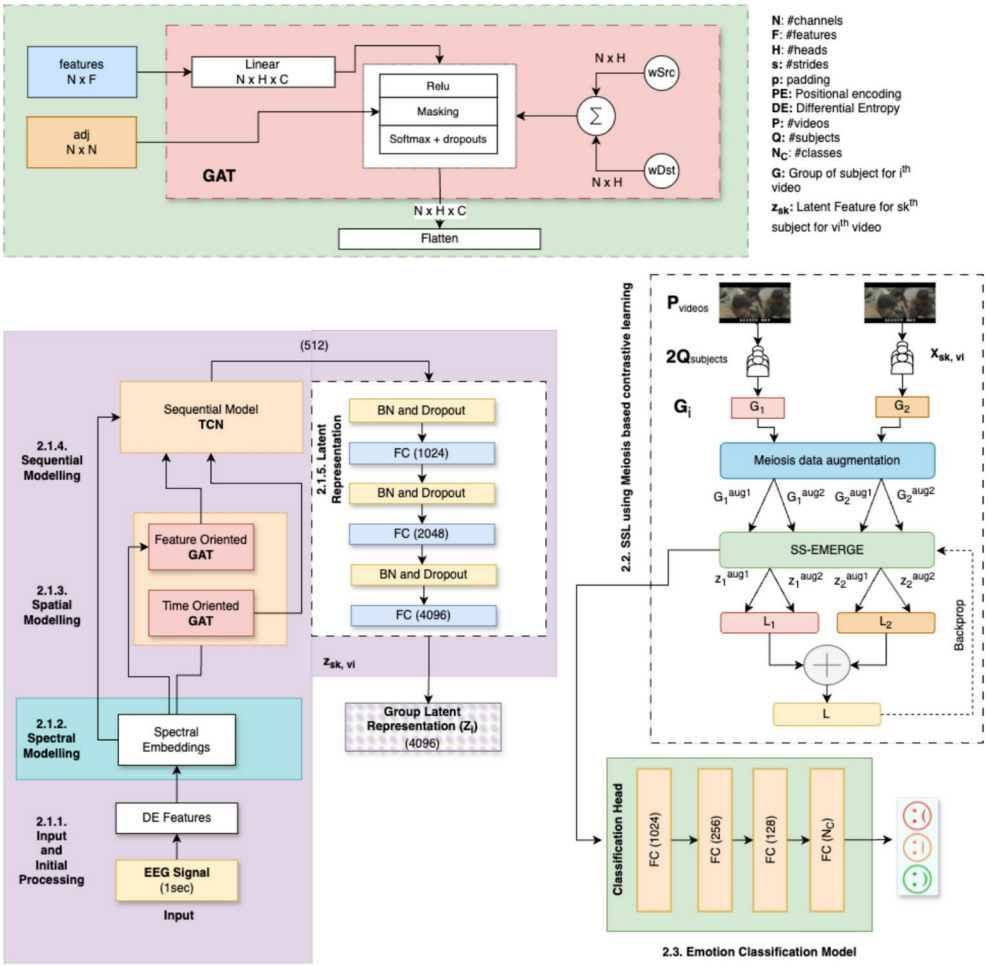


Fig. 2. SS-EMERGE backbone architecture for representation learning using SSL and fine-tuning for emotion classification label and associate with the text not clear.

illustrates the SS-EMERGE backbone architecture that learns the latent representations in Phase-1 (pretraining), followed by Phase-2 (fine-tuning), which fine-tunes the pre-trained network adding the classification for downstream emotion recognition task.

Input and initial processing
The input to the model consists of raw EEG signals segmented into 1-second intervals. These short segments help capture the immediate brain responses associated with different emotional states. From these EEG signals, Differential Entropy (DE)[9,30] features are extracted. DE features are robust in representing the spectral properties of EEG signals, making them particularly suitable for emotion recognition tasks.

Spectral domain modelling
SS-EMERGE leverages the Differential Entropy (DE) feature, which has demonstrated effectiveness in EEG-based emotion recognition tasks[31]. The model extracts DE features $X = (x_1, x_2, \ldots, x_N) \in \mathbb{R}^{N\times C\times F}$ from EEG signals in the spectral domain and transforms them into samples $\tilde{X} = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_N) \in \mathbb{R}^{N\times C\times T\times F}$ using an overlapping window of T seconds, where F represents frequency bands ($\delta : 1-4$ Hz, $\theta : 4-8$ Hz , $\alpha : 8-14$ Hz, $\beta : 14-31$ Hz, and $\gamma : 31-50$ Hz) converted by a Short-Time Fourier transform (STFT) for each channel C. In the spectral embedding layer, SS-EMERGE projects the i-th sample, $\tilde{x}_i$, which resides in $\mathbb{R}^{C\times T\times F}$, into a D-dimensional space using a linear layer. This process embeds the spectral information of the EEG signals into $\tilde{x}_i^{(s)} = \tilde{x}_i w^{(s)} + b^{(s)}$, where $w^{(s)}$ is the weight vector in $\mathbb{R}^{F\times D}$ and $b^{(s)}$ is the bias in $\mathbb{R}^{D}$. It finally reshapes $\tilde{x}_i$ to $C \times T \times D$dimensional embedding shape. Later, Positional Encoding (PE)[32] layer retains the temporal position of all channels C by adding sine-cosine encoding on each ith time step of signal with length T of every channel to facilitate later encoding.

Spatial domain modelling
Inspired by the Multivariate Time-series Anomaly Detection via Graph Attention Network (MTAD-GAT)[27], which first passes the time series through 1D Conv followed by two GAT networks. SS-EMERGE integrates two

types of Graph Attention Transformers (GAT)[26] to effectively capture spatial using Feature-Oriented GAT and temporal dependencies via Time Oriented GAT in the EEG signals.

1. **Feature Oriented GAT**: This layer constructs a graph where nodes represent different features simultaneously, and edges depict the relationships or interactions between these features. It effectively models the spatial relationships among various EEG features, enhancing the model's ability to understand how different features interact at a given moment in time.
2. **Time Oriented GAT**: This layer constructs a graph where nodes represent the same feature across different time points, and edges represent the temporal dependencies between these nodes. It captures the temporal progression of individual EEG features over time, allowing the model to track changes and developments in feature values across the duration of the signal[33].

Temporal domain modelling

In SS-EMERGE, **sequential modelling** is handled using a **Temporal Convolutional Network (TCN)** to capture temporal dependencies in EEG signals. Unlike recurrent architectures, TCN employs **causal convolutions**[34], ensuring that each output at a given time step is influenced only by current and past inputs while preserving the chronological order of EEG sequences.

The **TCN module** processes **spectrally and spatially enriched embeddings** using stacked **1D convolutional layers** with dilation, allowing the network to model **long-range dependencies** without the vanishing gradient problem of recurrent models. This makes it highly efficient for EEG-based emotion recognition.

As illustrated in Figure 2, sequential modelling in SS-EMERGE follows a structured pipeline:

1. **Input Features:** Spectral embeddings from the **Spectral modelling** module.
2. **Feature-Oriented and Time-Oriented GAT:** Extracts spatial relationships in EEG signals.
3. **Temporal Convolutional Network (TCN):** Models sequential dependencies over time.
4. **Batch Normalisation and Dropout Layers:** Prevent overfitting and stabilise learning.

Latent representation

Graph Attention Transformers (GATs) first process outputs through a series of Batch Normalization (BN) and Dropout Layers with a dropout rate of 0.2 before passing them to the Fully Connected (FC) layers. It starts with batch normalisation to stabilise the learning process and incorporates a dropout of 0.2 to prevent overfitting. Subsequent layers increase in neuron count from 1024 to 2048 and then 4096, each refining the features into a high-dimensional latent representation that encapsulates comprehensive EEG signal information.

Phase-1: SSL using Meiosis-based contrastive learning

Phase 1 (pretraining) learns the latent representation using Meiosis-based constrastive learning to the discriminative capabilities of a model by leveraging the augmented pairs created through Meiosis Data Augmentation. In the context of EEG data, constrastive learning learns distinct differences in brain responses that can signify various cognitive or emotional states in response to identical stimuli. Meiosis augmentation enhances EEG-based emotion recognition by cross-exchanging signal segments from different subjects within the same group. It creates augmented pairs, where positive samples come from subjects who watched the exact video (same stimuli), and negative samples come from different video groups (different stimuli). Meiosis-based Contrastive Learning helps the model to distinguish between similar and dissimilar stimuli, improving its ability to recognise emotions.

Meiosis data augmentation

Meiosis data augmentation aims to enhance EEG-based emotion recognition by augmenting one group sample to generate two groups that preserve the same stimuli-related features. It involves aligning stimuli in the EEG group to construct positive pairs, thereby increasing the meaningful difficulty of the model decoding EEG samples.

1. **Group Creation:**

   (a) *Sampling EEG Data:* Select $2Q$ subjects EEG recordings $(X_{s_1}, X_{s_2}, \ldots X_{s_{2Q}},)$ based on their exposure to specific video clips P videos $(v_1, v_2, \ldots, v_P)$, where each clip serves as a unique stimulus.

   (b) *Group Formation:* For each selected video clip v_i, form groups G_i consisting of EEG samples from the $2Q$ subjects. Thus, a group G_i for video clip v_i is defined as:

$$G_i = \{X_{s_1}^{v_i}, X_{s_2}^{v_i}, \ldots, X_{s_{2Q}}^{v_i}\}$$

Each group G_i shares similar stimuli-related features among its samples.

2. **Meiosis Data Augmentation:**

   (a) *Individual Pairing:* For an original group of EEG signals $G_i = \{X_{s_k}^{v_i} \mid k = 1, 2, \ldots, 2Q\}$ (corresponding to a video clip v_i), pair the signals randomly to form Q pairs: $\{X_{s_1}^{v_i}, X_{s_{1+Q}}^{v_i}\}, \{X_{s_2}^{v_i}, X_{s_{2+Q}}^{v_i}\}, \ldots, \{X_{s_Q}^{v_i}, X_{s_{2Q}}^{v_i}\}$.

(b) *Crossover:* Meiosis receives a randomly given split position c (where $1 < c < T - 1$) to perform the crossover transformation for each pair, exchanging data of the first c sampling points between the two signals. It results in augmented pairs $\{\{\tilde{X}^{v_i}_{s_k}, \tilde{X}^{v_i}_{s_{k+Q}}\} \mid k = 1, 2, \ldots, Q\}$.
(c) *Separation:* The transformed signals are then randomly divided into two groups. Paired transformed signals are placed in different groups X and Y, which results in two homologous groups of EEG $\tilde{G}^i_X = \{\tilde{X}^{v_i}_{s_k} \mid k = 1, 2, \ldots, Q\}$ and $\tilde{G}^i_Y = \{\tilde{X}^{v_i}_{s_{k+Q}} \mid k = Q+1, Q+2, \ldots, 2Q\}$ that share similar group-level stimuli-related features.

3. **Constructing the Minibatch:**

(a) Equation 1 shows that for one minibatch of P group samples G, $2P$ group samples $\tilde{G}$ are obtained through the Meiosis augmentation process.

$$\tilde{G} = \{\tilde{G}^i_t \mid i = 1, 2, \ldots, P; t \in \{1, 2\}\} = \text{Meiosis}(G) \tag{1}$$

(b) $\tilde{G}^i_X$ forms a positive pair with $\tilde{G}^i_Y$ and forms negative pairs with any other $2(P-1)$ group samples.

Contrastive learning setup
Contrastive learning uses meiosis-augmented pairs to train a model that learns to distinguish between pairs that share the same stimuli label (positive pairs) and pairs that come from different stimuli labels (negative pairs). Figure 2shows that the pretraining phase maximises the similarity between positive pairs and minimises the similarity between the negative pairs in the latent space[5]. The complete contrastive learning happens along with the Meiosis Augmentation as follows:

1. *Augmentation:* Meiosis augmentation generates pairs of EEG signals $\tilde{G}^i_X$ and $\tilde{G}^i_Y$ by cross-exchanging segments from different subjects within the same group. For instance, $\tilde{G}^i_X$ might contain segments from subject i and $\tilde{G}^i_Y$ from subject j, both corresponding to the same stimuli.
2. *Feature Extraction:* SS-EMERGE extracts latent features z_X for the augmented pairs $\tilde{G}^i_X$ and $\tilde{G}^i_Y$, generated by meosis augmentation[35].
3. *Similarity Measurement:* Equation 2 calculates the similarity between the latent representations z_X and z_Y using cosine similarity. It quantifies how close the representations of the augmented pairs are in the latent space[36].

$$s(z_X, z_Y) = \frac{z_X \cdot z_Y}{\|z_X\| \|z_Y\|} \tag{2}$$

4. *Contrastive Loss:*Finally, Phase-1 employs constrative loss to train the model using SimCLR[11] loss function. The loss function penalises large deviations between the similarity scores and their expected values based on the temperature-scaled cross-entropy in Equation 3:

$$\ell^X_i = -\log \frac{\exp(s(z^X_i, z^Y_j)/\tau)}{\sum_{j=1, j\neq i}^{P} \exp(s(z^X_i, z^X_j)/\tau) + \sum_{j=1}^{P} \exp(s(z^X_i, z^Y_j)/\tau)} \tag{3}$$

Here, P denotes the number of groups, τ is the temperature parameter, and $[j \neq i]$ is an indicator function that equals 1 if j is not equal to i[37].
Equation 4 then calculates the combined loss across all the groups in a batch of Stochastic Gradient Descent.

$$L = \frac{1}{2P} \sum_{i=1}^{P} (\ell^X_i + \ell^Y_i) \tag{4}$$

Algorithm 1 outlines the end-to-end Phase-1 pretraining using contrastive learning with meiosis augmentation. It starts by extracting group-level features and projecting them to latent space. It helps obtain group representations. The algorithm then minimises the distance between positive pairs and maximises amongst the negative pairs[38].

Phase-2: fine-tune emotion classification model

In Phase-2 (fine-tuning), the pre-trained SS-EMERGE model generates the latent representation but further fine-tunes using the classification head seen in Figure 2 for emotion classification modelling. It only learns the parameters of the classification head without changing other model parameters.

Model Architecture and Training Procedures:

1. **Classifier Head**: The classifier consists of additional fully connected layers that further process the latent representation of dimension D. These layers are followed by a softmax layer to produce the final emotion class probabilities (N_C).

(a) *Fully Connected Layers*: The layers with 1024, 256, and 128 neurons progressively transform and reduce the dimensionality of the latent representation.

Initialization:
Set the number of epochs E_1
Initialize the model parameters
for each epoch $t = 1, 2, \ldots, E_1$ **do**
    Video Clip Sampling:
    Sample P video clips per iteration until all video clips are enumerated
    EEG Sample Extraction:
    Extract $2P \times Q$ EEG samples $D = \{X_{sk}^{vi} | i = 1, 2, \ldots, P; k = 1, 2, \ldots, 2Q\}$
    Pack these samples into groups $G = \{G_i | i = 1, 2, \ldots, P\}$
    Meiosis Data Augmentation:
    Randomly generate a fixed split position c (where $1 < c < M - 1$)
    Apply Meiosis crossover to each group to generate augmented samples $\tilde{G} = \{\tilde{G}_i^t | i = 1, 2, \ldots, P; t \in \{X, Y\}\}$
    Feature Extraction and Projection:
    Extract group-level features
    Project features to latent space to obtain group representations
    Loss Calculation:
    Calculate the loss L using Equation 3
    Model Update: state Perform backpropagation to calculate gradients state Update the parameters of the model using the optimiser
end for

Algorithm 1. Pretraining with Meiosis-Based Contrastive Learning

(b) *Class Logits (FC (N_C))*: This layer outputs N_C neurons, where N_C corresponds to the number of emotion classes.
(c) *Softmax Layer*: Converts the logits into class probabilities using Softmax.

2. **Training Procedure**:

(a) Finally, Phase 2 fine-tunes the classifier head over the pre-trained model using labelled EEG data.
(b) The training process involves minimising a cross-entropy loss, which learns to discriminate different predicted class probabilities and the true emotion labels.
(c) It optimises the network using Stochastic Gradient Descent (SGD), with appropriate learning rate scheduling to ensure convergence shown in Table 2.
(d) It regularises the model training using a dropout of 0.5 and batch normalisation to prevent overfitting.

Phase	Parameter	Value
SSL Phase	Learning Rate	1×10^{-3}
	Momentum	0.1
	Batch Size	64
	Epochs	3288
	Meiosis Augmentation - *P*	16 (video clips per iteration) (SGMC[18])
	Meiosis Augmentation - *Q*	2 (samples per group) (SGMC[18])
Classification Phase	Learning Rate	1×10^{-3}
	Batch Size	256
	Epochs	100
	Optimization	Adam with learning rate scheduling
	Loss Function	Cross-entropy
	Dropout	0.5

Table 2. Hyperparameters of the SS-EMERGE Model.

Experimental setup

Dataset

This work conducts experiments on the SEED[21]and SEED-IV[22] datasets, known for their robust emotional state annotations and varying class distributions. These datasets provide a comprehensive and balanced foundation for evaluating the performance of emotion recognition models.

1. **SEED-IV Dataset** [21]: Consists of EEG recordings from 15 subjects (7 males and 8 females) as they watched film clips designed to evoke specific emotions. Each subject contributed 45 samples, evenly distributed among three emotion labels: happy, sad, and neutral, resulting in a total of 675 samples[21]. The signals were recorded using a 62-channel ESI NeuroScan System, with a sampling rate of 1,000 Hz. To facilitate analysis, the signals were downsampled to 200 Hz and bandpass filtered between 0–75 Hz to remove unrelated artifacts.
2. **SEED-IV Dataset** [22]: Includes data from the same 15 subjects, with each providing 72 samples distributed across four emotion labels: happy, sad, fear, and neutral, leading to a total of 1,080 samples[22]. The experimental setup and data processing methods were similar to those used in the SEED[21] dataset. However, the bandpass frequency filter was set to 1–75 Hz for SEED-IV, and the DE features were normalised to ensure consistency across samples.

The balanced distribution of samples across subjects and emotion classes in both datasets ensures an unbiased training and evaluation process.

Data preparation

For both SEED[21]and SEED-IV[22] datasets, each trial of EEG signal in each channel undergoes L2 normalisation to ensure that the data from different trials are on a similar scale. The movie videos are divided into 1-second windows. Given the varying lengths of the trial videos, adjacent windows are segmented from front to back according to the time axis until the coverage of windows exceeds the video range.

Evaluation strategy

To rigorously assess the generalisation capabilities and performance stability of the SS-EMERGE model, this work employs a robust evaluation strategy involving subject-mixed and subject-independent testing frameworks. The Leave-One-Subject-Out (LOSO) cross-validation scheme is utilised to ensure comprehensive evaluation:

1. **Subject-Independent / Cross-Subject Evaluation (LOSO)**:
   - **Fully-supervised Learning**: For each cycle of LOSO, the model is trained on EEG data from $N-1$ subjects using all available labels. The model is then tested on the left-out subject, assessing its adaptability to new individual data unseen during the training phase.
   - **Self-Supervised Learning**: Begins with pretraining a base model on all subjects without labels to capture generalisable features. The model is then fine-tuned on minimal labelled data from the left-out subject, optimising the blend of learned representations with subject-specific adjustments.
2. **Subject-Mixed / Trial-Random Evaluation**:
   - Conducted using a random train-test split (80-20%) to evaluate the model's architecture and ability to generalise across a random data selection, focusing not specifically on individual subject differences.

Training parameters for SSL and classification model

Phase-1 (pretraining) and Phase-2 (fine-tuning) select the hyperparameters used by SGMC[18] along with the default settings of the optimiser. These settings yield good subject-independent and subject-dependent performance without the rigorous need for hyperparameter tuning. Table 2 lists all the hyperparameters with their corresponding default values during pretraining and fine-tuning of the Emotion Classification model.

Results and discussion

The results presented in this section evaluate the performance of SS-EMERGE in different experimental settings in the SEED[21]and SEED-IV[22] datasets. The evaluation focuses on three key aspects as follows:

- Experiment 1: Subject-Mixed and Cross-Subject Evaluation of Base-Model Architecture Using fully-supervised Learning
- Experiment 2: Comparative Analysis of Self-Supervised vs Full-Supervised Learning for Cross-Subject Adaptation
- Experiment 3: Foundational pretraining by combining SEED and SEED-IV.

Experiment 1: choosing a base model architecture

EEG-based emotion recognition is inherently challenging due to subject-specific variations. A cross-subject generalisation is particularly difficult, as models must adapt to significant intersubject variability. This experiment compares the ResNet used by SGMC[18] with the proposed SS-EMERGE, a fully-supervised model trained for EEG emotion classification. The objective is to determine the most effective fully-supervised backbone model for emotion recognition for cross-subject.

Model	Setup	Positive (%)	Neutral (%)	Negative (%)	Average (%)
ResNet	Subject-Mixed	89.7	90.0	89.8	89.83
	Cross-Subject	68.5	69.8	65.2	67.83
SS-EMERGE	Subject-Mixed	96.2	92.5	94.1	94.27
	Cross-Subject	87.4	83.1	81.9	84.13

Table 3. Performance of models on SEED[21] dataset under subject-independent and cross-subject setups.

#Samples	SGMC	SS-EMERGE
SEED Dataset[21]		
10% (fine-tuned)	73.29%	75.50%
50% (fine-tuned)	82.91%	86.13%
100% (fine-tuned)	90.75%	92.35%
SEED-IV Dataset[22]		
10% (fine-tuned)	52.30%	58.10%
50% (fine-tuned)	68.05%	76.75%
100% (fine-tuned)	75.20%	81.51%

Table 4. Self-Supervised Learning Comparison of SGMC[18]and SS-EMERGE on SEED[21]and SEED-IV[22] datasets.

Unlike previous studies that rely primarily on subject-independent evaluations, this work evaluates cross-subject, ensuring that the model is trained on $N-1$ samples of a partial sample of subjects and tested on rest. This approach rigorously assesses cross-subject generalisation.

From Table 3, ResNet achieves 67.83% accuracy under the cross-subject evaluation, demonstrating its limited ability to generalise across subjects. In contrast, SS-EMERGE outperforms ResNet, achieving 84.13%, highlighting its effectiveness as a fully-supervised backbone model for EEG-based emotion recognition. Both models exhibit stronger performance in the random-split setting due to similarities in intra-subject training and testing data.

Experiment 2: cross-subject evaluation under self-supervised and fully-supervised learning

Having established SS-EMERGE as the superior fully-supervised backbone in Experiment 1, this experiment evaluates SS-EMERGE, a self-supervised learning (SSL) approach that combines contrastive learning with classification. SS-EMERGE is compared against SGMC, which also employs SSL, and SS-EMERGE, the strongest fully-supervised model from Experiment 1.

The SEED[21]and SEED-IV[22] datasets are used to analyse how SSL compares with fully-supervised learning under cross-subject conditions. The model is fine-tuned using varying percentages of labelled data (10%, 50%, and 100%) to evaluate performance under different data availability scenarios.

From Table 4, SS-EMERGE consistently outperforms both SGMC and SS-EMERGE across all labelled data settings. Notably, when fine-tuned with only 10% labelled data, SS-EMERGE achieves 75.5% accuracy on SEED, compared to 74.12% for SS-EMERGE and 73.29% for SGMC. Also, when compared with fully-supervised counterpart, SEED[21] achieves 84.13% cross-subject performance as compared to 92.35% of SSL.

Experiment 3: foundational pretraining using SEED and SEED-IV datasets

This experiment evaluates the impact of foundational pretraining using the combined datasets of SEED[21]and SEED-IV[22]. Given that both datasets share identical channel configurations, data integration during pretraining is streamlined. The distinct video stimuli in each dataset require a specialized approach for stimulus identification. To address this, a composite identifier that merges the dataset and video ID is used, ensuring accurate alignment with subject-specific emotional responses.

Baseline models

The baseline models for this study include:

- **Vanilla SGMC**: This model operates solely in the temporal domain using a ResNet architecture, focusing on capturing dynamic temporal features.
- **SS-EMERGE**: An advanced model that integrates features across temporal (T), spectral (F), and spatial (S) domains using a Graph Attention Transformer (GAT)[28].

Ablation studies

Ablation studies are performed to assess the impact of each domain:

1. **SS-EMERGE.FT**: Focuses on the spectral and temporal domains, evaluating the models ability to utilize these features for emotion recognition.

#Samples	SGMC T	SS-EMERGE.FT FT(Proposed)	SS-EMERGE.TS TS(Proposed)	SS-EMERGE FTS(Proposed)
SEED Dataset[21]				
10%	73.29%	73.82%	73.51%	**74.11%**
50%	83.71%	84.02%	83.84%	**84.42%**
100%	83.04%	85.56%	85.12%	**88.87%**
SEED-IV Dataset[22]				
10%	58.33%	56.81%	55.53%	**57.92%**
50%	66.44%	62.32%	61.02%	**63.54%**
100%	77.64%	76.31%	75.08%	**82.75%**

Table 5. SSL Performance with foundational pretraining for SEED[21]and SEED-IV[22].

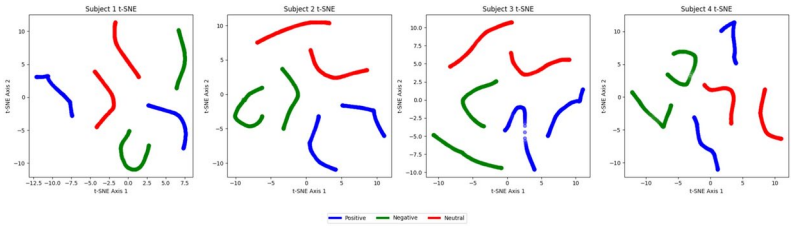


Fig. 3. TSNE visualization of SEED[21] embeddings.

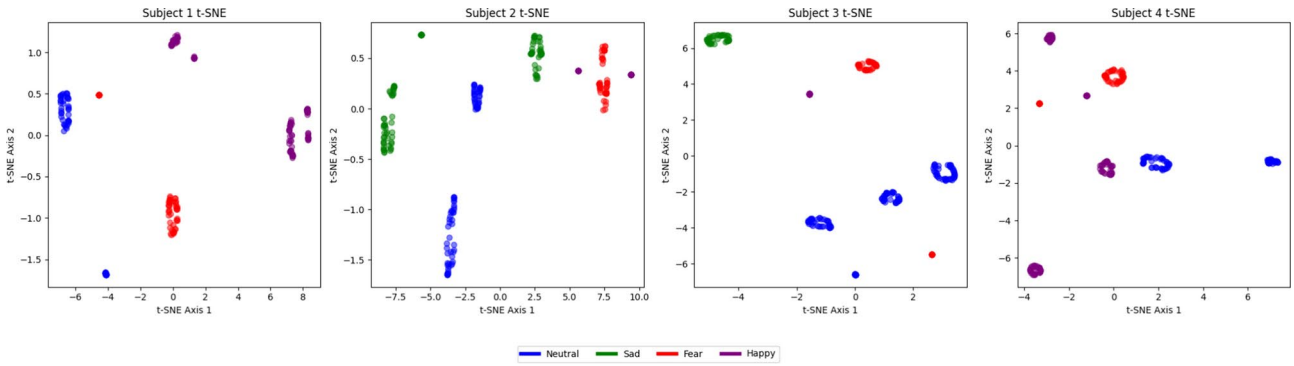


Fig. 4. TSNE visualization of SEED-IV[21] embeddings.

2. **SS-EMERGE.TS**: Analyzes the integration of temporal and spatial domains, omitting spectral features to determine the importance of spatial information.

These models are evaluated on a subject-independent basis due to computational constraints, which also guide the domain-specific ablation studies.

The SEED[21]and SEED-IV[22] datasets are processed under identical protocols to ensure uniformity, as detailed in subsection 3.2. Stimuli are uniquely tagged with dataset-specific prefixes, facilitating clear differentiation and integration into the SGMC model without modifications. Foundational pretraining employs contrastive loss to enhance feature discrimination across emotional stimuli without using labels.

In the fine-tuning phase, labels are reintroduced, tailoring the model to the distinct emotional patterns of each dataset. Performance is evaluated directly on the subjects from the training data, avoiding cross-subject validation.

As shown in Table 5, pretraining on the combined datasets significantly improves model accuracy 88.80% on SEED and 81.50% on SEED-IV[21,22], demonstrating the efficacy of cross-domain pretraining. This approach represents the first attempt to merge and assess these datasets for cross-subject performance.

Experiment 4: visualisation of underlying embedding

The t-Distributed Stochastic Neighbor Embedding (t-SNE) visualisations in Figures 3 and 4illustrate the feature representations from the SS-EMERGE on the SEED[21]and SEED-IV[22] dataset. The t-SNE projections reduce the 4096-dimensional feature space to a 2-dimensional space for visualisation. The clustering of SS-EMERGE learnt

features shows distinct and well-separated clusters. The more precise separation of clusters of the SS-EMERGE indicates that it enhances the model's ability to capture the discriminative features due to its rich representation learning.

Conclusion and future work

This study presents the SS-EMERGE model, a novel self-supervised learning approach using multidimension graph neural networks to enhance EEG-based emotion recognition. Evaluated on the SEED and SEED-IV datasets, SS-EMERGE consistently outperforms the SGMC baseline and fully-supervised models, particularly when labelled data is limited. The model shows substantial performance gains with just 10% and 50% labelled data and achieves its highest accuracy with full data utilisation, highlighting its potential for applications with scarce labelled resources.

Furthermore, SS-EMERGE demonstrates its capability as a foundational model for emotion recognition by efficiently leveraging combined datasets during pretraining. Future research could explore integrating additional diverse EEG datasets to improve model robustness further and investigate the efficacy of meiosis-based contrastive learning for enhanced group-level feature representation. This direction is promising for advancing emotion recognition technologies in real-world scenarios.

Data availability and access

The data used in this study is available on SEED[21] and SEED-IV[22]. Access to the datasets can be requested through the following link: https://bcmi.sjtu.edu.cn/home/seed/.

Received: 22 October 2024; Accepted: 14 April 2025
Published online: 24 April 2025

References

1. LeCun, Yann, Bengio, Yoshua & Hinton, Geoffrey. Deep Learning. *Nature* **521**(7553), 436–444 (2015).
2. Multiple Authors: Self-Supervised Learning for Visual Understanding. Nature Reviews Methods Primers (2023)
3. Devlin, Jacob, Chang, Ming-Wei, Lee, Kenton, & Toutanova, Kristina. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pp. 4171–4186 (2019).
4. Authors, Multiple. *Natural Language Processing: State of the Art* (Current Trends and Challenges, Multimedia Tools and Applications, 2023).
5. Oord, Aaron van den, Li, Yazhe, & Vinyals, Oriol. Representation Learning with Contrastive Predictive Coding. arXiv preprint arXiv:1807.03748 (2018).
6. Multiple Authors: Self-Supervised Learning in Audio. In: ICML 2023 Proceedings (2023). https://proceedings.mlr.press/v139.
7. Zheng, Wei-Long, & Lu, Bao-Liang. Investigating Critical Frequency Bands and Channels for EEG-Based Emotion Recognition with Deep Neural Networks. In: IEEE Transactions on Autonomous Mental Development (2015).
8. Li, Xin, Zhang, Xiangyu, Song, Xue, Li, Zhouyun, & Wu, Xiaohua Tony. Cross-Subject Emotion Recognition Using Deep Adaptation Networks. In: 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII), pp. 1–7 (2019). IEEE.
9. Zheng, Wei-Long., Liu, Jiahui & Lu, Bao-Liang. EmotionMeter: A Multimodal Framework for Recognizing Human Emotions. *IEEE Transactions on Cybernetics* **49**(3), 1110–1122 (2018).
10. Hadsell, Raia, Chopra, Sumit, & LeCun, Yann. Dimensionality Reduction by Learning an Invariant Mapping. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), vol. 2, pp. 1735–1742 (2006). IEEE.
11. Chen, Ting, Kornblith, Simon, Norouzi, Mohammad, & Hinton, Geoffrey. A Simple Framework for Contrastive Learning of Visual Representations. In: International Conference on Machine Learning, pp. 1597–1607 (2020). PMLR.
12. Zheng, Wei-Long., Zhu, Jiangjiang & Lu, Bao-Liang. Multichannel EEG-Based Emotion Recognition via Group Sparse Canonical Correlation Analysis. *IEEE Transactions on Cognitive and Developmental Systems* **9**(3), 281–290 (2017).
13. Kostas, Dimitrios, & Rudzicz, Frank. Thinker Invariance: Enabling Deep Neural Networks for BCI Across More than 100 Subjects. arXiv preprint arXiv:1811.00928 (2018).
14. Schirrmeister, Robin Tibor, Springenberg, Jost Tobias, Fiederer, Lukas, Dominique Josef, Glasstetter, Martin, Eggensperger, Katharina, Tangermann, Michael, Hutter, Frank, Burgard, Wolfram, Ball, Tonio. Deep Learning with Convolutional Neural Networks for EEG Decoding and Visualization. In: Human Brain Mapping, vol. 38, pp. 5391–5420 (2017). Wiley Online Library.
15. Bashivan, Pouya, Rish, Irina, Yeasin, Md, & Codella, Noel. Learning Representations from EEG with Deep Recurrent-Convolutional Neural Networks. In: International Conference on Learning Representations (2015).
16. Rafiei, M.H., Gauthier, L., Adeli, H., & Takabi, D. Self-supervised learning for electroencephalography. IEEE Transactions on Neural Networks and Learning Systems **PP**, 1–15 (2022) https://doi.org/10.1109/TNNLS.2022.3190448.
17. Li, Rui, Wang, Yiting, Zheng, Wei-Long, & Lu, Bao-Liang. A Multi-View Spectral-Spatial-Temporal Masked Autoencoder for Decoding Emotions with Self-Supervised Learning. In: Proceedings of the 30th ACM International Conference on Multimedia, pp. 4002–4010 (2022). ACM.
18. Zheng, Wei-Long, Duan, Jun-Cheng, & Lu, Bao-Liang. Self-Supervised Group Meiosis Contrastive Learning for EEG-Based Emotion Recognition. In: 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1335–1339 (2021). IEEE.
19. Li, R., Wang, Y., Zheng, W.-L., & Lu, B.-L. Graph-Based Multi-Space Self-Supervised Learning for EEG Representation Learning. In: Proceedings of the 29th ACM International Conference on Multimedia, pp. 1065–1073 (2021). ACM.
20. Song, Tinglei, Zheng, Wei-Long, Lu, Bao-Liang, Zong, Yixuan. EEG Emotion Recognition Using Dynamical Graph Convolutional Neural Networks. In: 2018 24th International Conference on Pattern Recognition (ICPR), pp. 3599–3604 (2018). IEEE.
21. Zheng, W. L. et al.: Investigating EEG-Based Emotion Recognition via Ensemble Learning. IEEE Transactions on Affective Computing (2015).
22. Zheng, Wei-Long., Zhu, Jia-Yi. & Lu, Bao-Liang. EmotionMeter: A Multimodal Framework for Recognizing Human Emotions. *IEEE Transactions on Cybernetics* **49**(3), 1110–1122 (2018).
23. Oord, A., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrenner, N., Senior, A., & Kavukcuoglu, K. Wavenet: A generative model for raw audio. arXiv preprint arXiv:1609.03499 (2016).
24. Bai, S., Kolter, J.Z., & Koltun, V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. arXiv preprint arXiv:1803.01271 (2018).

25. Ingolfsson, T.M., Wang, X., Hanna, G., Gugat, R., McAllister, J., & Frossard, P. Eeg-tcnet: An accurate temporal convolutional network for embedded motor-imagery brain-machine interfaces. arXiv preprint arXiv:2006.00622 (2020).
26. Velickovic, Petar, Cucurull, Guillem, Casanova, Arantxa, Romero, Adriana, Liò, Pietro, & Bengio, Yoshua. Graph Attention Networks. In: International Conference on Learning Representations (ICLR) (2018).
27. Hsu, T.-H., Tseng, Y.-C., & Chen, J.-J. Graph attention transformer for unsupervised multivariate time series anomaly detection. In: Proceedings of MACLEAN: MAChine Learning for EArth ObservatioN Workshop 2022, Co-located with the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD 2022), Grenoble, France (2022). https://ceur-ws.org/Vol-3343/paper3.pdf.
28. Gao, Z. et al.: Vital Signs Monitoring of Multiple People Using a FMCW Radar with Beamforming. IEEE Transactions on Microwave Theory and Techniques (2019).
29. Shen, Y. et al.: Comprehensive EEG-Based Emotion Recognition. IEEE Transactions on Affective Computing (2021).
30. Alarcao, Sergio M., & Fonseca, Manuel J. Emotions Recognition Using EEG Signals: A Survey. IEEE Transactions on Affective Computing **10**(3), 374–393 (2017).
31. Duan, L. et al.: EEG Emotion Recognition Based on Deep Learning Techniques. Journal of Neural Engineering (2022).
32. Vaswani, A. et al.: Attention Is All You Need. In: NeurIPS (2017).
33. Brody, S. et al.: Attentive Temporal Graph Convolutional Networks for Dynamic Graphs. In: ICLR (2020).
34. Oord, A. V. D. et al.: WaveNet: A Generative Model for Raw Audio. arXiv preprint arXiv:1609.03499 (2016).
35. Misra, I. et al.: Self-Supervised Learning with Augmented Multiview Embeddings. In: NeurIPS (2020).
36. Grill, J. B. et al.: Bootstrap Your Own Latent: A New Approach to Self-Supervised Learning. In: NeurIPS (2020).
37. Chen, X. et al.: Improved Baselines with Momentum Contrastive Learning. arXiv preprint arXiv:2003.04297 (2020).
38. Tian, Y. et al.: Contrastive Multiview Coding. In: ECCV (2020).

Author contributions

The authors confirm contribution to the paper: study conception and design: Chirag Ahuja; analysis and interpretation of results: Chirag Ahuja; draft manuscript preparation: Chirag Ahuja, Divyashikha Sethia. All authors reviewed the results and approved the final version of the manuscript.

Funding

The authors did not receive support from any organisation for the submitted work. No funding was received to assist with preparing this manuscript. No funding was received for conducting this study. No funds, grants, or other support was received.

Declaration

Competing interests

The authors have no relevant financial or non-financial interests to disclose. The authors have no conflicts of interest to declare relevant to this article's content. All authors certify that they have no affiliations with or involvement in any organisation or entity with any financial or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no financial or proprietary interests in any material discussed in this article.

Additional information

Correspondence and requests for materials should be addressed to C.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2025

2025 3rd IEEE International Conference on Industrial Electronics: Developments & Applications (ICIDeA) | 979-8-3315-3320-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICIDEA64800.2025.10962818

Stabilization and Performance Analysis of Double Link-Rotary Inverted Pendulum using LQR-I Controller

Anjali Rai
Electrical Engineering Department
Delhi Technological University
New Delhi, India
anjalirai_2k22phdee01@dtu.ac.in

Bharat Bhushan
Electrical Engineering Department
Delhi Technological University
New Delhi, India
bharat@dce.ac.in

Bhavnesh Jaint
Electrical Engineering Department
Delhi Technological University
New Delhi, India
bhavneshjaint@dtu.ac.in

***Abstract*—This paper presents an optimal LQR-Integral based controller (LQR-I) for balancing control of the Double Link Rotary Inverted Pendulum (DL-RIP) system. It is one of the challenging topics in control and robotics due to its nonlinearity and instability, also complex system due to 3 degree of freedom rotary system. This system uses a motor to control and balance two rigid links serially connected pendulums in a vertically upward direction. Pole-zero mapping is used to determine the stability of the system. In this paper, a performance analysis of the system is done with an LQR-I controller using three reference paths (Square, Sawtooth, and Sinusoidal wave), two of which are discontinuous and one continuous path. The system shows less overshoot and deflection to stabilize the pendulum vertically upwardly with a continuous sinusoidal reference path compared to a square and sawtooth wave.**

Keywords—LQR-I, DL-RIP, eigenvalues.

I. Introduction

An underactuated system is a type that has fewer actuators than the degree of freedom and has attracted many researchers in the control engineering field. Challenges occur in controlling these systems due to their complexity, path-tracking instability, and nonlinearity [1]. Some examples of these systems are robotics systems (flying robots, rotary pendulums, etc), marine vessels (ships like oil tankers, autonomous underwater vehicles), and aerospace vehicles. One of the well-known underactuated systems is the inverted pendulum, which has fewer actuators than the control variable and benchmark system used in research for broad performance analysis of different controllers. Different types of inverted pendulum systems are seen in the literature view, which can be grouped into two parts: dynamic cart-type and rotary-type [2]. Controlling the pendulum's position on a cart in linear motion has been done using a conventional proportional integral derivative (PID) controller. The issues in conventional control, such as high overshoot time taken to settle, are resolved using optimal and intelligent controllers. An LQR controller [3] was applied to control the arm and pendulum angle of the rotary inverted pendulum and to enhance the performance of the system's ode hybrid intelligent controller. A double-link serially connected rotary system has been designed and developed to enhance the system's complexity by checking the controllers' robustness and broad exploration [4]. In a double-link pendulum, the movement of one link also affects the second link's movement for stabilization at a vertically zero position. Section II contains mathematical modeling, the Controller is explained in section III, and the Result and Discussion are in section IV.

II. Mathematical Modeling of Double link Rotary Inverted Pendulum (DL-RIP)

The system configuration mainly depends on two factors: the method of actuators, which are two ways viewed in literature, linear and angular, and how many degrees of freedom the system has. Linear inverted pendulum links are directly coupled to the motor shaft, which means at least two degrees of freedom system, one for base position and another for angle pendulum, respectively [5]. For a higher degree of freedom, links in the system are increased, or a single link can move in various directions. As it increases, the complexity in the modeling of the system and its controlling is increased, and if the system is rotary, it becomes more difficult [6].

A. Model Convention

The rotary arm pivot is attached to the servo system and actuated. The servo turns in an anti-clockwise direction when motor voltage is greater than zero (Vm >0); thus, the arm also increases positively, with length Lr, movement of inertia Jr, and its angle theta when rotated in an anti-clockwise direction because both movements are interdependent [7]. The whole assembly of the double pendulum is connected to the rotary arm. The first short-sized, bottom pendulum has a length of L_1, mass m_1, and center of mass l_{m1} with a moment of inertia about its center J_1 connected to the base of the rotary arm. The second upper side pendulum has a length of L_2, mass m_2, and center of mass l_{m2} with a moment of inertia about its center J_2 connected to the end of the first pendulum with the hinge of mass m_h. The angle of the bottom pendulum is alpha ($\alpha>0$), and the top one is phi ($\varphi>0$), which is zero when the pendulum is perfectly balanced in a vertical upright position.

For highly complicated systems like robotic manipulators with multiple joints, a systematic method named the Lagrange method is used to determine the equation of motion of systems.

979-8-3315-3320-5/25/$31.00 ©2025 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:24:17 UTC from IEEE Xplore. Restrictions apply.

Particularly, the equation that explains the motion of the rotary arm and pendulum angle with respect to motor voltage is obtained through the Euler-Lagrange equation, as mentioned in equation (1).

$$\frac{\partial^2 L}{\partial t \partial \dot{q}_i} - \frac{\delta L}{\delta q_i} = Q_i \tag{1}$$

Generalized coordinates $q_{i,}$ also called Lagrange coordinates, are three for this system, as present in equation (2), and their first derivates are corresponding velocities. With the generalized equation (1), the Euler-Lagrange dynamics for this system are mentioned in equation (3)

$$q^T = [\theta(t), \alpha(t), \phi(t)] \tag{2}$$

$$\left.\begin{aligned} \frac{\partial^2 L}{\partial t \partial \dot{\theta}} - \frac{\delta L}{\delta \theta} &= Q_1 \\ \frac{\partial^2 L}{\partial t \partial \dot{\alpha}} - \frac{\delta L}{\delta \alpha} &= Q_2 \\ \frac{\partial^2 L}{\partial t \partial \dot{\phi}} - \frac{\delta L}{\delta \phi} &= Q_3 \end{aligned}\right\} \tag{3}$$

Now, the generalized force Q_i is used to describe the non-conservative force like viscous damping force (B), the torque generated by a motor (ζ_m) applied to a system with respect to generalized coordinates [8]. The generalized force acting on:

(a) Rotary arm

$$\left. Q_1 = \zeta_m - B_r\left(\frac{d}{dt}\theta(t)\right)\right\} \tag{4}$$

(b) Bottom and top pendulum

$$\left.\begin{aligned} Q_2 &= -B_1\left(\frac{d}{dt}\alpha(t)\right) \\ Q_3 &= -B_2\left(\frac{d}{dt}\phi(t)\right) \end{aligned}\right\} \tag{5}$$

To determine the Lagrange equation of the system, first needs to find out its total kinetic and potential energy, and their difference is defined as Lagrange dynamics [9]. The total potential energy of the system (PT) is the summation of the total potential energy of the elastic of the system, link 1, link 2, and hinge. The total kinetic energy of the system (KT) is the summation of kinetic energy due to the rotation of the arm due to the rotational and translation motion of the pendulum's first and second hinge, as shown in equation (6).

$$\left.\begin{aligned} P_T &= P_{elastic} + P_1 + P_2 + P_{hinge} \\ K_T &- K_{arm} + K_{rot1} + K_{rot2} + K_{trans1} + K_{trans2} \\ L &= P_T - K_T \end{aligned}\right\} \tag{6}$$

After computing the Lagrange equation and generalized force coordinate, these values are put in equation (3) to determine the non-linear equation of motion for the system by computing all derivative parameters. To linearize, the non-linear equation uses a linearized function, as mentioned in equation (7).

$$\left.\begin{aligned} f_{line.} = f(z_o) &+ \left(\frac{\partial^2 f(z)}{\partial z_1}\right)_{z=z_o} (z_1 - a) \\ &+ \lim_{x\to\infty}\left(\frac{\partial^2 f(z)}{\partial z_2}\right)_{z=z_o} (z_2 - b) \end{aligned}\right\} \tag{7}$$

Linear State-space equations are mentioned in equation (9), where x is the state, and u is the control input [10]. The state space variable chosen should at least contain generalized coordinates, and their first derivative, as given in equation (8), y, is the system's output.

$$\begin{aligned} x^T &= \begin{bmatrix} \theta(t) & \alpha(t) & \phi(t) & \dot{\theta}(t) & \dot{\alpha}(t) & \dot{\phi}(t) \end{bmatrix} \\ y^T &= \begin{bmatrix} \theta(t) & \alpha(t) & \phi(t) \end{bmatrix} \end{aligned} \tag{8}$$

$$\sum := \begin{cases} \dfrac{dx}{dt} = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases} \Leftrightarrow \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{9}$$

After linearizing the nonlinear equation and solving acceleration terms, the state space matrix of the rotary double inverted pendulum was obtained, as presented below.

$$A = {}^{1}\!/_{F_t} \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & a_{42} & a_{43} & 0 & 0 & 0 \\ 0 & a_{52} & a_{53} & 0 & 0 & 0 \\ 0 & a_{62} & a_{63} & 0 & 0 & 0 \end{bmatrix} \tag{10}$$

where, $$a_{42} = L_r g(m_1^2 l_1^2 + 2m_1^2 m_h^4 L_1^5 m_2^2 l_1^3) \tag{11}$$

$$a_{43} = -m_1 l_1 L_r m_2 g(-l_1 + L_1) \tag{12}$$

$$\left.\begin{aligned} a_{52} = g\{&L_r^2 L_1 m_2 (m_1 + m_h) + L_1 L_r^2 (m_h^2 + m_1^2) \\ &+ J_r L_1 (m_1 + m_h) + 2L_1 L_r^2 m_h m_1\} \end{aligned}\right\} \tag{13}$$

$$a_{53} = -m_2 g\{L_1 L_r^2 (m_2 - m_1) + J_r L_1\} \tag{14}$$

$$a_{62} = -{}^{g}\!/_{L_2} \begin{Bmatrix} L_1^2 L_r^2 (m_1^2 - m_1 m_2) + L_r^2 L_1 L_2 m_2 (m_1 + m_2) \\ + J_r L_1 L_2 (2m_1 + m_h) + 2L_r^2 L_1 L_2 m_h m_1 \\ + L_r^2 L_1 L_2 (m_1^2 + m_h^2) + L_r^2 L_1 (m_1 m_2 - m_1^2) \end{Bmatrix} \tag{15}$$

$$\begin{aligned} a_{63} = {}^{g}\!/_{L_2} \{&J_r m_2 L_1 (L_2 + L_1) - 2L_1^2 L_r^2 m_1 (m_2 + m_h) \\ &+ L_1^2 J_r (m_1 + m_h)\} \end{aligned} \tag{16}$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{17}$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:24:17 UTC from IEEE Xplore. Restrictions apply.

$$B = \frac{1}{F_t}\begin{bmatrix} 0 \\ 0 \\ 0 \\ L_1^2(m_h + m_1) \\ L_r L_1(m_h + m_1) \\ -\frac{L_r}{L_2}\{L_2 L_1(m_h + m_1)\} \end{bmatrix} \quad (18)$$

The parameter values are given in Table I to determine the system's numerical matrix.

TABLE I. SYSTEM PARAMETER VALUES

Symbol	Values
Masses (kilogram 'kg')	
m_1	0.097
m_2	0.127
m_h	0.1410
Length (meter 'm')	
L_1	0.2
l_{m1}	0.1635
L_2	0.3365
l_{m2}	0.1778
L_r	0.2159
Viscous Damping (N-m-s/rad)	
B_1	0.0024
B_2	0.0024
Moment of Inertia (kg-m^2)	
J_r	9.98* 10^-4
J_1	0.0041
J_2	0.0053

III. CONTROLLER

The system's stability is checked using the state space characteristic equation, calculated as the determinant of (s*I-A), where I is the identity matrix and plots the system's pole-zero and Nyquist plot, as mentioned in Fig. 1 and 2. The open loop eigenvalues are [0 -29.52 12.27 7.23 -10.19 -2.14]. The eigenvalues show that two poles lie on the right-hand side, and the system's step response diverges from the Nyquist plot. It is also clear that the system is not stable. Hence, the system is highly unstable.

To stabilize the system, firstly shift all the poles on the right-hand side using the pole placement technique, and after that, use the optimal optimization algorithm to find the gain of the system and apply a feedback loop by implementing the integral term to control or stabilize the pendulum in a vertically upward direction with minimum deflection, which means using a LQR feedback controller with a small integral mode shows in Fig. 3.

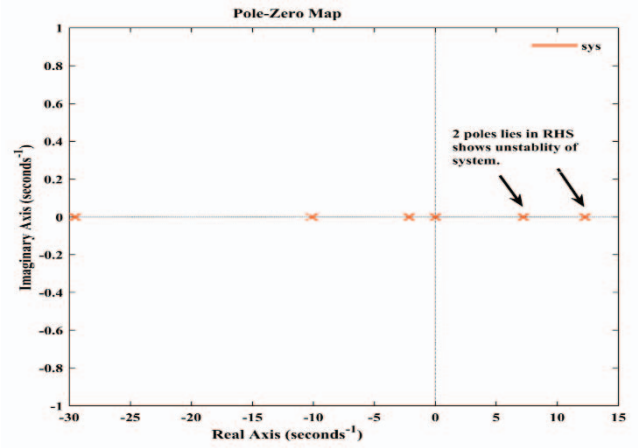


Fig. 1. Pole- Zero Map

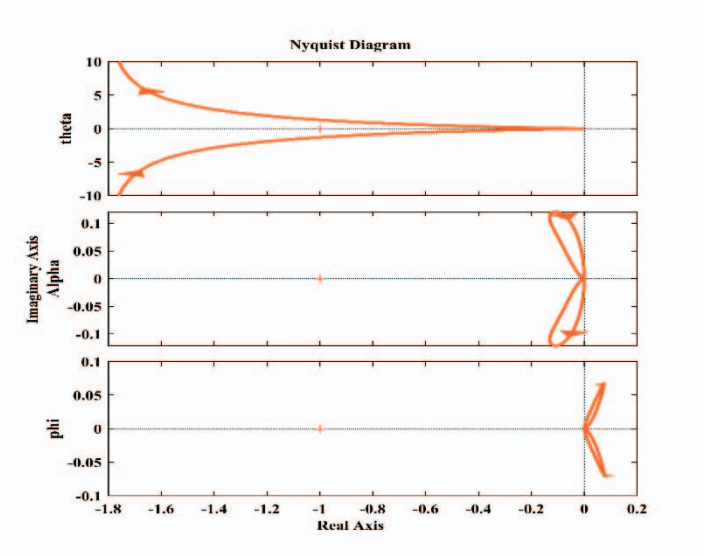


Fig. 2. Nyquist plot of the system

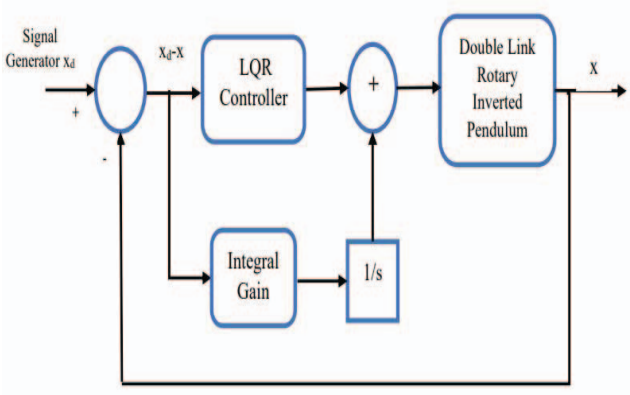


Fig. 3. Block Diagram of Controller Implementation

In the LQR controller, mathematical algorithm-controlled feedback gain is used to design optimal control, enabling the closed-loop system to be stable and high-performance. It uses a weighting parameter instead of eigenvalue placement. It minimizes the cost function that maps the variable to a real number, as mentioned in equation (19), which uses all state variables. Q and R denote positive definite weighting matrix. Equation (20) represents the matrix algebraic Riccatti equation

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:24:17 UTC from IEEE Xplore. Restrictions apply.

whose solution 'V' is required to determine the optimal feedback gain 'k'.

$$E = \int_0^{\infty} x(t)^T Q * x(t) + u(t)^T R * u(t) d(t) \tag{19}$$

$$A^T V + VA + BR^{-1}B^T V + Q = 0 \tag{20}$$

$$u = -\sum k_j (x_d - x_j);\ k = R^{-1}B^T A^T V \tag{21}$$

The feedback loop stabilizes the servo to the desired position (θ_d) while minimizing the pendulum deflection. The reference state is defined as:

$$x_d = \theta_d \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad \text{and} \, u = k(x_d - x) \tag{22}$$

After introducing the integral term, the overall system state-variable and reference state are mentioned in equation (23).

$$\rho^T = \left[\theta(t) \quad \alpha(t) \quad \phi(t) \quad \dot{\theta}(t) \quad \dot{\alpha}(t) \quad \dot{\phi}(t) \int \theta(t)\right]$$
$$\rho_d = \left[\theta_d(t) \quad 0 \quad 0 \quad 0 \quad 0\,0 \quad \int \theta(t)\right] \tag{23}$$

IV. Result and Discussion

This section analyzes the response of the DL-RIP system using an introduced optimal integral controller by taking different signals as reference paths like square, sinusoidal, and saw-tooth. After tunning the weighting matrix of the LQR controller, taking R= 20 and Q is equal to diag[5 1 1 5 12 0 0.9], set Q(7,7) = 0.9 to introduce a small integral gain to track servo error, with the help of MATLAB code obtains the gain value k of the controller shows in equation (24) and closed-loop eigenvalue of the system is given below in equation (25).

$$k = \begin{bmatrix} 0.12 & -16.47 & -37.68 & 0.437 & -12.54 & -4.43 & 0.127 \end{bmatrix} \tag{24}$$

$$E = [-0.4221+0.38i,\ -0.4221-0.38i,\ -3.348,\ -6.3040,\ -10+3.055i,\ -10-3.055i,\ -50] \tag{25}$$

Taking the amplitude of gain block 30 to generate a step with an amplitude of 30 degrees. The pendulum maintains its balance about the vertical position (zero angle) while tracking the 30-degree servo angle, as shown in response to the rotary arm system with setpoint square, sin, and sawtooth wave.

In the case of a square setpoint, the arm maintains its balance at 47. 48 degrees in 11 sec with an overshoot of 13 percent, but with a sawtooth wave, it maintained balance at 37.67 degrees in 12 sec with an 11 percent overshoot, less than a square case. In both cases, the discontinuity occurs in the path, and at the discontinuity point, it is disturbed with their track and again achieved as required, shown in Fig. 4 and 5. When tested with sinusoidal setpoint, it maintained balance at 32.57 degrees in 13 sec with 1.89 % overshoot only; hence, when no discontinuity occurs in the desired path, less overshot occurs during the stabilization of the system, and also, less error occurs, as in this case 2.57 shows in Fig. 6. Also compare this LQR-I controller response with existing LQR, the response with LQR-I controller is better than existing one as shown in Fig.6. The overshoot and settling time less.

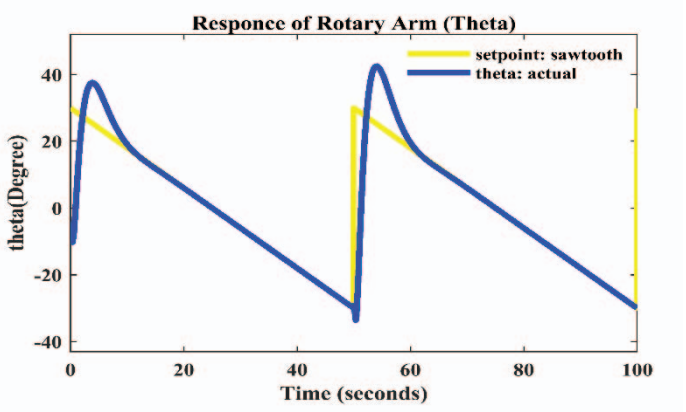


Fig. 4. Rotary Arm Response with Sawtooth Setpoint

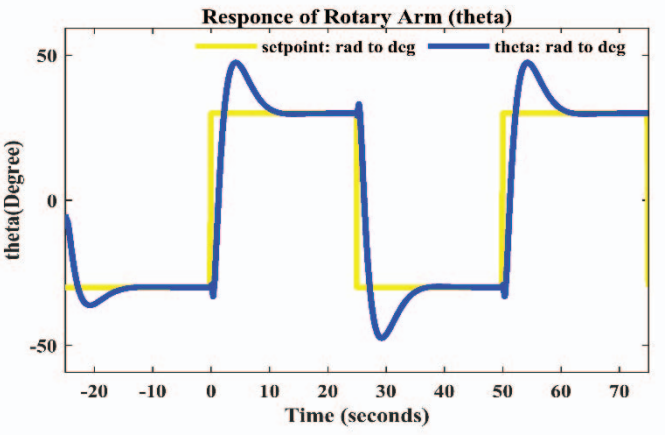


Fig. 5. Rotary Arm Response with Square Setpoint

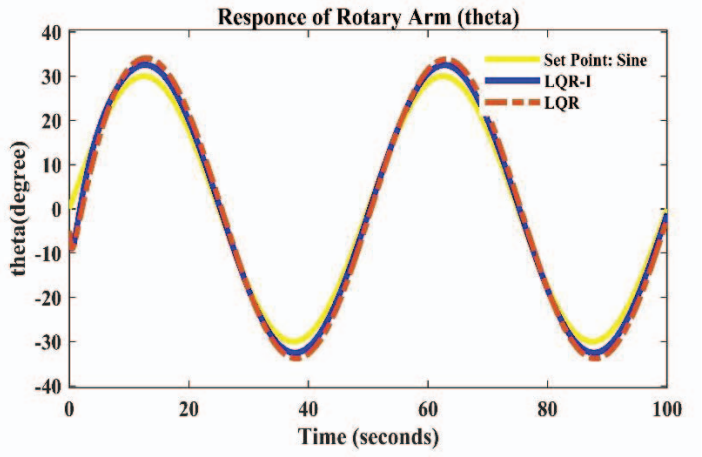


Fig. 6. Rotary Arm Response with Sinusoidal Setpoint

In all three cases, the first and second pendulum stabilization behavior in the vertical upward direction is presented in Fig. 7 and 8. In the sinusoidal reference case, the pendulum stabilizes with less overshoot and less error but takes more time than the other two cases. Also, the system tracks the desired path in all three cases, but with sinusoidal, small deflection occurs with the desired path, and less disturbance occurs.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:24:17 UTC from IEEE Xplore. Restrictions apply.

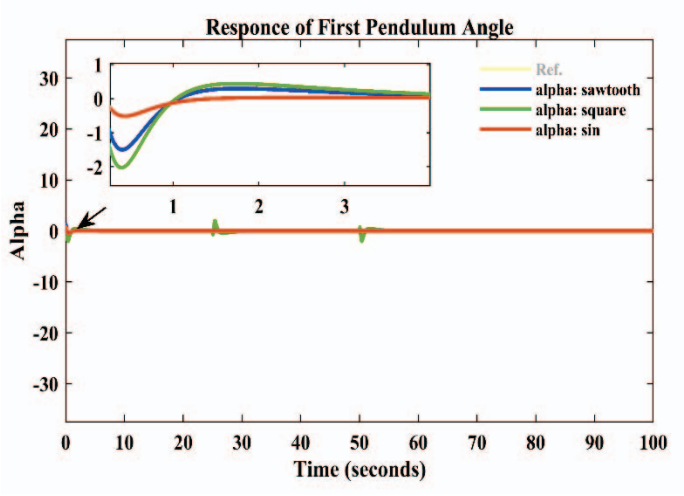


Fig. 7. Response of first pendulum angle in three cases

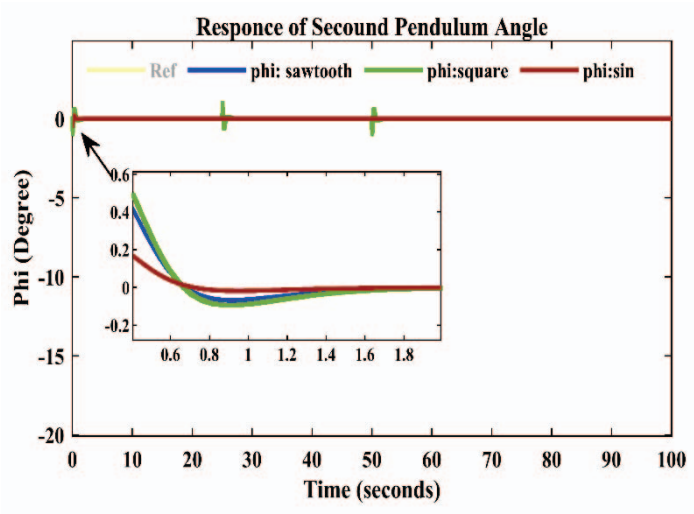


Fig. 8. Response of second pendulum angle in three cases

Tables II, III, and IV mention the pendulums' and rotary arms' time characteristic response.

TABLE II. TIME CHARACTERISTICS RESPONSE OF ROTARY ARM

Theta	Settling time (sec)	Rise time (sec)	Peak Angle Deflection	Percentage overshoot
Square	11	0.84	47.48	13
Sine	13	0.38	32.57	1.89
LQR Sine	14.05	0.45	33.52	4
Sawtooth	12	0.379	37.60	11

TABLE III. TIME CHARACTERISTICS RESPONSE OF IST PENDULUM

Alpha	Settling time (sec)	Rise time (sec)	Peak	Percentage overshoot
Square	0.97	2.02	2.5	2
Sine	4.58	0.192	0.42	1.57
Sawtooth	2.68	0.0612	1.215	1.94

TABLE IV. TIME CHARACTERISTICS RESPONSE OF IIND PENDULUM

phi	Settling time (sec)	Rise time (sec)	Peak	Percentage overshoot
Square	1.5	0.035	0.97	4
Sine	1.29	0.0937	0.838	1.38
Sawtooth	1	0.0428	1.93	1.96

V. CONCLUSION

The proposed DL-RIP system demonstrates stable vertical upward movement when utilizing the LQR-I controller. The analysis of three distinct input cases—sinusoidal, square, and sawtooth waveforms—shows that the system achieves superior tracking and performance with sinusoidal inputs, delivering a smoother and more accurate response. Furthermore, the LQR-I controller outperforms the standard LQR controller, as evidenced by its reduced overshoot of 1.89% and a faster rise time of 0.38 seconds. These results highlight the effectiveness of the LQR-I controller in improving stability and dynamic response for the DL-RIP system, making it a robust solution for applications requiring precise control and minimal oscillations.

REFERENCES

[1] A. Pandey, and M. Adhyaru, "Robust control design for rotary inverted pendulum with unmatched uncertainty," International Journal of Dynamics and Control, vol. 11.3, pp. 1166-1177, 2023.

[2] A. Muhammad et al., "Swing up and stabilization control of rotary inverted pendulum based on energy balance, fuzzy logic, and LQR controllers," Measurement and Control, vol. 54.9-10, pp. 1356-1370, 2023.

[3] S. Sondarangallage, and M. Parnichkun, "Control of rotary double inverted pendulum system using LQR sliding surface based sliding mode controller," Journal of Control and Decision, vol. 9.1, pp. 89-101, 2022.

[4] T. Wiput, and P. Ruangurai, "Design, modeling and parameters identification of rotary-type double inverted pendulum," 10th International Congress on Advanced Applied Informatics (IIAI-AAI), IEEE, 2021.

[5] Kasera, Shrey, A. Kumar, and L.B. Prasad, "Trajectory tracking of 3-DOF industrial robot manipulator by sliding mode control," 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON), IEEE, 2017.

[6] Hazem, Z. Ben, M.J. Fotuhi, and Z. Bingül, "Development of a Fuzzy-LQR and Fuzzy-LQG stability control for a double link rotary inverted pendulum," Journal of the Franklin Institute, vol. 357.15, pp. 10529-10556, 2021.

[7] T. Van-Si, C.H Nguyen, T. Duong, H. Nam Le, and M.T. Vo, "Comparative Study of Linear Quadratic Gaussian and Linear Quadratic Regulator Controllers for Rotary Double Inverted Pendulum in Parallel Type," 7th International Conference on Green Technology and Sustainable Development (GTSD), pp. 407-412, IEEE, 2024.

[8] J. Slávka, and J. Sarnovský, "Modelling of classical and rotary inverted pendulum systems–a generalized approach," Journal of Electrical Engineering, vol. 64.1, pp. 12-19, 2013.

[9] N. Arulmozhi, and A. A Victoire, "Optimization reinforced PID-sliding mode controller for rotary inverted pendulum," IEEE Access, vol. 11, pp. 24420-24430, 2023.

[10] H. Ruben, R.G Hernandez, and F. Jurado, "Modeling, Simulation, and Control of a Rotary Inverted Pendulum: A Reinforcement Learning-Based Control Approach," Modelling 5, vol. 4, pp. 1824-1852, 2024.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:24:17 UTC from IEEE Xplore. Restrictions apply.

Stuttering Detection Using LSTM and LSTM-Attention based Convolutional Neural Network

Monalisa Maity
Department of Applied Mathematics
Delhi Technological University
monalisamaity_mc21a12_31@dtu.ac.in

Ishita Gupta
Department of Applied Mathematics
Delhi Technological University
ishitagupta_mc21a12_01@dtu.ac.in

Dinesh Kumar Vishwakarma
Department of Information Technology
Delhi Technological University
dinesh@dtu.ac.in

2025 10th International Conference on Signal Processing and Communication (ICSC) | 979-8-3315-1189-0/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICSC64553.2025.10968952

***Abstract*— Stuttering is a common speech disorder which results in involuntary repetitions, prolongations, blocks or interjections while speaking. Detection and evaluation using traditional methods is time-consuming. Therefore, we conduct a study on 2 classification based deep learning models for stuttering detection: 1) a model which comprises of a Convolutional Neural Network (CNN) followed by Long-Short Term Memory (LSTM) and Attention layers 2) a model which comprises of a CNN followed by LSTM layer only. We used the SEP-28K dataset for training and evaluating our models. These models take mel spectrogram as input and perform binary classification for each disfluency label of the SEP-28K dataset. The CNN + LSTM + Attention model achieved a higher accuracy of 70.47% in detecting prolongation, 60.80% in block and 56.90% in word repetition compared to the CNN + LSTM model, which achieved 69.43%, 60.34% and 55.23% for the same labels, respectively. However, the CNN + LSTM model outdid the attention-based model in classifying fluent/no-stutter instances with 66.52% accuracy against 61.50% and sound repetition with 65.22% accuracy against 64.25%. For interjection, both the models had similar performance in terms of accuracy. For F1 scores, the CNN + LSTM + Attention model showed improvement in block, prolongation, sound and word repetition classifications, while the CNN + LSTM model observed better results for fluent/no-stutter and interjection detection.**

Keywords—Stuttering, CNN, LSTM, Attention, mel spectrogram, SEP-28K

I. Introduction

Stuttering causes speech disfluency which affects the day-to-day lives of people suffering from it. People who stutter are prone to anxiety, low self-esteem and several other mental health issues because they often feel shame when they are not able to speak fluently **[1]**. The most common symptoms that indicate stuttering include: Prolongation, Repetition, Block and Interjection. Prolongation is a type of stuttering in which a person holds out a sound for too long. Repetition involves repeating a sound or word multiple times during speech. Blocks are gaps or breaks in speech. Interjections are usually the filler-words used by the people who stutter to hide stuttering like 'umm', 'uh' and 'well'. There can be multiple causes of stuttering. Developmental stuttering usually begins in early childhood, starting with episodic stuttering which may progress to blocks and prolongations **[2]**. Acquired stuttering consists of sound repetition without typical anxiety or secondary behaviour. Stuttering can also be triggered by traumatic events **[3]** and Neurogenic stuttering is usually caused by brain injuries and diseases **[4]**.

Thus, the diagnosis of stuttering is important as it can help in detecting underlying diseases as well. Speech-language pathologists (SLP) are the healthcare professionals who diagnose and treat patients that suffer from stuttering. However, stuttering detection is a time-consuming and laborious task because SLPs need to interact with patients in different situations to evaluate their speech and behaviour. Therefore, we conduct a study on LSTM and LSTM-Attention based CNN models for stuttering detection. Our models were trained on the SEP-28K **[5]** dataset.

Table 1: Types of Disfluencies

Disfluency Type	*Example*
Prolongation	*I eat cheeeeeerry.*
Blocks	*I…… eat cherry…*
Interjections	*I eat umm cherry.*
Sound Repetition	*I e-e-e-eat cherry.*
Word Repetition	*I eat eat eat cherry.*

Our key contributions are as follows: 1) Implementing CNN + LSTM and CNN + LSTM + Attention based models for stuttering detection, to classify whether a person is stuttering or not in an audio clip. 2) We used the SEP-28K dataset for conducting our experiments- trained and evaluated our models for different stuttering types and made comparisons between both models.

The rest of the paper is organized as follows: In Section-2, we have discussed about the work that has been done on Stuttering Detection. In Section-3, we describe the SEP-28K dataset and in Section-4, we explain the architecture of the LSTM-attention model. In Section-5, we present the experimental results and Section-6, concludes the work.

II. Related Work

Different deep learning methods have been proposed in the literature for stuttering detection **[6]**.

Kourkounakis et al. **[7]** proposed a model FluentNet that consists of a Squeeze-and-Excitation Residual convolutional neural network which facilitates the learning of strong spectral frame-level representations, followed by a set of bidirectional long short-term memory layers that aid in learning effective temporal relationships. An attention layer has been used in the model to identify the important portions of the speech to

979-8-3315-1189-0/25/$31.00 ©2025 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:24:43 UTC from IEEE Xplore. Restrictions apply.

enhance performance. It was implemented on the UCLASS**[8]** dataset as well as LibriStutter: a disfluency dataset based on the public LibriSpeech dataset with synthesized stutters.

Lea et al. **[5]** prepared a large stuttering dataset named SEP−28k and used the ConvLSTM (Convolutional LSTM) model to detect various types of stuttering. On SEP-28K, the model performed poorly for blocks and word repetitions as compared to other disfluency labels.

Al-Banna et al. **[9]** also proposed a model for stuttering detection to evaluate stuttering severity. The model accepted log mel spectrogram and used a 2D atrous convolutional network to learn speech spectral and temporal features. They evaluated the performance of the model on UCLASS and FluencyBank **[10]** datasets. According to their experimental results the model performed better than pre-existing methods for prolongation in terms of F1 score on the FluencyBank and UCLASS datasets.

Hamdi et al**. [13]** proposed an augmentation pipeline which was applied on the pre-filtered data and used pitch-shifting technique to augment the raw signal as well as spectral data augmentation technique to augment the computed mel-spectrograms.

III. DATA

In this paper we reference the dataset SEP-28K which is a publicly available stuttering events in podcasts dataset prepared by Lea et al**. [5]** from a set of 385 podcasts. From each podcast, 3 seconds long segments are chosen which gives us an aggregate of 28,177 segments.

SEP-28K consists of stuttering clips labelled according to the following categories: disfluent and non-disfluent. In this paper we concern ourselves only with the disfluent labels of this dataset. These labels are prolongations, repetitions, blocks, interjections, fluent. Each label has values 0, 1, 2 or 3 for the number of listeners (out of 3) who agreed with the respective label.

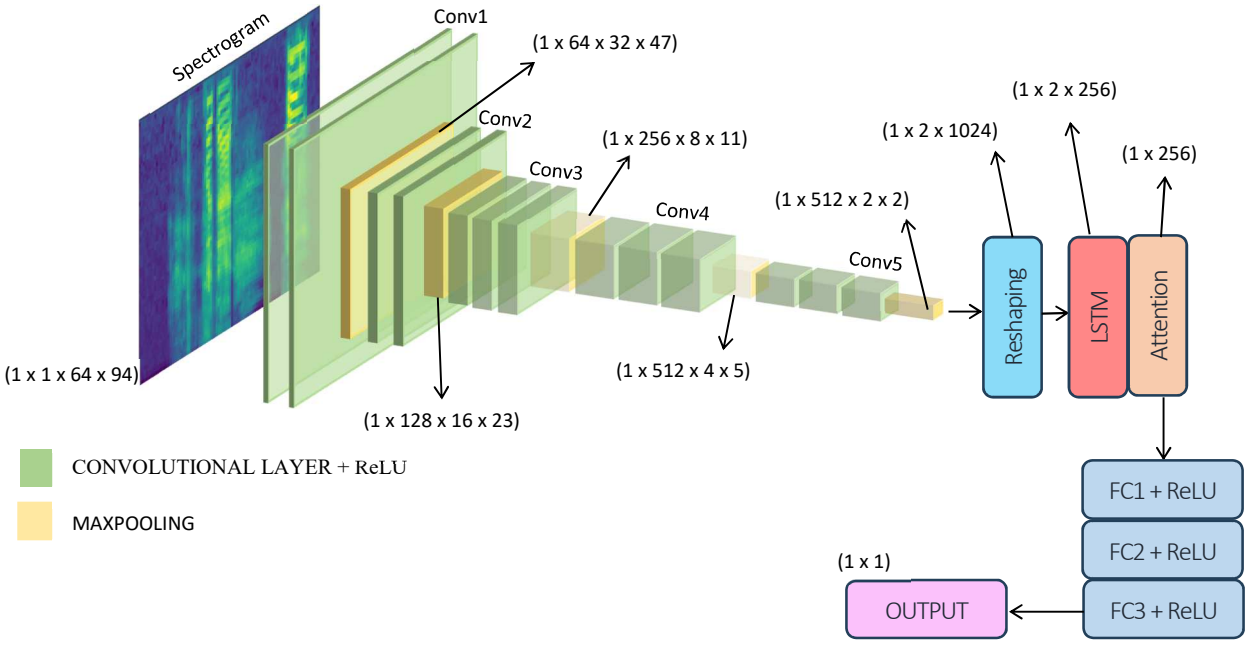


Fig 1: Model architecture- CNN+ LSTM+ Attention

Sheikh et al**. [11]** addressed the class imbalance problem in stuttering detection via a multi-branching scheme and by weighting the contribution of classes in the overall loss function, and conducted their experiments on the SEP-28K dataset over the baseline (StutterNet **[12]**). They used data augmentation to deal with data scarcity. Training the model with augmented data gave better results than the multi-branching StutterNet in terms of macro F1-score. They also proposed a new model, a multi-contextual (MC) StutterNet. It captured multiple contexts of the speech clips. This model had a better performance than the single context based multi-branching model for the dataset.

An LSTM-Attention based model similar to the one we use for our experiments, has been used for Covid-19 detection.

We have chosen samples from shows 'HeStutters', 'IStutterSowhat', 'MyStutteringLife', 'StutterTalk' and 'WomenWhoStutter'. This reduces the total samples in our experiment to 21,120. We randomly select 80% of these samples to train the model whilst the remaining 20% are reserved for testing the model. We have also converted the label values into binary i.e. 0 and 1 by setting values equal to and greater than 1 as 1, and the values less than 1 as 0. This allows us ease of classification into true and false categories.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:24:43 UTC from IEEE Xplore. Restrictions apply.

IV. Model Architecture.

The LSTM-Attention model is a hybrid architecture designed to integrate CNN for feature extraction, the LSTM component for capturing the temporal features and the Attention layer to improve performance by identifying the most important features. The CNN component is based on the VGG-16 **[14]** architecture, consisting of 5 blocks of convolutional layers. Each mel spectrogram of the dimensions 64 x 94 is passed through the CNN layers and then the output is reshaped into a tensor of dimensions 2 x 1024. The reshaped output obtained from the CNN layers is then passed through 2 LSTM layers each consisting of 256 units. An attention mechanism **[13]** is then applied to the LSTM output. The attention layer includes a linear layer with 256 units, a Tanh activation, and a Softmax function. The attention output is element-wise multiplied with the LSTM output and summed. The output of the attention layer is then passed through 3 fully-connected layers with 4096 units, 4096 units and 1unit respectively. Similarly, we have the LSTM-only model, but the attention layer is removed.

Table 2: Experimental Results – Accuracy(%)

Model	Disfluency / Stuttering Label of SEP-28K					
	Block	Prolongation	Sound Repetition	Word Repetition	Interjection	Fluent/ No stutter
VGG-16+LSTM+Attention	**60.80**	**70.47**	64.25	**56.90**	72.11	61.50
VGG-16+LSTM	60.34	69.43	**65.22**	55.23	72.11	**66.52**

Table 3: Experimental Results - F1 Score(%)

Model	Disfluency / Stuttering Label of SEP-28K					
	Block	Prolongation	Sound Repetition	Word Repetition	Interjection	Fluent/ No stutter
VGG-16+LSTM+Attention	**57.87**	**61.07**	**41.69**	**33.67**	61.55	70.77
VGG-16+LSTM	56.53	58.01	41.40	31.11	**61.82**	**77.49**

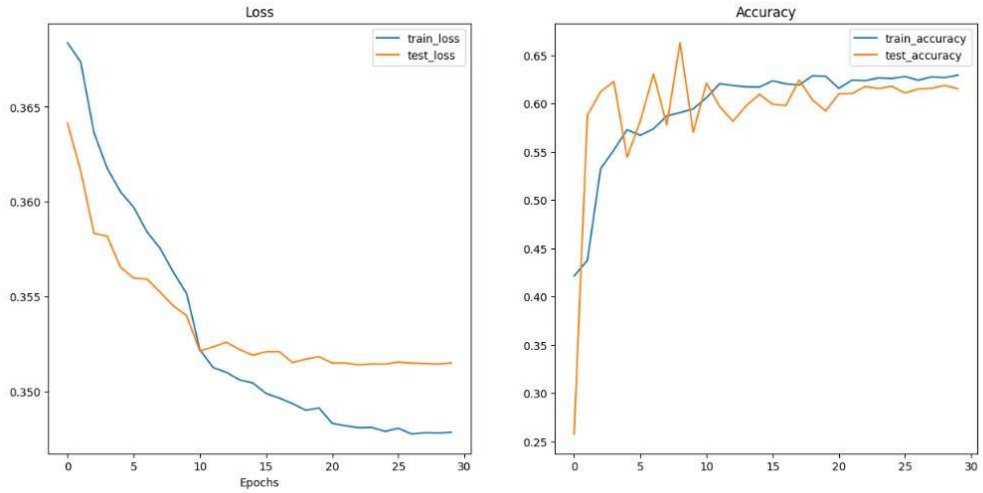


Fig 2: Training and validation loss/accuracy for CNN-LSTM-Attention on fluent-label data

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:24:43 UTC from IEEE Xplore. Restrictions apply.

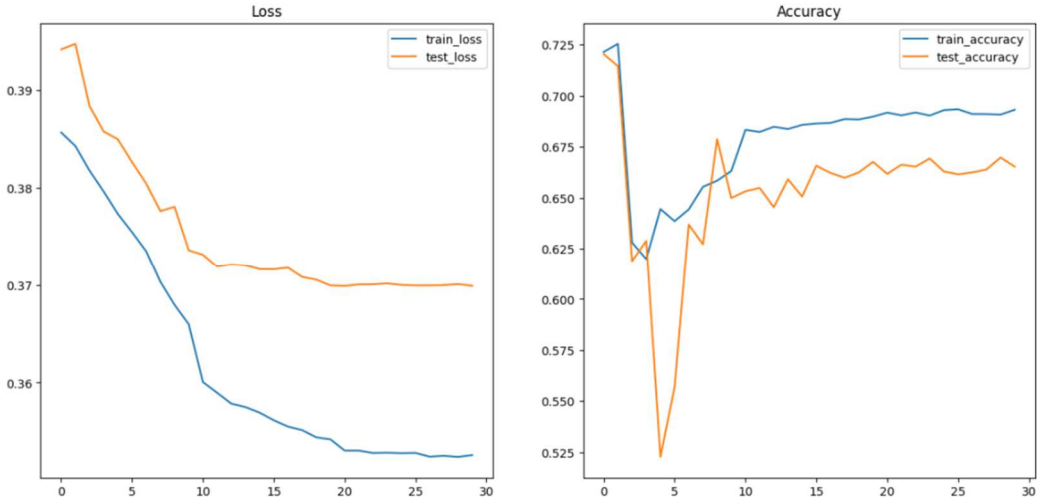


Fig 3: Training and validation loss/accuracy for CNN - LSTM on fluent-label data

V. Experimental Results

A. Setup

We used PyTorch for processing and converting the audio clips into mel spectrograms, as well as building the models. The dataset was highly imbalanced, so we used a weighted loss function. The weight for the positive class was calculated by dividing the number of negative samples by positive samples. We used the Adam optimizer set at a learning rate of 0.00001 and weighted binary cross-entropy loss to train the models for 30 epochs with each batch of size 32 samples.

B. Evaluation and Results

To evaluate the models' performance, we used two metrics: accuracy and F1-score. We used 80% of the data for training and the remaining 20% for testing the models.

We trained and evaluated our models using the SEP-28K dataset. We carried out two experiments: the first with a CNN + LSTM model, and the second with a CNN + LSTM + Attention model. This was done for each disfluency (prolongations, sound repetitions, word repetition, blocks, interjections, and fluent) separately. For each disfluency, the models performed binary classification. **Tables 2, 3** show the results of the experiments. **Figures 2, 3** show the training and validation losses and accuracies for our 2 models for the fluent/no-stutter label.

The CNN + LSTM + Attention model performed better in terms of accuracy for word repetition, prolongation, and block compared to the CNN+LSTM model. While the CNN+LSTM model performed better for sound repetition and no stutter labels. For the interjection label both models have similar accuracy.

VI. Conclusions

This paper presents a study on stuttering detection using an LSTM-Attention based hybrid CNN model and an LSTM based hybrid CNN model. Our models perform binary classification by taking a mel spectrogram as an input to identify whether a person has one of the disfluencies (Prolongation, Block, Interjection and Repetition). The CNN component of our models extracts the relevant features, the LSTM component captures the temporal features and the Attention layer improves the performance by identifying the most important features.

We compared the results of two experimental results, one with the CNN - LSTM model and the other with the CNN-LSTM-Attention model. We used the SEP-28K dataset for all our experiments. In order to deal with the imbalanced dataset, we used a weighted loss function. The CNN - LSTM - Attention model had better performance in terms of accuracy for Word repetition, Prolongation, and Block detection as compared to the CNN - LSTM model. While the CNN - LSTM model performed better for sound repetition and no stutter labels.

In our future work we will be exploring different model architectures similar to **[15][16][17]** and evaluate their performance against our baseline models.

References

[1] U. Singh and S. S. Kumar, 'Perception of Stuttering in Individuals With Stuttering', *Indian J Otolaryngol Head Neck Surg*, vol. 74, no. S3, pp. 4880–4890, Dec. 2022, doi: 10.1007/s12070-020-02344-4.

[2] N. Gordon, 'Stuttering: incidence and causes', *Developmental medicine and child neurology*, vol. 44, no. 4, pp. 278–282, 2002.

[3] D. Ward, *Stuttering and cluttering: frameworks for understanding and treatment*. Psychology Press, 2017. Accessed: Jun. 14, 2024. [Online]. Available: https://www.taylorfrancis.com/books/mono/10.4324/9781315727073/stuttering-cluttering-second-edition-david-ward

[4] C. Cruz, H. Amorim, G. Beça, and R. Nunes, 'Neurogenic stuttering: A review of the literature', 2018, Accessed: Jun. 14, 2024. [Online]. Available: https://www.researchgate.net/profile/Renato-Nunes-4/publication/322527188_Neurogenic_stuttering_A_review_of_the_liter

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:24:43 UTC from IEEE Xplore. Restrictions apply.

ature/links/5d78e29d92851cacdb319032/Neurogenic-stuttering-A-review-of-the-literature.pdf

[5] C. Lea, V. Mitra, A. Joshi, S. Kajarekar, and J. P. Bigham, 'Sep-28k: A dataset for stuttering event detection from podcasts with people who stutter', in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2021, pp. 6798–6802. Accessed: Jun. 14, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9413520/

[6] S. A. Sheikh, M. Sahidullah, F. Hirsch, and S. Ouni, 'Machine learning for stuttering identification: Review, challenges and future directions', *Neurocomputing*, vol. 514, pp. 385–402, 2022.

[7] T. Kourkounakis, A. Hajavi, and A. Etemad, 'FluentNet: End-to-End Detection of Speech Disfluency with Deep Learning'. arXiv, Sep. 23, 2020. Accessed: Jun. 14, 2024. [Online]. Available: http://arxiv.org/abs/2009.11394

[8] P. Howell, S. Davis, and J. Bartrip, 'The UCLASS archive of stuttered speech.', *Journal of speech, language, and hearing research : JSLHR*, Mar. 2009, doi: 10.1044/1092-4388(2009/07-0129).

[9] A.-K. Al-Banna, E. Edirisinghe, and H. Fang, 'Stuttering detection using atrous convolutional neural networks', in *2022 13th International Conference on Information and Communication Systems (ICICS)*, IEEE, 2022, pp. 252–256. Accessed: Jun. 14, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9811183/

[10] N. B. Ratner and B. MacWhinney, 'Fluency Bank: a new resource for fluency research and practice', *Journal of fluency disorders*, vol. 56, p. 69, Jun. 2018, doi: 10.1016/j.jfludis.2018.03.002.

[11] S. A. Sheikh, M. Sahidullah, F. Hirsch, and S. Ouni, 'Advancing stuttering detection via data augmentation, class-balanced loss and multi-contextual deep learning', *IEEE Journal of Biomedical and Health Informatics*, 2023, Accessed: Jun. 14, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10050738/

[12] S. A. Sheikh, M. Sahidullah, F. Hirsch, and S. Ouni, 'Stutternet: Stuttering detection using time delay neural network', in *2021 29th European Signal Processing Conference (EUSIPCO)*, IEEE, 2021, pp. 426–430. Accessed: Jun. 14, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9616063/

[13] S. Hamdi, M. Oussalah, A. Moussaoui, and M. Saidi, 'Attention-based hybrid CNN-LSTM and spectral data augmentation for COVID-19 diagnosis from cough sound', *J Intell Inf Syst*, vol. 59, no. 2, pp. 367–389, Oct. 2022, doi: 10.1007/s10844-022-00707-7.

[14] K. Simonyan and A. Zisserman, 'Very Deep Convolutional Networks for Large-Scale Image Recognition'. arXiv, Apr. 10, 2015. doi: 10.48550/arXiv.1409.1556.

[15] A. Pandey and D. K. Vishwakarma, 'Attention-based Model for Multi-modal sentiment recognition using Text-Image Pairs', in *2023 4th International Conference on Innovative Trends in Information Technology (ICITIIT)*, IEEE, 2023, pp. 1–5. Accessed: Jun. 05, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10068626/

[16] A. Pandey and D. K. Vishwakarma, 'Progress, achievements, and challenges in multimodal sentiment analysis using deep learning: A survey', *Applied Soft Computing*, vol. 152, p. 111206, Feb. 2024, doi: 10.1016/j.asoc.2023.111206.

[17] A. Pandey and D. K. Vishwakarma, 'VABDC-Net: A framework for Visual-Caption Sentiment Recognition via spatio-depth visual attention and bi-directional caption processing', *Knowledge-Based Systems*, vol. 269, p. 110515, 2023.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on May 01,2025 at 07:24:43 UTC from IEEE Xplore. Restrictions apply.

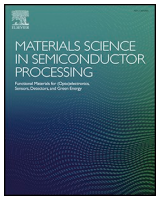
Materials Science in Semiconductor Processing 195 (2025) 109607



Contents lists available at ScienceDirect

Materials Science in Semiconductor Processing

journal homepage: www.elsevier.com/locate/mssp



Substantial enhancement of the electroactive phase and dielectric properties at low loading levels of SnO_2/MXene nanocomposite as a novel nanofiller in the PVDF matrix

Gagan Sharma, Deshraj Meena *

Clean Energy Research Laboratory, Department of Applied Physics, Delhi Technological University, Delhi, 110042, Delhi, India

ARTICLE INFO

Keywords:
SNMX nanocomposite
Free standing films
Piezoelectric flexible nanogenerators
Dielectric properties

ABSTRACT

Herein, we investigate the impact on electroactive phase and dielectric properties of polyvinylidene fluoride (PVDF) on addition of hydrothermally synthesized SnO_2/MXene (SNMX) nanocomposite with different concentration (0.25 wt%, 0.50 wt%, 0.75 wt%) as nanofillers. The pure and SNMX/PVDF composite films were fabricated using solution casting method. Morphological and structural analyses revealed a modified spherulitic structure, indicating a transformation from the non-polar phase (52 %) to the polar ($\beta + \gamma$) phase (86 %). The dielectric measurements were conducted over a frequency range of 20 Hz to 1 MHz at different operating temperature to examine the various relaxations occurring in the polymer. At 25 °C and an applied field frequency of 20 Hz, the relative dielectric permittivity of the SNMX/PVDF composite film (0.75 wt%) elevates approximately 25-fold relative to the pristine PVDF film, with a concomitant reduction in loss tangent. Further, the relative dielectric permittivity of all films intensifies with rising operating temperature. At 100 °C, the relative dielectric permittivity amplifies from 165 to 250 compared to room temperature for the 0.75 wt% nanofiller concentration. The increment in dielectric properties with addition of SNMX are attributed to a) the enhancement of the electroactive phase, which favours molecular reorientation, b) the presence of SnO_2 semiconductive layer on the surface of MXene flakes that prevent formation of conductive pathways, thus improves the interfacial polarization through the Maxwell–Wagner–Sillars (MWS) effect in addition to dipolar polarization. A piezoelectric flexible nanogenerator (PENG) fabricated from the 0.75 wt% SNMX/PVDF nanocomposite film generated a maximum peak to peak voltage is 105 V.

1. Introduction

In recent years, flexible dielectric materials have garnered significant interest due to their potential applications in the rapidly developing field of flexible nanoelectronics. In this context, dielectric polymers have attracted considerable attention because of their desirable physical and chemical properties [1–3]. However, polymers generally exhibit lower dielectric constants compared to inorganic materials. For instance, polyvinylidene fluoride (PVDF), a prototypical dielectric polymer, has a dielectric constant of approximately 6.8 and a loss of about 0.0184 at 1 kHz, owing to its fluorinated structure [4]. In contrast, inorganic materials such as ZnO, TiO_2 and $BaTiO_3$ demonstrate much higher dielectric constants at room temperature [5–7]. Despite excellent dielectric properties, inorganic ceramics possess limitations, including brittleness, limited tunability, susceptibility to dielectric breakdown, and low toughness. As a result, their use in flexible nanoelectronics is constrained [8].

In contrast, PVDF offers strong flexibility, and excellent mechanical strength which make it suitable for flexible nanodevices. However, due its low dielectric constant than the inorganic materials hinder its application to nanoelectronics devices. Thus, researchers are trying to improve the dielectric constant and electroactive phase of PVDF. It exists in three phases: α, β, and γ with the β and γ phases being electroactive [9]. In recent times, polymer nanocomposites, incorporating inorganic nanofillers into the polymer matrix, enhancing PVDF's dielectric properties while retaining the advantages of both materials [1]. However, increasing the filler content raises the dielectric constant, but it also leads to higher dielectric loss, reduced flexibility and crystallinity [3, 10].

The choice of nanofiller also influences phase transitions within the

* Corresponding author.
E-mail address: deshrajmeena@dtu.ac.in (D. Meena).

https://doi.org/10.1016/j.mssp.2025.109607
Received 13 February 2025; Received in revised form 7 April 2025; Accepted 21 April 2025
Available online 26 April 2025
1369-8001/© 2025 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

polymer. SnO_2, nanostructures have been established as an efficient inorganic nanofiller [11,12]. For example, Zhe et al. reported a high dielectric constant of approximately 300 at 10 kHz for SnO_2/PVDF nanocomposites with a high nanofiller concentration of 0.16 vol% [11]. Adding nanostructure of inorganic materials significantly elevates specific surface interactions, which improves PVDF's electroactivity thus the dielectric properties [12].

Similarly, $Ti_3C_2T_x$ MXene has gained attention as an effective conductive filler due to its exceptional electrical conductivity of approximately 6500 S/cm and an elastic modulus of around 0.33 TPa [13,14]. MXene-based polymer nanocomposites are capable of achieving high dielectric constants with low percolation thresholds. However, as the filler loading approaches the percolation threshold, there is a significant boost in dielectric constant, dielectric loss, and electrical conductivity. This underscores the necessity for surface modifications, such as the incorporation of insulating or semiconductive layers, to optimize the material's performance [13–15].

Recently, incorporation of inorganic nanomaterials/MXene nanocomposites as nanofillers have attracted considerable interest among researchers due to their notable advantages. Specifically, the dispersion of semiconductive layer of inorganic nanomaterials on MXene nanosheets obstruct the establishment of conductive pathways. Thereby suppressing the electrical conductivity and enhances the charges at surfaces, thus improves the dielectric properties [15–17]. Like as Feng et al. synthesized $BaTiO_3$/MXene/PVDF composites, incorporating 8 wt % $BaTiO_3$ and 2 wt% MXene, which exhibited a dielectric constant of 77 and a loss of 0.15 at 100 Hz [16]. Similarly, at 1 kHz, the MXene/-TiO_2/MoS_2/P(VDF-HFP) nanosheet-loaded nanocomposite films, with an optimal loading of 8.0 wt%, demonstrated a high dielectric constant of 944 while maintaining a low dielectric loss of 0.19 [15]. Similarly, the incorporation of filler into the PVDF matrix has proven to be highly effective in enhancing the piezoelectric performance of PVDF. For instance, Kumari P. et al. Reported the improved output power 0.14 mW/cm^2 of PVDF films upon the addition of $BaTiO_3$ nanofillers [18].

Interestingly, to the best of our knowledge, the impact of SNMX nanocomposite fillers on dielectric behaviour of PVDF films has not been previously reported. Additionally, the most studies focus on higher filler concentrations, typically exceeding 5 wt%, which can reduce crystallinity, alter electroactive properties, and form networks within the polymer matrix [2]. These changes can affect electrical conductivity and augments dielectric loss. To overcome these challenges, this study employs a very low filler concentration of SNMX nanofillers, demonstrating high dielectric constant and low loss at lower filler loading [10]. The reduced filler content did not compromise the transparency, flexibility, or crystallinity of the nanocomposite films.

In this study, we have synthesized a nanocomposite of SnO_2 nanoparticles with 20 wt% of $Ti_3C_2T_x$ MXene nanosheets using hydrothermal method. The synthesized SNMX nanocomposite was used as nanofiller for the fabrication of PVDF composite films with three different weight concentrations (0.25 wt%, 0.50 wt%, and 0.75 wt%). Notably, the SNMX nanocomposite has been shown to preferentially induce the electroactive β and γ phases in PVDF. Additionally, significant improvements in the dielectric properties have been observed, with relative dielectric permittivity (ε_r) of 150 and 165 and loss tangent (tanδ) of 0.8 and 2.1 at 20 Hz for the 0.50 wt% and 0.75 wt% SNMX/PVDF composite films, respectively. Temperature-dependent studies indicate that at 100 °C, the relative dielectric permittivity of the 0.75 wt% SNMX/PVDF composite film reaches 250. The enhanced electroactive behaviour and dielectric properties, coupled with the exceptional durability of the SNMX/PVDF films, suggest their potential application as flexible power sources in portable electronic devices employing piezoelectric technology. Notably, the piezoelectric flexible nanogenerator (PENG) made from the 0.75 wt% SNMX/PVDF nanocomposite film produces a peak-to-peak voltage of 105 V.

2. Experimental section

2.1. Materials

All reagents used in the experiment were of analytical grade and were employed without any modifications, as received. MAX phase Ti_3AlC_2 (≥90 %, particle size of ≤100 μm), hydrofluoric acid (HF, 40 wt %), stannic chloride pentahydrate ($SnCl_4.5H_2O$), sodium hydroxide (NaOH), N-dimethylformamide (RG, 99.8 %) (DMF), Urea (CH_4N_2O) and hydrochloric acid (HCl) were purchased from Sigma-Aldrich Chemical Co. Ltd. The used polyvinylidene fluoride (PVDF) powder was purchased from Alfa Aesar.

2.2. Synthesis of tin oxide nanostructure (SnO_2)

A low-cost hydrothermal method was used to synthesize SnO_2 nanoparticles as shown in Fig. 1 (a). 4 g of $SnCl_4.5H_2O$ was dissolved in 30 mL ethanol and magnetically stirred to obtain a clear solution. Subsequently 1.095 g of citric acid was added into the solution and further stirred for 2 h. The resultant white solution was transferred to an autoclave filled with 70 % capacity and treat it hydrothermally for 8 h at 160 °C. After the reaction time, the precipitates were centrifuged and washed multiple times with DI water and ethanol. The collected precipitate was dried at 60 °C for 12 h.

2.3. Synthesis of MXene

$Ti_3C_2T_x$ sheets were produced by selectively etching the Al layer from the MAX phase Ti_3AlC_2 and then delaminating it using an intense layer delamination method [14]. First, 1 g of MAX phase powder was mixed with 48 % concentrated hydrogen fluoride (HF) for 24 h at room temperature using a magnetic stirrer. The mixture was filtered and the obtained precipitates washed multiple times with deionized water and centrifuged until the pH reached 7. Finally, the filtered black slurry was dried for 6 h at 80 °C and grinded into a fine powder to obtain $Ti_3C_2T_x$ MXene sheets. The used synthesis process is depicated in Fig. 1 (b).

2.4. Synthesis of SnO_2/MXene (SNMX) nanocomposite

The SnO_2/MXene nanocomposite was synthesized using the hydrothermal method. This process involved sonicating SnO_2 and $Ti_3C_2T_X$ (20 wt%) in 100 mL of distilled water, followed by the addition of 0.32 g of urea and dropwise addition of 40 μL of HCl. The resulting solution was then sealed in an autoclave and heated at 120 °C for 8 h. The resulting powder was extracted, centrifuged, and then sequentially cleaned with ethanol and deionized water. It was then allowed to dry overnight at 60 °C in an oven. The produced SnO_2/MXene nanocomposite are referred to as SNMX nanocomposite.

2.5. Synthesis of PVDF film

PVDF powder was dissolved in a 10 wt% solution of N-dimethylformamide (DMF) to produce the flexible, free-standing PVDF films. The mixture was stirred at ambient temperature for a duration of 4 h. The resulting solution was then applied to cleaned glass slides and annealed at 80 °C in a microwave oven for 2 h. Subsequently, the films were detached to yield flexible, free-standing films.

2.6. Synthesis of SNMX/PVDF composite films

Polyvinylidene fluoride (PVDF), SNMX powder, and DMF were used in the preparation of flexible composite films. PVDF powder was first dissolved in DMF for 1 h at room temperature using a magnetic stirrer, resulting in a homogeneous and clear solution. On the other hand, SNMX powder was added to DMF in varying concentrations (0.25 wt%, 0.50 wt %, and 0.75 wt%) and ultrasonically agitated for 1 h at room

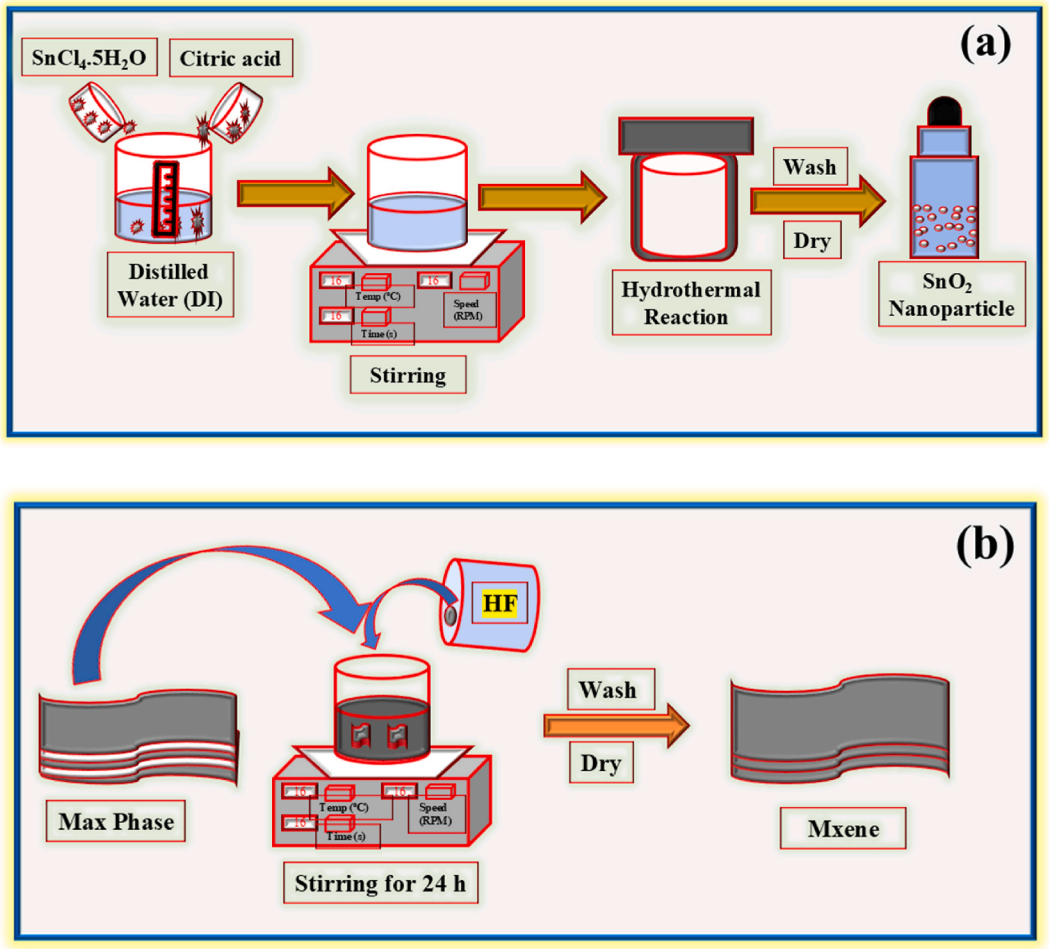


Fig. 1. A stepwise depiction of the synthesis processes for (a) SnO_2 nanoparticles and (b) Mxene flakes.

temperature. The PVDF solution was then combined with the synthesized different wt% SNMX solutions and stirred for 40 h at room temperature to get a slurry type solution of SNMX/PVDF composite. The resulting solution was then applied to cleaned glass slides and annealed at 80 °C in a microwave oven for 2 h. Subsequently, the films were detached to obtained flexible, free-standing films. Further, the thickness of the synthesized films was measured using digital vernier calliper and the measured thickness values of the synthesized composite films are tabulated in Table 1.

2.7. Fabrication of the piezoelectric flexible generator

The flexible composite films were sectioned into small pieces, each measuring 2.5 cm × 2.5 cm, and aluminium tape was applied to both surfaces to function as the top and bottom electrodes for the piezoelectric flexible nanogenerator (PENG). Copper wires were then attached to the aluminium tape using copper tape to measure the output signals when pressure was applied and the device was connected to an external circuit. Fig. 2 (a & b) illustrates both the schematic diagram of the fabricated PENG device and a digital photograph of the prepared

Table 1
Thickness values of Pure PVDF and composite films.

Sample	Pure PVDF	0.25 wt% SNMX	0.50 wt% SNMX	0.75 wt% SNMX
Thickness (mm)	0.19	0.23	0.20	0.25

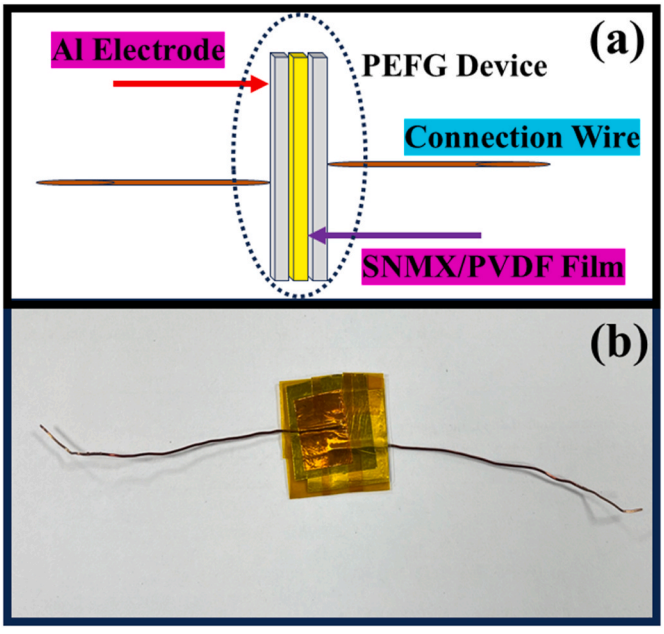


Fig. 2. (a) Schematic representation of the structure of the PENG device and (b) digital image of the fabricated PENG device.

PENG devices.

2.8. Characterizations

The crystal phases of the synthesized samples were determined using a Rigaku Ultima IV X-ray diffractometer (XRD) equipped with a Cu Kα (λ = 1.5418 Å) radiation source. The scans were performed at a rate of 2° per minute across a 2θ range of 5°–70°. An energy-dispersive X-ray spectroscopy (EDX) system (SwiftED3000, Oxford) was used for elemental mapping and compositional analysis, and an FEI Quanta 200 field emission scanning electron microscope (FESEM) was used to analyse the surface morphology of the samples. Fourier transform infrared (FTIR) spectroscopy (NICOLET 5700) was used to analyse the electroactive phases of the PVDF and composite films. The dielectric characteristics were measured using an impedance analyzer (E4990A), with a frequency range of 20 Hz to 10 MHz.

3. Results and discussion

3.1. X-ray diffraction

Structural characterization of the MAX phase Ti_3AlC_2 (before HF etching) and synthesized MXene $Ti_3C_2T_x$ (after HF etching) were performed using X-ray diffraction (XRD), as shown in Fig. 3(a). XRD pattern of the MAX phase is in good agreement with the data presented in JCPDS card no. 00-052-0875.A characteristic sharp peak corresponding to the Al layer is observed in the MAX phase at 2θ = 38.92° along the preferred (104) direction, which needs to be fully etched out during the synthesis process. The other prominent peaks at 2θ = 9.4°, 38.04°, 41.74°, 60.2°, and 74.2° correspond to the (002), (101), (105), (110), and (118) planes, respectively. After HF treatment, the characteristic diffraction peak at 2θ = 38.92° (104) corresponding to Al disappears in the XRD pattern of MXene $Ti_3C_2T_x$, confirming the complete removal of Al from the MAX phase. Most notably, the diffraction peak for the (002) plane shifts backward after HF etching, from 2θ = 9.4° in the MAX phase to 2θ = 9.02° in the MXene. This peak shift is attributed to the intensify in the c-lattice parameter of the MXene and indicates the formation of an interlayer gap in the layered structure. The XRD peaks match well with JCPDS card no. 00-032-1383, confirming the successful synthesis of MXene $Ti_3C_2T_x$ sheets from the MAX phase by HF etching [13–19].

Fig. 3(b) displays the XRD pattern of nanocrystalline SnO_2 nanoparticles (NPs). The strong diffraction peaks indicate excellent crystallinity and tetragonal rutile phase of the SnO_2. The diffraction peaks at 2θ = 26.58°, 33.94°, 38.02°, 51.80°, 54.75°, 58.09°, and 65.87° are well indexed to (110), (101), (200), (211), (220), (002), and (310) planes, respectively, according to JCPDS Card No. 41–1445 [2,11]. The SNMX composite exhibited almost all of the distinct peaks of both MXene and SnO_2, suggesting successful synthesis of SNMX nanocomposite [20]. The modest broadening of the peaks, is attributed to an increase in d-spacing and a rise in the orderliness of the crystal planes.

Fig. 4(a) presents the X-ray diffraction (XRD) patterns for all the films under investigation, spanning a 2θ range from 10° to 50°.The α, β, and γ phases constitute the majority of the PVDF and SNMX/PVDF composite films at varying concentrations. The diffraction peaks at 18.4°, 20.0° and 26.6° which can be indexed to the (020), (110), and (021) planes, respectively, correspond to the α phase of PVDF. The XRD pattern displays two moderate peaks at approximately 18.5° and 39.0°, and a strong peak at 20.3°, which can be affiliated to the γ phase. These diffraction peaks correspond to the (110/101), (020), and (211) planes, respectively. In contrast, a weak peak at 38.8°, corresponding to the (002) plane, and at 20.6°, corresponding to the (110/200) plane, are predominantly associated with the β phase. Fig. 4(b) is the magnified version of the XRD within the range of 15°–25°. The pattern reveals a noticeable decrease in the intensity of peaks associated with the α phase, while the intensities of peaks corresponding to the β and γ phases elevate significantly with the rising concentration of SNMX. The intensifies in

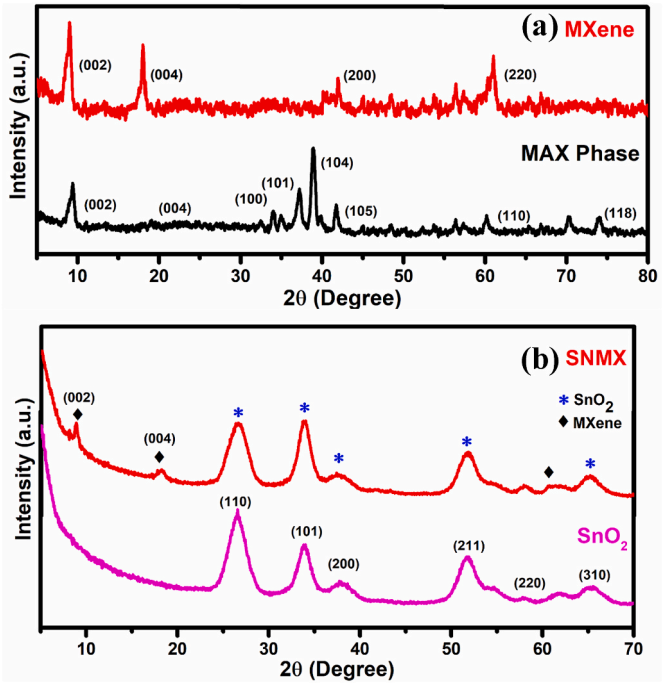


Fig. 3. XRD patterns of (a) MAX powder (Ti_3AlC_2) compared to MXene, and (b) comparison between SnO_2 nanoparticles and SNMX nanocomposite.

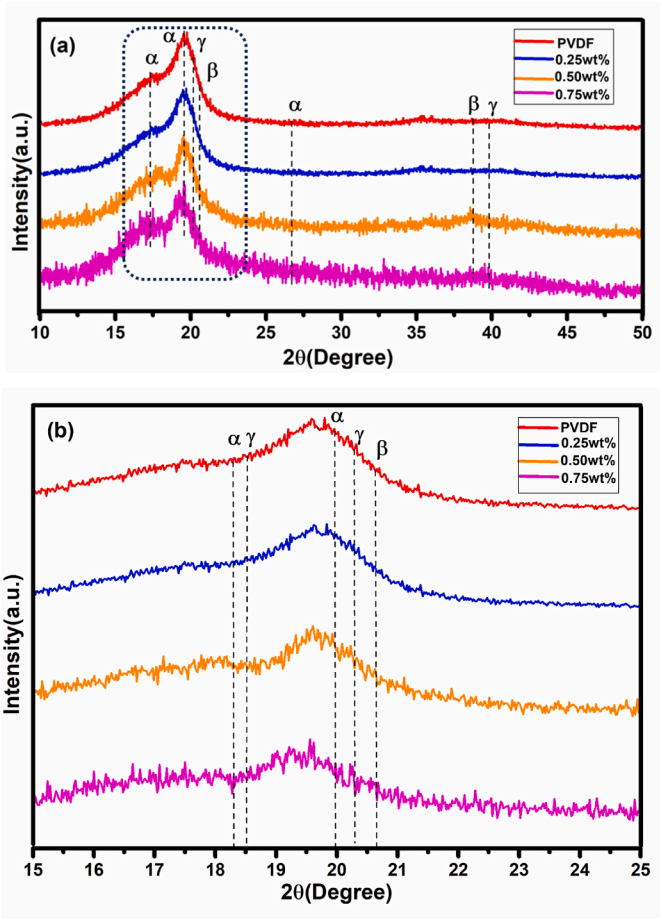


Fig. 4. XRD pattern of (a) representative samples and (b) magnified view highlighting the phases.

intensity of the (β + γ) phase suggesting the phase transition in the composite films due to the presence of nanofiller [9,10,21]. Further, the intensity ratio of the EA (β + γ) phase to the α phase has been estimated, and it was found that the electroactive (EA) phase of the highest 0.75 wt % SNMX/PVDF films is 1.3 times greater than that of the pure PVDF film. The obtained EA phase ratio corresponding to PVDF, 0.25 wt% SNMX/PVDF, 0.50 wt% SNMX/PVDF and 0.75 wt% SNMX/PVDF are 0.58, 0.62, 0.73, and 0.80 respectively. It is also important to highlight that the crystallinity of the PVDF films progressively decreases as the concentration of the nanofiller increases. This decrement in the crystallinity may be due to formation of new nucleation sites. However, as the concentration of nanofiller is very low so decrement in the crystallinity is very low as compared to reported data [15]. Thus, the obtained XRD results confirms that EA phase of the films increases without significant loss the crystallinity of the polymer.

3.2. Fourier transform infrared spectroscopy (FTIR)

Fig. 5 (a) displays a plot of the FTIR spectrum of pure PVDF and SNMX/PVDF films in the range of 400–1500 cm^{-1}. The spectra show peaks that correspond to the different phases (α, β and γ) of the PVDF. The exclusive peaks for the α phase are at approximately 488,534,616,764, 976 and 1207 cm^{-1}; for the β phase, they are at 473 and 1275 cm^{-1}; and for the γ phase, they are at 482, 811 and 1234 cm^{-1}. The absorbance for the β+γ phases is notably stronger than that of the α phase, as indicated by the peaks at 837 and 511 cm^{-1}. Consequently, the electroactive β and γ phases can be characterized by the presence of these two peaks. Several studies suggest that the peaks at 1431 and 1429 cm^{-1} can be used as characteristic bands for the β and γ phases, respectively [9,10,21]. To clearly observe how the loading of SNMX affects the intensity of the unique peaks, magnified spectra in the ranges of 740 cm^{-1} to 860 cm^{-1} and 1140 cm^{-1} to 1320 cm^{-1} have been plotted and are shown in Fig. 5 (b & c). It can be seen that the intensity of the α phase peaks at 764 cm^{-1} and 976 cm^{-1} decreases, while the intensity of the β+γ phase (837 cm^{-1}) and γ phase (1235 cm^{-1}) increases with the increment in the percentage of SNMX in the PVDF matrix, suggesting transformation of α phase into β andγphases. The absorption peak at 764 cm^{-1} cans be attributed to the bending of the CF_2 skeleton, whereas that at 837 cm^{-1} originates from the rocking of the CH_2, the stretching of the C-C skeleton and CF_2, these two peaks correspond to the non-polar α- and polar β-phases [5]. In fact, for the 0.75 wt% film, the α phase peaks nearly disappear due to the incorporation of SNMX in PVDF matrix. The relative fraction (F_{EA}) of the EA phases (β andγ) in terms of crystalline components has been calculated using equation (1) for all the films [22].

$$F_{EA} = \frac{I_{837}}{\left(\frac{K_{837}}{K_{764}}\right) I_{764} + I_{837}} \times 100\% \quad (1)$$

In this context, F_{EA} represents the fraction of the electroactive phase, while I_{837} and I_{764} denote the peak intensities at wavenumbers 837 cm^{-1} and 764 cm^{-1}, respectively. The values of K_{837} (7.7 × 10^4 cm^2/mol) and K_{764} (6.1 × 10^4 cm^2/mol) correspond to the molecular absorption coefficients at these respective wavenumbers.

The fraction formula indicates that when SNMX nanofiller incorporated into the PVDF matrix, the F_{EA} intensifies. The calculated F_{EA} values for various samples have been plotted in Fig. 5(d). The F_{EA} of the 0.75 wt % film (86 %) is 1.5 times higher compared to that of the pure film (52 %). These results align with the XRD results, which showed a decrease in the α content of the composite films and an increase in their EA phase. The rise in the electroactive phase suggests that the molecules have reoriented [10]. In this case, the polarization induced by the SNMX nanofillers promotes this reorientation. The material's elevates dielectric constant, which will be discussed subsequently, further illustrates this reorientation and the associated enhancement of the electroactive

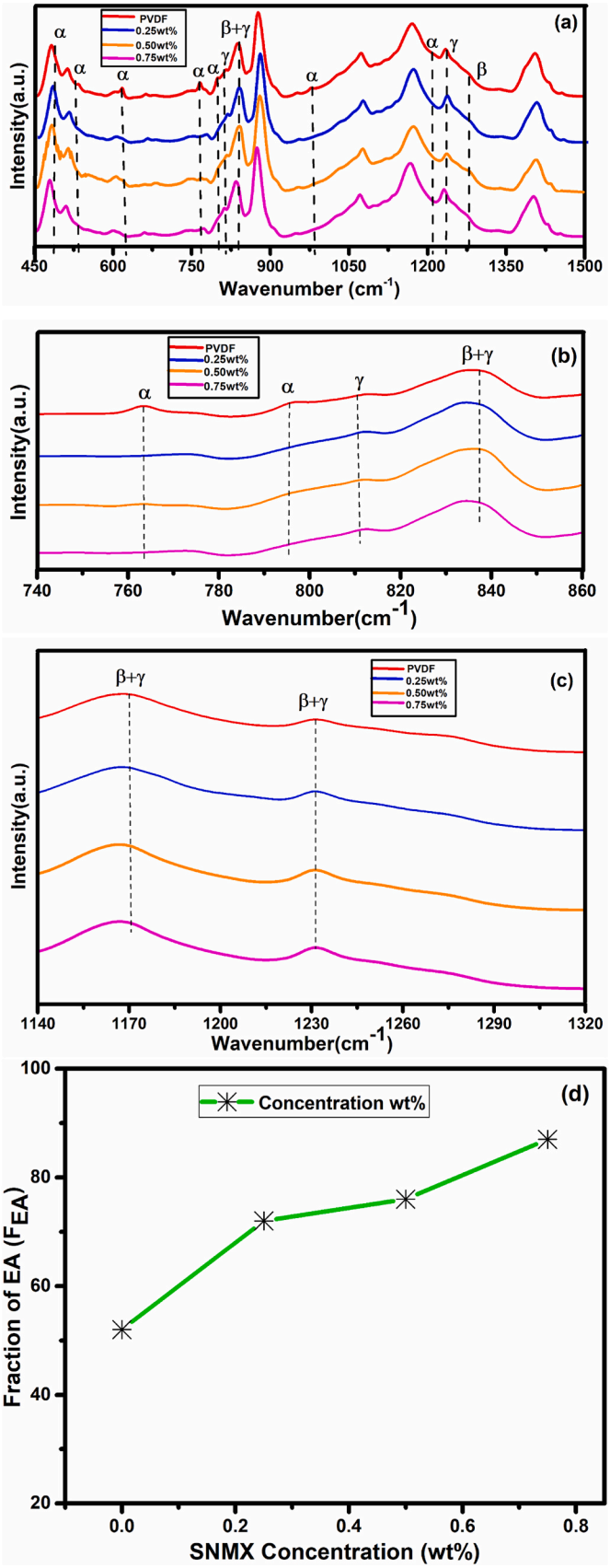


Fig. 5. FTIR spectra of all samples across different ranges: (a) 400–1500 cm^{-1}, (b) 740–860 cm^{-1}, (c) 1140–1320 cm^{-1}, and (d) plot of the fraction of EA versus SNMX concentration (wt%).

(EA) content.

3.3. Morphological and compositional analysis

FESEM analysis was employed to investigate the morphology and structure of exfoliated MXene, SnO_2 nanoparticles (NPs) and the SNMX nanocomposite. Fig. 6(a–f) shows the obtained images at low and high magnification. After the HF treatment of the MAX phase (which exhibited an uneven surface and tightly packed structure), transformed into an accordion-like morphology of MXene, as depicted in Fig. 6 (a & b). Additionally, the parallel arrangement of MXene layers resulted from the specific etching direction of "Al" [19]. The nanocrystalline morphology of SnO_2 nanoparticles (NPs) is shown in Fig. 6 (c & d). Further, the FESEM analysis confirmed the composite synthesis by the high dispersion of NPs among the MXene flakes. It can be clearly observed that the SnO_2 NPs are integrated in between the MXene layers preventing their restacking. However, some aggregated SnO_2 NPs are dispersed on top of the MXene sheets. EDX was employed to verify the elemental composition of the SNMX composite. The peaks for Sn, Ti, C, and O are observed in the EDS spectra within the 1–10 keV region as shown in Fig. 7(a). The atomic percentages of the SNMX nanocomposite are listed in the Table in the spectra [20]. The distribution of SnO_2 nanoparticles across the MXene is confirmed by the elemental mapping of the SNMX nanocomposite, which shows the presence of mixed elements of SnO_2 and MXene in the characterized area and represented in Fig. 7(b–g).

FESEM images of pure PVDF and the SNMX/PVDF composite films are presented in Fig. 8(a–d). The porous and spherulitic microstructure observed in both pure and SNMX/PVDF composite films. The morphologies of all the composite films are nearly identical. It can be observed as the SNMX filler concentration increases the surface of the films become more porous thereby suggesting uniform dispersion of SNMX nanocomposite within the PVDF matrix. Which facilitates the migration of charges within the composite film and aids in aligning the dipoles in a specific direction [22,23]. Additionally, the increment in the number of pores further indicates that incorporation of SNMX nanocomposite enhances the surface roughness of the composite films.

3.4. Impedance spectroscopic: dielectric parameters

Dielectric spectroscopy has been employed to investigate the dielectric behaviour of the prepared films. The dielectric properties of bulk or nanoscale solid materials are influenced by the spatial distribution and alignment of the local electric field within the system. Consequently, the temperature dependence and frequency-dependent relative dielectric permittivity and loss tangent offer valuable insights into the defect behaviour, structural changes and transport mechanisms within the solid [24–26].

The dielectric constant exhibits a complex nature and can be expressed as Equation (2):

$$\varepsilon^* = \varepsilon_r - j\varepsilon_r'' \tag{2}$$

The real part of the complex dielectric constant (ε^*), also known as the relative dielectric permittivity (ε_r), measures the electrical energy stored in the material when a necessary alternating electric field is applied. In contrast, the imaginary part (ε"), the dielectric loss, describes the energy dissipated in the material, as the molecules undergoing polarization are unable to match the rate of oscillation of the external electric field [27].

The relative dielectric permittivity (ε_r) and loss tangent for the materials were determined using Equations (3) and (4), respectively.

$$\varepsilon_r = \frac{Cd}{A\varepsilon_0} \tag{3}$$

$$tan\,\delta = \frac{\varepsilon_r''}{\varepsilon_r} \tag{4}$$

where A is the electrode area, ε_0 is the permittivity of free space (8.85×10^{-12} F/m), d is the sample thickness, C is the capacitance [23–26].

Fig. 9(a–d) shows the variation of the relative dielectric permittivity with frequency (20 Hz - 10 MHz) of the applied electric field at different operating temperatures for all the films. It is evident that the relative dielectric permittivity (ε_r) intensifies with temperature across the entire frequency range, while it diminishes with higher frequency. Furthermore, the incorporation of the SNMX nanocomposite into the PVDF matrix significantly enhances the relative dielectric permittivity of the films. At room temperature and 20 Hz, the relative dielectric permittivity elevates from 7 for pristine PVDF to 35 with 0.25 wt%, 147 with 0.50 wt%, and 165 with 0.75 wt% SNMX/PVDF composite films. Increase in relative dielectric permittivity with the addition of SNMX may be attributed to enhanced Maxwell–Wagner–Sillars (MWS) interfacial polarization and easy charge migration at lower applied frequency [24–26]. As the nanofiller concentration increases, polarization driven by the alignment of interfacial charges between the polymer and the

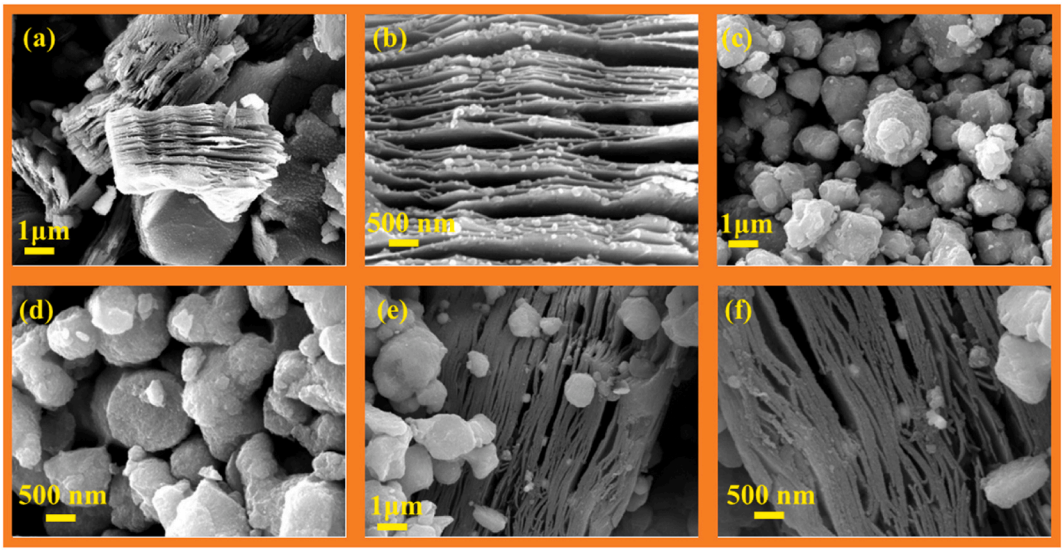


Fig. 6. FESEM images showing (a & b) the accordion-like morphology of exfoliated MXene, (c & d) SnO_2 nanoparticles, and (e & f) MXene flakes functionalized with SnO_2 nanoparticles.

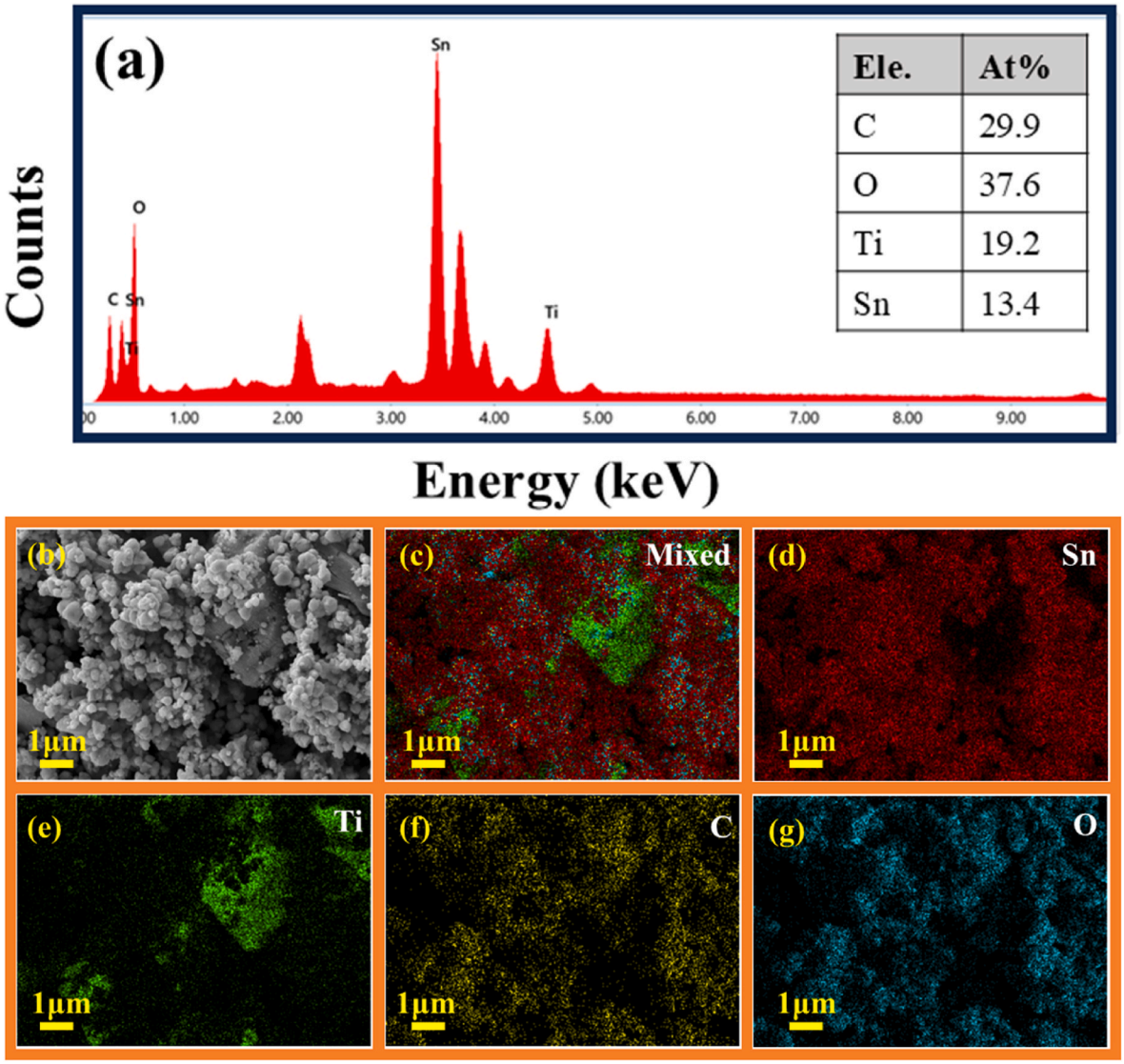


Fig. 7. (a) EDX analysis, and (b to g) elemental mapping of the SNMX nanocomposite.

SNMX nanofiller intensifies. Furthermore, as β and γ phases of PVDF are known to exhibit the highest spontaneous polarization. Thereby, the enhancement in EA phase means increment in the polar phase with loading of SNMX nanofillers favouring the reorientation of the molecules in applied external electric field. Thus, the film with high EA content demonstrates the maximum dielectric constant [10,23].

At higher operating temperatures, ε_r continues to proliferate, reaching approximately 250 at 100 °C and 20 Hz, while the relative dielectric permittivity of pure PVDF increases to 60 under the same conditions. At room temperature, the dielectric constant boosts by about 25 times, whereas at 100 °C, it increases approximately 4 times compared to pure PVDF films. These results indicate that even a small amount of SNMX nanocomposite as nanofillers significantly improves the relative dielectric permittivity of PVDF at room temperature. The observed escalation in the relative dielectric permittivity with rising operating temperature can be ascribed to the enhanced segmental motion of the molecular chains in PVDF [23,28].

While additional nanofiller would likely increase the relative dielectric permittivity, it would also result in higher dielectric loss. The variation of the loss tangent (tanδ) with applied frequency (20 Hz - 10 MHz) is plotted for all the films at different working temperature and displayed in Fig. 10(a–d). It has been observed that the loss tangent amplifies as the nanofiller SNMX concentration grows in the PVDF. At 20 Hz and room temperature, the loss tangent of the nanocomposite films with 0.25 wt%, 0.50 wt% and 0.75 wt% SNMX is approximately 0.5, 0.75 and 2.3, respectively, while that of pure PVDF is around 0.2 under the same conditions. The frequency-dependent changes in loss tangent may attributed to different relaxation mechanisms within the materials. One is the α-relaxation in the crystalline region, accounts for the loss tangent at low frequencies. The increase in dielectric loss at higher frequencies suggests a second relaxation peak corresponding to α-relaxation in the amorphous region, occurring in the 1 MHz–10 MHz range [23]. It is interesting to note that the obtained losses are significantly lower with enhanced relative dielectric permittivity as compare to the previous reported data [10,11]. The lower losses are attributed to significant enhancement in the EA phase with low concentration of the SNMX fillers without reducing the crystallinity of the PVDF. For lower filler loadings, up to the percolation threshold, the filler is encapsulated within the PVDF matrix, acting as small, individual capacitors that collectively enhance both ε_r and tanδ [15]. The reduced relative dielectric permittivity and loss tangent of SNMX/PVDF nanocomposites, as compared to PVDF, can be attributed to mobility constraints. The presence of nanoparticles hinders the movement of polymer chains, resulting in an increase in relaxation time and a shift of the relaxation process to lower frequencies. As a result, nanocomposites with a SNMX content exhibit a decrease in dielectric loss [23].

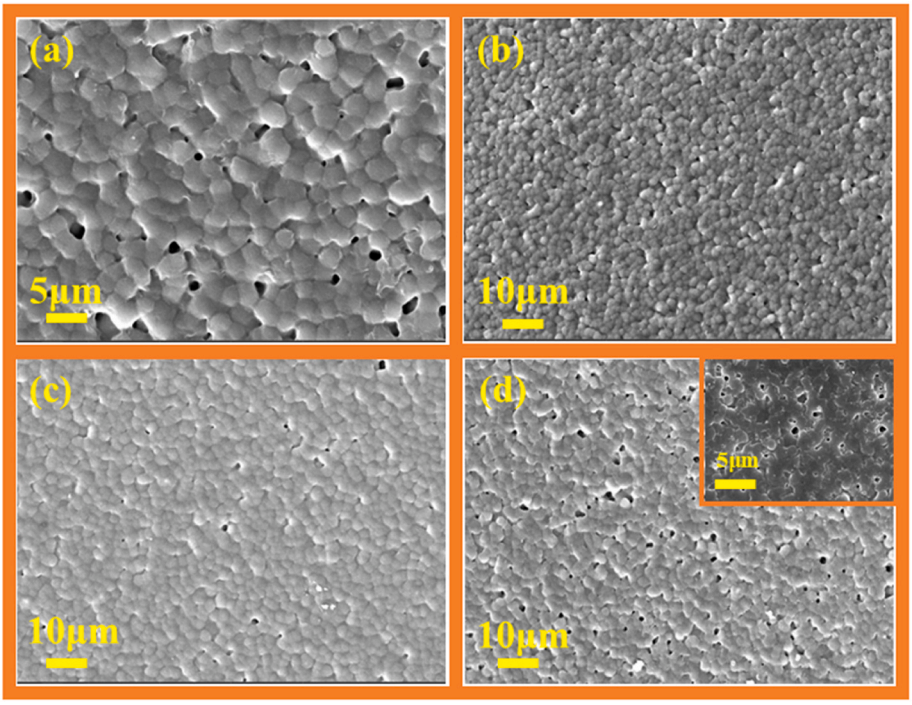


Fig. 8. FESEM images of (a) pure PVDF, (b) 0.25 wt% SNMX/PVDF, (c) 0.50 wt% SNMX/PVDF, and (d) 0.75 wt% SNMX/PVDF.

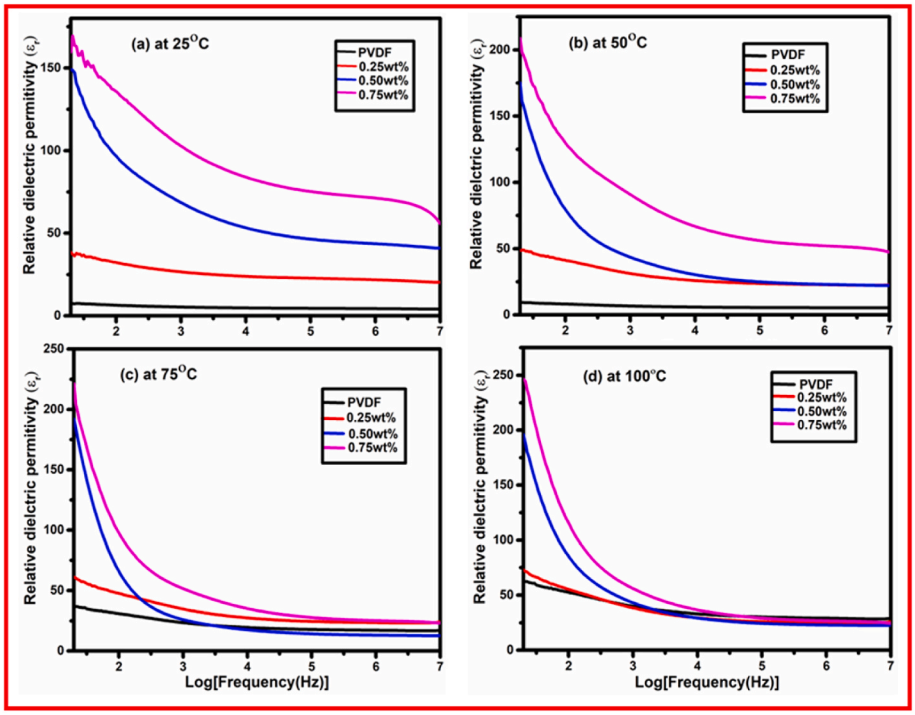


Fig. 9. Relative dielectric permittivity ε_r spectra of PVDF and SNMX/PVDF composites films at (a) 25 °C, (b) 50 °C, (c) 75 °C, and (d) 100 °C.

Further, the temperature-dependent loss tangent behaviour suggests that the losses increase with elevated operating temperatures for all films, except for the 0.75 wt% SNMX nanocomposite film. The 0.75 wt% film demonstrate highest loss tangent (~2.3) at room temperature and lowest loss tangent (~1.1) at 50 °C. However, on further increment in the operating temperature (75 and 100 °C) the loss increases to ~ 1.5 and remains same. Also, the losses of 0.25 wt% and 0.50 wt% films at 100 °C becomes twice that of the at room temperature. The obtained relative dielectric permittivity and loss tangent values at different operating temperature at a frequency of 20 Hz is demonstrated in Fig. 11 (a and b).

A comparison of recent studies on the enhancement of the relative dielectric permittivity of PVDF using different nanofillers is provided in Table 2. Although, some reports demonstrate high relative dielectric permittivity with higher loss tangent at substantial wt% of nanofillers. That leads to degradation in crystallinity and flexibility of the PVDF

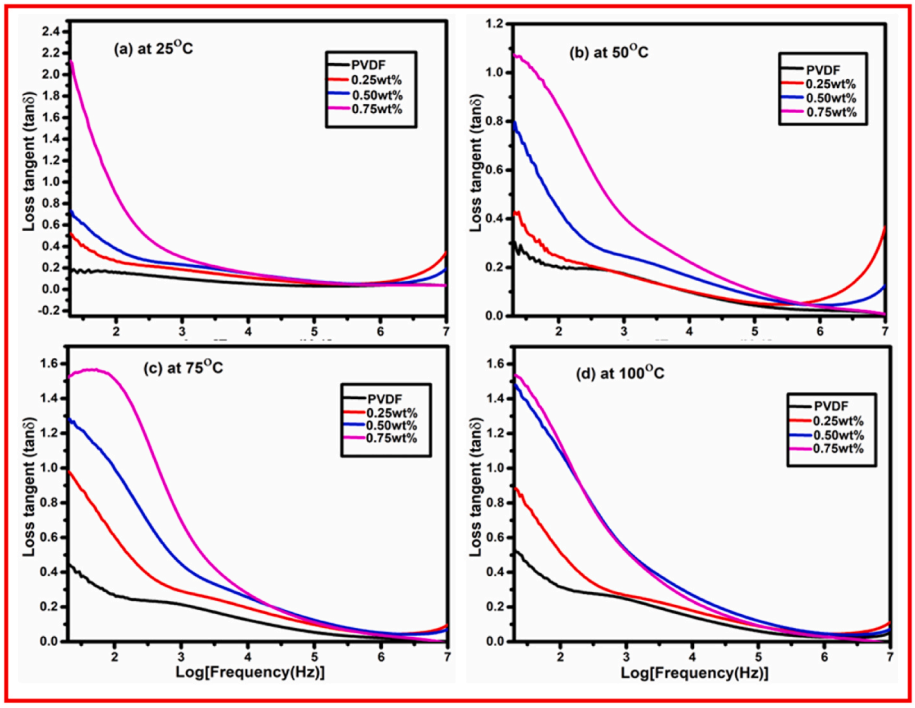


Fig. 10. Loss tangent spectra of PVDF and SNMX/PVDF composites films at (a) 25 °C, (b) 50 °C, (c) 75 °C, and (d) 100 °C.

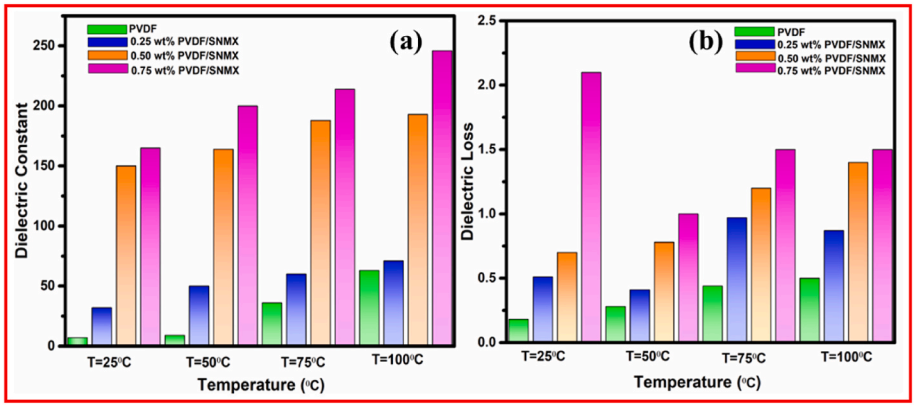


Fig. 11. Relative dielectric permittivity and loss tangent values at various temperatures measured at a frequency of 20 Hz.

Table 2
A comparative analysis of the dielectric parameters with PVDF composite films at room temperature.

S.No.	Nanofiller/PVDF Composite	Nanofiller Conc.	Dielectric Parameters		Freq.	Ref.
			Dielectric constant	Dielectric loss		
1.	$BaTiO_3/SnO_2$/PVDF	60 vol%	145	0.8	10 kHz	[29]
2.	SnO_2/PVDF	0.16 vol%	300	0.82	10 kHz	[11]
3.	SnO_2/PVDF	15 wt%	40	0.34	40 Hz	[2]
4	DA@SnO_2/PVDF	1 wt%	6		1 kHz	[30]
5	SnO_2/PVDF	0.1 wt%	145	0.8	10 Hz	[10]
6	$BaTiO_3$/MXene/PVDF	8 wt%	77	0.14	100 Hz	[16]
7	MnO_2/MXene/PVDF	7 wt%	37	0.2	1 kHz	[31]
8	$CaCu_3Ti_4O_1$/MXene/PVDF	30 vol%	28	0.061	10 kHz	[17]
9	SnO_2@MXene/PVDF	0.25 wt%	38	0.7	20 Hz	This work
10	SnO_2@MXene/PVDF	0.50 wt%	150	0.6	20 Hz	This work
11	SnO_2@MXene/PVDF	0.75 wt%	165	2.1	20 Hz	This work

films.

Thus, the obtained dielectric results and findings enhance our understanding the influence of SNMX nanofillers in PVDF matrix. The results suggest that the synthesized films have the potential to act as flexible dielectric material in various industries, including electronics, energy storage, and telecommunications.

4. Piezoelectric flexible nanogenerator (PENG) studies

An oscilloscope was used to measure the generated voltage output after constructing the piezoelectric flexible nanogenerator (PENG) device using the synthesized SNMX/PVDF flexible composite films. The graphs in Fig. 12(a–d) display the open-circuit voltage produced by the devices with different concentrations of SNMX nanocomposites in the PVDF matrix when force was applied using a manual taper. The plots demonstrate that the open-circuit voltage intensifies as the SNMX content in the PVDF matrix rises. The observed peak-to-peak voltage (PPV) outputs for the composite films are 65 ± 1.5 V, 97 ± 1.5 V and 105 ± 1.5 V, corresponding to 0.25 wt%, 0.50 wt% and 0.75 wt% SNMX nanofillers in the PVDF matrix, respectively. While the virgin PVDF film PPV output is 32 ± 1.5 V. The voltage output from the PENG device follows an alternating voltage waveform [2,10,18].

The range of load resistances influences the resulting power density. Therefore, power density of 0.75 wt% SNMX/PVDF composite films were investigated across the load resistances from 500 kΩ to 10 MΩ. The obtained variations in power density as a function of load resistance is presented in Fig. 13. The formula used to calculate the output power density of the PENG device is given in Equation (5). Where V is the voltage across the load resistance R and A is the effective area of electrodes.

$$P = \frac{V^2}{R \times A} \tag{5}$$

The power density increases up to 3 MΩ, reaching a maximum of 0.44 μW/cm^2, after which it decreases as the load resistance continues to increase. According to the maximum power transfer theorem, maximum power is achieved when the internal impedance of the source matches the external load resistance [22].

The PPV output of 0.75 wt% SNMX/PVDF composite film is enhanced by more than three times than that of the pristine PVDF film. It has been reported that boosting the dielectric constant, reducing the film thickness, and improving the surface roughness of the active material are highly effective in improving the performance of the PENG [24,28,32]. Further, the nanofillers are commonly incorporated into polymer matrices to enhance the dielectric constant and the surface roughness of polymers. Thus, this enhancement in PPV output may be attributed to increased relative dielectric permittivity and improved surface roughness of the 0.75 wt% SNMX/PVDF film as suggested by the FESEM. The energy harvested by the PENG can be used to power small circuits, functioning as a nanogenerator.

5. Conclusion

SNMX/PVDF composite films, made of PVDF and SNMX nanocomposites, have been studied for their electroactive phase formation and significant enhancement of dielectric characteristics over a wide frequency range (20 Hz–1 MHz). When 0.75 wt% SNMX is added to the PVDF matrix, the SNMX/PVDF composite film exhibits the maximum electroactive phase yield (F_{EA} = 86 %). FTIR and XRD analyses confirm that the electroactive phases in the SNMX/PVDF films have nucleated. In particular, the SNMX/PVDF films loaded with 0.50 wt% and 0.75 wt% SNMX nanocomposite at 20Hz and room temperature exhibit the high relative dielectric permittivity, 150 and 165, respectively, with notably low loss tangent. The temperature-dependent dielectric characteristics of the manufactured films shows that the relative dielectric permittivity of composite films increase rapidly as compare to pure PVDF. The largest peak-to-peak voltage produced by a piezoelectric flexible nanogenerator (PENG) made from 0.75 wt% SNMX/PVDF nanocomposite film is 105V. Therefore, the SNMX/PVDF composite films with increased electroactive phase, superior dielectric properties and enhanced PPV as PENG suggest that they may/can meet the needs of next-generation electronic components, such as power conditioning, flexible film capacitors, and pulsed power generators.

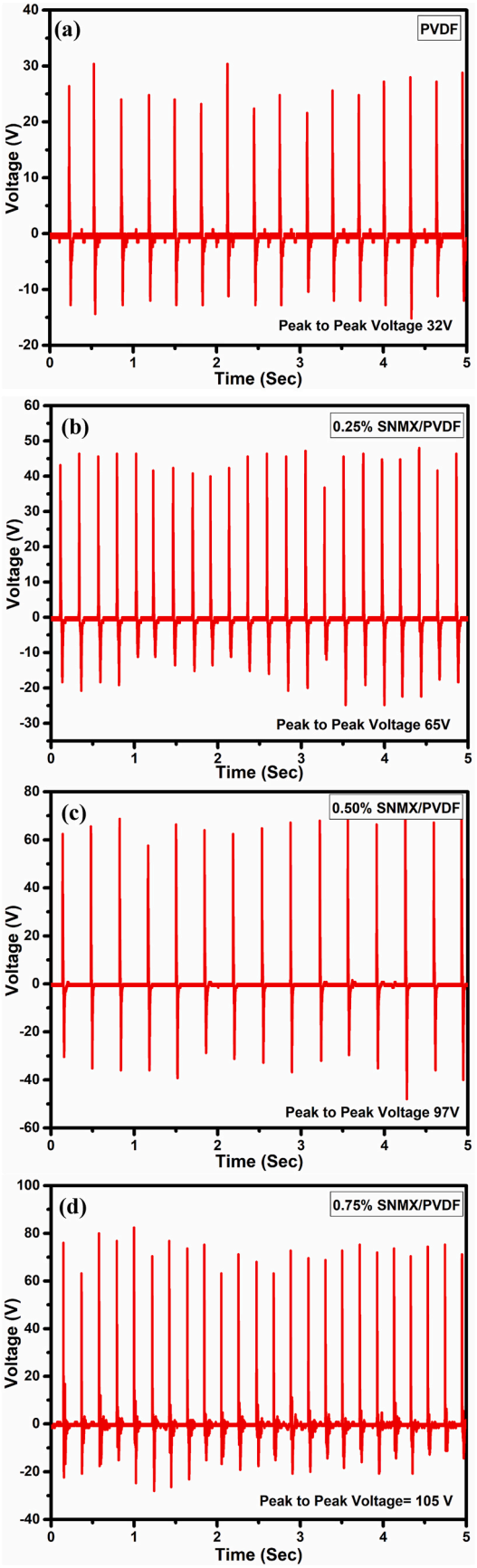


Fig. 12. Peak to Peak Voltage output characteristics of (a) pure PVDF, (b) SNMX (0.25 wt%), (c) SNMX (0.50 wt%), and (d) SNMX (0.75 wt%) composites films.

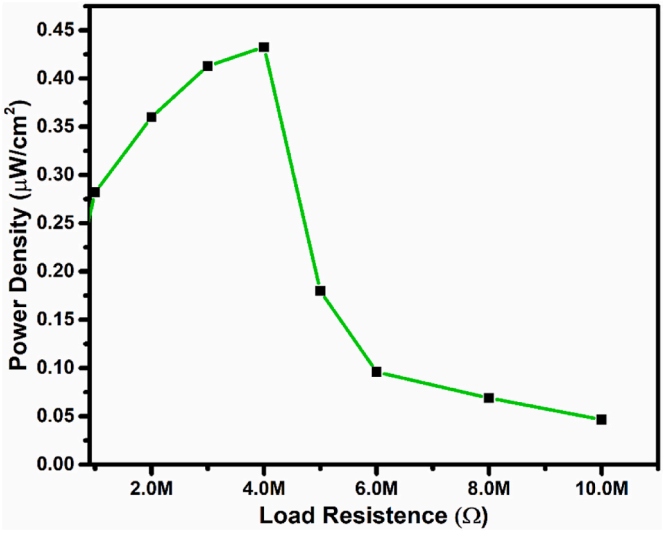


Fig. 13. Variation of power density of the 0.75 wt% SNMX/PVDF composite film as a function of load resistance.

CRediT authorship contribution statement

Gagan Sharma: Writing – original draft, Visualization, Methodology, Investigation, Conceptualization. **Deshraj Meena:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors acknowledge the use of fabrication & characterization facilities at IIT Delhi under Indian Nanoelectronics Users' Programme-Idea to Innovation (INUP-i2i) project no. (RP04176G) funded by the Ministry of Electronics and Information Technology (MeitY). Gagan Sharma expressed his gratitude to Delhi Technological University (DTU), India, for providing the DTU fellowship (2K21/PHDAP/02).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mssp.2025.109607.

Data availability

No data was used for the research described in the article.

References

[1] A. Sengupta, S. Das, S. Dasgupta, P. Sengupta, P. Datta, Flexible nanogenerator from electrospun PVDF–polycarbazole nanofiber membranes for human motion energy-harvesting device applications, ACS Biomater. Sci. Eng. 7 (4) (2021) 1673–1685.
[2] E. Kar, N. Bose, B. Dutta, S. Banerjee, N. Mukherjee, S. Mukherjee, 2D SnO2 nanosheet/PVDF composite based flexible, self-cleaning piezoelectric energy harvester, Energy Convers. Manag. 184 (2019) 600–608.
[3] A. Thakur, J. Mandeep, S. Dam, N.C. Shekar, G. Amarendra, S. Hussain, A. Saha, Enhancing the electroactive phases in freestanding flexible films of MoS2/PVDF, Polymer Crystallization 4 (2) (2021) e10164.
[4] A. Anand, D. Meena, K.K. Dey, M.C. Bhatnagar, Enhanced piezoelectricity properties of reduced graphene oxide (RGO) loaded polyvinylidene fluoride (PVDF) nanocomposite films for nanogenerator application, J. Polym. Res. 27 (2020) 1–11.
[5] N.D. Kulkarni, P. Kumari, Development of highly flexible PVDF-TiO2 nanocomposites for piezoelectric nanogenerator applications, Mater. Res. Bull. 157 (2023) 112039.
[6] M. Choi, G. Murillo, S. Hwang, J.W. Kim, J.H. Jung, C.Y. Chen, M. Lee, Mechanical and electrical characterization of PVDF-ZnO hybrid structure for application to nanogenerator, Nano Energy 33 (2017) 462–468.
[7] N.H. Yusoff, R.A.M. Osman, M.S. Idris, K.N.D.K. Muhsen, N.I.M. Nor, Dielectric and structural analysis of hexagonal and tetragonal phase BaTiO3, in: AIP Conference Proceedings, vol. 2203, AIP Publishing, 2020, January. No. 1.
[8] J. Tang, H. Zhang, Z. Yang, Z. Mao, J. Zhang, Comparation of barium titanate and multi-walled carbon nanotubes modified PVDF composites: a perspective of dielectric, crystalline, rheological and mechanical properties, Ceram. Int. 50 (1) (2024) 1059–1068.
[9] X. Cai, T. Lei, D. Sun, L. Lin, A critical analysis of the α, β and γ phases in poly (vinylidene fluoride) using FTIR, RSC Adv. 7 (25) (2017) 15382–15389.
[10] A. Thakur, M. Jangra, S. Dam, S. Hussain, Enhanced dielectric constant with addition of a low amount of SnO2 nanoparticles as fillers in a PVDF matrix with low dielectric loss, Nano-Structures & Nano-Objects 34 (2023) 100978.
[11] Z. Liu, F. Wang, H. Zhu, Enhanced dielectric properties of polyvinylidene fluoride with addition of SnO2 nanoparticles, Phys. Status Solidi Rapid Res. Lett. 10 (10) (2016) 753–756.
[12] S. Sarkar, S. Garain, D. Mandal, K.K. Chattopadhyay, Electro-active phase formation in PVDF–BiVO 4 flexible nanocomposite films for high energy density storage application, RSC Adv. 4 (89) (2014) 48220–48227.
[13] A.D. Dillon, M.J. Ghidiu, A.L. Krick, J. Griggs, S.J. May, Y. Gogotsi, A.T. Fafarman, Highly conductive optical quality solution-processed films of 2D titanium carbide, Adv. Funct. Mater. 26 (23) (2016) 4162–4168.
[14] A. Lipatov, H. Lu, M. Alhabeb, B. Anasori, A. Gruverman, Y. Gogotsi, A. Sinitskii, Elastic properties of 2D Ti3C2T x MXene monolayers and bilayers, Sci. Adv. 4 (6) (2018) eaat0491.
[15] W. Ma, K. Yang, C. Zhou, H. Li, Enhanced dielectric constant and suppressed electrical conductivity in polymer nanocomposite films via loading MXene/TiO2/MoS2 nanosheets, Ceram. Int. 48 (8) (2022) 10447–10457.
[16] Y. Feng, Q. Deng, C. Peng, Q. Wu, High dielectric and breakdown properties achieved in ternary BaTiO3/MXene/PVDF nanocomposites with low-concentration fillers from enhanced interface polarization, Ceram. Int. 45 (6) (2019) 7923–7930.
[17] A. Tsyganov, M. Vikulova, I. Zotov, I. Burmistrov, A. Gorokhovsky, N. Gorshkov, Synergistic effect of CaCu3Ti4O12 ceramic and Ti3C2Tx MXene nanoflakes on the dielectric properties of poly (vinylidene fluoride) composites, Mater. Chem. Phys. 313 (2024) 128714.
[18] M. Kumar, N.D. Kulkarni, P. Kumari, Piezoelectric performance enhancement of electrospun functionally graded PVDF/BaTiO3 based flexible nanogenerators, Mater. Res. Bull. 174 (2024) 112739.
[19] A. Mondal, M. Faraz, M. Dahiya, N. Khare, Double-layer electronegative structure-based triboelectric nanogenerator for enhanced performance using combined effect of enhanced charge generation and improved charge trapping, ACS Appl. Mater. Interfaces 16 (38) (2024) 50659–50670.
[20] S. Gasso, M.K. Sohal, A. Mahajan, MXene modulated SnO2 gas sensor for ultra-responsive room-temperature detection of NO2, Sensor. Actuator. B Chem. 357 (2022) 131427.
[21] T. Furukawa, Ferroelectric properties of vinylidene fluoride copolymers, Phase Transitions A Multinatl. J. 18 (3–4) (1989) 143–211.
[22] A. Kumar, R. Sharma, Piezoelectric flexible generator based on Mn-doped ZnO/PVDF composite films for energy harvesting application, Curr. Appl. Phys. 68 (2024) 159–168.
[23] S. Chen, V.K. Thakur, A.A. Skordos, Fluorine and nitrogen doping of zinc oxide to enhance dielectric storage of PVDF based particulate composites, Mater. Sci. Eng., B 302 (2024) 117244.
[24] H. Parangusan, D. Al-Sowaidi, E.H. Elhadrami, D. Ponnamma, J. Bhadra, Development of flexible PVDF/BaTiO3-MoS2 polymer nanocomposites for energy harvesting and gas sensing applications, J. Mater. Sci. Mater. Electron. 35 (10) (2024) 679.
[25] P.L. Reddy, K. Deshmukh, T. Kovářík, D. Reiger, N.A. Nambiraj, K.P. Sk, Enhanced dielectric properties of green synthesized Nickel Sulphide (NiS) nanoparticles integrated polyvinylalcohol nanocomposites, Mater. Res. Express 7 (6) (2020) 064007.
[26] R. Meena, R.S. Dhaka, Dielectric properties and impedance spectroscopy of NASICON type Na3Zr2Si2PO12, Ceram. Int. 48 (23) (2022) 35150–35159.
[27] S. Sangwan, R. Meena, R. Bokolia, V. Singh, D. Meena, Exploration of structural and dielectric properties of orthorhombic Ta2O5 nanoplatelets towards the potential optoelectronic devices, Mater. Today Commun. 38 (2024) 108468.
[28] R. Sahoo, S. Mishra, A. Ramadoss, S. Mohanty, S. Mahapatra, S.K. Nayak, Temperature-dependent dielectric properties of metal-doped ZnO nanofiller reinforced PVDF nanocomposites, Mater. Res. Bull. 132 (2020) 111005.
[29] J.W. Zha, X. Meng, D. Wang, Z.M. Dang, R.K. Li, Dielectric properties of poly (vinylidene fluoride) nanocomposites filled with surface coated BaTiO3 by SnO2 nanodots, Appl. Phys. Lett. 104 (7) (2014).
[30] M.V. Paranjape, S.A. Graham, H. Patnam, P. Manchi, J.S. Yu, Dopamine treated SnO2/PVDF composite films for hybrid mechanical energy harvester, Compos. Sci. Technol. 221 (2022) 109323.
[31] L. Cao, Y. Chen, S. Feng, Y. Su, Reduced dielectric loss of MXene/PVDF composites with adding MnO2 nanorods, Ceram. Int. 50 (22) (2024) 47655–47665.
[32] N.D. Kulkarni, A. Saha, P. Kumari, The development of a low-cost, sustainable bamboo-based flexible bio composite for impact sensing and mechanical energy harvesting applications, J. Appl. Polym. Sci. 140 (29) (2023) e54040.

Optics Communications 586 (2025) 131875



Contents lists available at ScienceDirect

Optics Communications

journal homepage: www.elsevier.com/locate/optcom





Temperature-dependent coupling characteristics in a silicon photonic crystal waveguide coupler

Varnam Sherawat [a], Yogita Kalra [a], Renuka Bokolia [b], Ravindra Kumar Sinha [a,c,*]

[a] TIFAC-Centre for Relevance and Excellence in Fiber Optics and Optical Communication, Department of Applied Physics, Delhi Technological University, Delhi, India
[b] Department of Applied Physics, Delhi Technological University, Delhi, India
[c] Gautam Buddha University, Gautam Budh Nagar, UP, Greater Noida, 201312, India

ARTICLE INFO

Keywords:
Coupling length
Photonic crystal waveguide (PCW)
Coupler
Temperature
Thermo-optic effect
Coupled mode theory (CMT)

ABSTRACT

This paper presents a systematic analysis of temperature-dependent coupling characteristics in a photonic crystal waveguide (PCW) coupler. Unlike previous works that primarily rely on either numerical simulations or experimental studies, our approach integrates an analytical framework based on Coupled Mode Theory (CMT) with computational validation using COMSOL Multiphysics. The proposed PCW coupler, composed of silicon rods in air, exhibits a significant reduction in coupling length, enhancing its suitability for photonic integrated circuits (PICs). Through a combination of Plane Wave Expansion (PWE) simulations and wavelength-domain modelling, we demonstrate that the coupling length decreases with increasing temperature, indicating enhanced efficiency for thermally sensitive applications. The study not only provides theoretical insights into thermo-optic coupling in PCWs but also offers a practical design strategy for compact photonic devices operating at variable temperatures.

1. Introduction

The increasing interest in developing photonic crystal (PC) based devices and their applications is predominantly due to the versatility in the control and manipulation of light in customized microstructures. Moreover, continual advancements in PC fabrication technology play a crucial role in the actualization of various key PC applications such as sensing physical quantities such as temperature [1], pressure [2,3], and refractive indices of bio-analytes [4]. Other notable applications include waveguide devices and add/drop multiplexing devices [5], PC nanocavity lasers [6], optical switching [7], etc. Control over the propagation of light on a subwavelength scale using the photonic bandgap (PBG) effect has enabled the integration of these applications in photonic integrated circuits (PICs) [8,9]. The intricate light confinement in photonic crystal waveguides (PCWs) facilitates the realization of complicated PICs [10].

Coupled mode theory (CMT) offers a framework to describe the behaviour of light as it propagates, focusing on the interactions between various modes in optical systems. This theory is particularly useful for modelling how modes in waveguides and resonators interact, facilitating the calculation of coupling coefficients and an understanding of interference phenomena [11]. CMT is applicable to optical cavities linked with waveguides, where the interaction between cavity modes and waveguide modes is explained through the field equivalence principle, providing a precise depiction of light behaviour in these systems [12]. The temporal coupled-mode theory (TCMT) has been developed to account for optical phenomena in structures like metasurfaces, addressing asymmetry by integrating the propagation phase into the coupling terms, thereby offering insights into the interaction between light and matter [13]. Additionally, CMT has been adapted to explore acousto-optic interactions in optomechanical crystals, enabling the directional control of elastic waves over optical waves and the study of infinite optical modes resulting from nonlinear interactions [14].

While previous studies have extensively explored the behavior of PCWS under various external factors, most investigations either focus on purely numerical simulations or experimental validations without an analytical framework. The absence of an integrated approach limits the generalizability of results and their applicability to photonic integrated circuits (PICs). Furthermore, conventional studies often overlook the explicit temperature-dependent derivation of coupling length, which is critical in real-world applications where thermal variations influence device performance. Our work addresses these gaps by developing a

* Corresponding author. TIFAC-Centre for Relevance and Excellence in Fiber Optics and Optical Communication, Department of Applied Physics, Delhi Technological University, Delhi, India.
E-mail address: dr_rk_sinha@yahoo.com (R.K. Sinha).

https://doi.org/10.1016/j.optcom.2025.131875
Received 28 November 2024; Received in revised form 7 April 2025; Accepted 13 April 2025
Available online 14 April 2025
0030-4018/© 2025 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

hybrid analytical-numerical approach, leveraging the well-established Coupled Mode Theory (CMT) alongside computational simulations in COMSOL. This study has been used to deduce a temperature-dependent expression for coupling length for the proposed PCW coupler design. This methodology allows for a more systematic and scalable analysis, ensuring that theoretical predictions align with simulated outcomes, thereby enhancing the design flexibility of temperature-sensitive photonic devices.

The study proposed in this paper introduces a novel methodology to theoretically analyze how temperature influences the coupling characteristics of waveguide directional couplers within two-dimensional PCs. The design features two parallel photonic crystal waveguides constructed from silicon (Si) rods in a hexagonal lattice arrangement with air as the background medium, both sharing identical geometries and separated by a row of rods with a reduced radius. By leveraging the thermo-optic effect inherent in Si, this research investigates the variations in the coupling parameters such as the coupling length resulting from changes in temperature. The coupling length is computed for the optical wavelength of 1.33 μm, across various temperatures within the range 300 K–400 K, by utilizing the wavelength domain simulations in the commercially available software COMSOL Multiphysics. This study offers important perspectives on enhancing PCWs for temperature-sensitive applications in PICs.

2. Theoretical design

The proposed design consists of a hexagonal lattice of dielectric silicon (Si) rods in air with a lattice constant 'a' and radius of the rods 'r', taken as 0.2a. Two rows of Si rods are removed, with a row of smaller-sized rods placed in between them to form a defect layer. Hence, a pair of parallel PCWs is formed, with one waveguide each on either side of the defect layer, constituting a PCW directional coupler. The schematic diagram of the proposed PCWs is shown in Fig. 1.

2.1. Design optimization

The design parameters were systematically optimized to ensure strong field confinement and minimal coupling losses. Using parametric simulations in COMSOL and Plane Wave Expansion (PWE) methods, the following optimizations were made.

- The lattice constant 'a' has been varied in the range 400–500 nm. The value a = 450 nm is found to align the midgap corresponding to the photonic bandgap with the optical wavelength ~1.33 μm. This value yields the best trade-off between compactness and efficient mode propagation.
- The radius of the silicon rods 'r' is chosen to be r = 0.2a to obtain a wide bandgap for TM mode operation.
- The defect rod radius r_d has been varied from 0.1a to 0.2a. An optimal value of r_d = 0.15a results in maximal field overlap and enhanced coupling efficiency.
- The spacing between the waveguides, which corresponds to one row of reduced-radius rods, has been selected to balance between strong coupling and minimal crosstalk.

The optimized parameters for the proposed design have been summarized in Table 1.

Table 1
Optimized Structural parameters for the proposed PCW coupler.

Parameter	Symbol	Optimized Value	Effect on Coupling Behavior
Lattice constant	a	450 nm	Ensures mode matching with λ = 1.33 μm
Rod radius	r	0.2a	Maintains wide photonic bandgap
Defect radius	r_d	0.15a	Maximizes coupling efficiency
Operating wavelength	λ	1.33 μm	Matches the midgap frequency

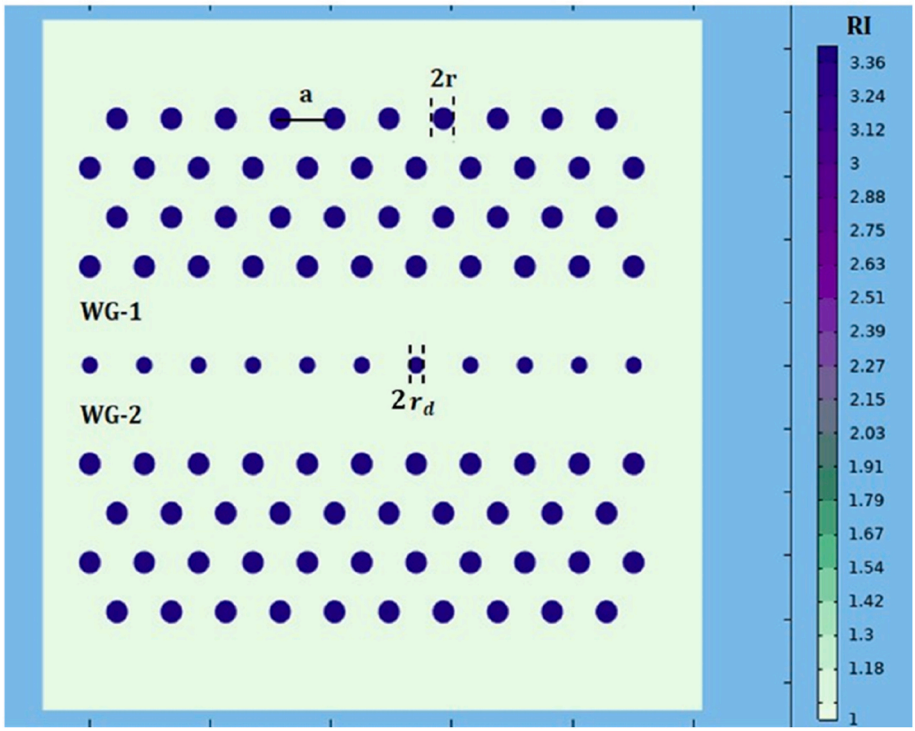


Fig. 1. Schematic design and refractive index profile of the proposed structure at 300 K.

2.2. Analytical modelling

The analytical modelling is based on the Coupled Mode Theory (CMT), which provides a mathematical framework to describe light propagation and power transfer between two coupled waveguides. In the case of two identical PCWs placed in proximity, the even and odd supermodes exhibit different propagation constants β_{even} and β_{odd}.The power transfer between the waveguides is governed by the phase difference between these supermodes, and the complete coupling of power occurs at specific interaction lengths, known as the coupling length L_c, given by:

$$L_c = \frac{\pi}{\Delta\beta} = \frac{\pi}{\beta_{odd} - \beta_{even}}$$

Here, $\Delta\beta$ is directly influenced by the thermo-optic induced change in the refractive index of silicon rods. The increase in temperature modifies the index contrast, leading to a larger $\Delta\beta$, and hence a reduced coupling length.

The difference in propagation constants $\Delta\beta$ due to increase in temperature can also be expressed as:

$$\Delta\beta\,(T) = \beta_{odd}(T) - \beta_{even}(T)$$

As temperature increases, the refractive index of Si increases linearly. The following relation is a widely accepted linear approximation for the refractive index of silicon as a function of temperature, particularly valid in the near-infrared region (~1.3–1.55 μm) and within moderate temperature ranges (300–400 K) [15]:

$$n(T) = n_0 + \gamma(\mathrm{T} - T_0)$$

where $\gamma = 1.86 \times 10^{-4}\ K^{-1}$ is the thermo-optic coefficient for silicon. This change in $n(T)$ alters β values, due to shift in modal dispersion and thus, causes $\Delta\beta$ to increase. Consequently, the coupling length L_c decreases, as deduced by Equation (7) derived in Section 4.2. This dependency is verified with numerical simulations the results of which are represented in Fig. 6 in Section 4.3.

3. Temperature dependence & thermo-optic effect in silicon

In this study, Equation (1) below is utilised to determine the refractive index (RI) and the corresponding dielectric constants of the dielectric rods across temperatures ranging from 300 K to 400 K. The refractive index of silicon in the infrared (IR) region [1] can be expressed as:

$$n = 3.38\left(1 + 3.9 \times 10^{-5} \times T\right)\ 77\mathrm{K} < \mathrm{T} < 400\ \mathrm{K} \qquad (1)$$

where T represents the temperature of the silicon rods. This linear dependence of RI on temperature is depicted in Fig. 2. The values of refractive index and dielectric constant derived from this equation are then applied in the PWE (Plane Wave Expansion) simulations to obtain the photonic bandgap (PBG) of the proposed structure. Further, wavelength domain simulations in COMSOL Multiphysics have been performed at different temperatures in the specified range to study the

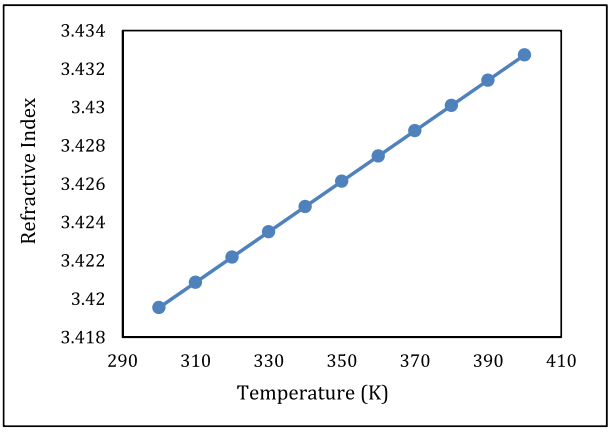


Fig. 2. Linear temperature dependence of RI due to thermo-optic effect in Si.

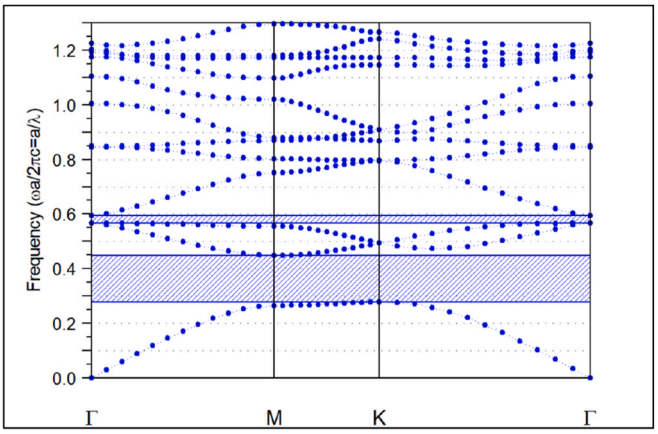


Fig. 3. The shaded region indicates the photonic bandgap of the PCW structure at 300 K corresponding to the TM mode.

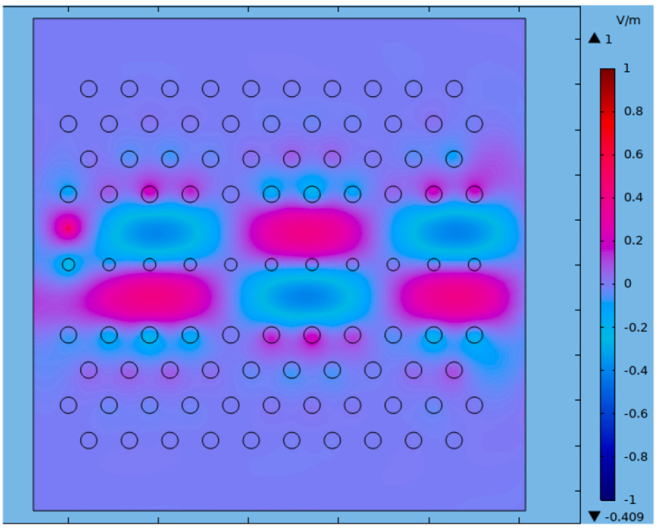


Fig. 4. Propagation of Electric Field (E_z) at λ = 1.33 μm along the interaction length of the PCWs (periodic power exchange). The contour map scale on the RHS represents the field values in V/m.

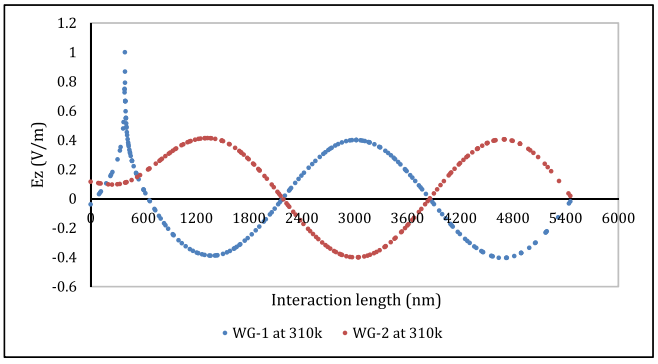


Fig. 5. Field amplitude (E_z) as a function of distance (z) along the direction of wave propagation in the PCWs at a temperature of 310 K.

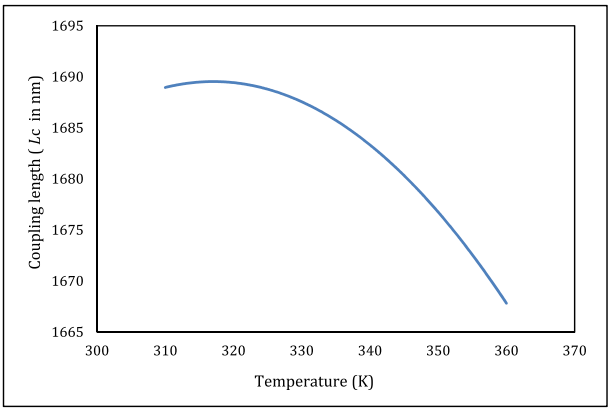


Fig. 6. Variation of coupling length with temperature for the proposed PC coupler.

effect of temperature variation on the coupling between the PCWs.

4. Results and discussion

The photonic bandgap of the PCW structure has been calculated using PWE simulations. The structure shows a wide bandgap for the transverse magnetic (TM) mode ranging from 0.27 to 0.45 in the normalized frequency units (a/λ), where λ is the wavelength in vacuum. The dispersion diagram in Fig. 3 shows the bandgap of the PCW structure at room temperature (~300 K), corresponding to a midgap wavelength of ~1.33 μm for the optimized 'a' and 'r' values. It is inferred that two guided TM modes emerge inside the PBG, originating from the splitting of an isolated single waveguide mode. These are the supermodes of the composite PCW structure, named as the odd and even modes as per the plane of symmetry between the waveguides. The central row of rods has been modified to tune the dispersion properties of the supermodes, and thus the coupling characteristics.

When two identical waveguides (WGs) are positioned close enough for their fields to interact, light can be transferred from one waveguide to the other through a process known as evanescent field coupling [16]. Efficient optical power transfer between the waveguides occurs when their mode profiles are similar, as explained by the coupled mode theory (CMT) [17–20]. For parallel waveguide directional couplers, the maximum power transfer $\boldsymbol{T}$ from the initial waveguide is described by the following relationship [21]:

$$\boldsymbol{T} = \left(1 - \frac{\boldsymbol{k}^2}{\boldsymbol{k}^2 + \Delta\boldsymbol{\beta}^2/4}\right) \tag{2}$$

where $\boldsymbol{k}$ represents the coupling coefficient, a key parameter proportional to the field overlap between the two adjacent waveguides (refer to Equation ($\boldsymbol{A13}$) derived in the Appendix). $\Delta\boldsymbol{\beta}$ is the difference in their respective propagation constants, also known as the *phase mismatch term* [22], and is expressed as:

$$\Delta\boldsymbol{\beta} = \frac{\boldsymbol{2\pi}}{\boldsymbol{\lambda}}\,\Delta\boldsymbol{n} \tag{3}$$

where, λ represents the operating wavelength.

4.1. Switching in coupler

A 2 × 2 directional coupler comprising of two closely positioned waveguides is characterized by the odd and even eigenmodes or supermodes. CMT stipulates that on placing two parallel and identical PCWs close to each other, the defect mode splits into an even and odd mode. The odd and even modes have different propagation constants (refer to Equation ($\boldsymbol{A6}$) in Appendix), namely $\boldsymbol{\beta_o}$ and $\boldsymbol{\beta_e}$ respectively, for every frequency. Switching happens when the relative path length of light is altered by half a wavelength, corresponding to a phase difference of $\boldsymbol{\pi}$ [23]. For a device length L, the change in refractive index required to switch from one state to another is given by $\Delta n.\boldsymbol{L} = \boldsymbol{\lambda/2}$.

The device proposed here uses a directional coupler in closely spaced PCWs to switch light. The propagation of a light in one of the WGs can be described using the superposition of the fundamental even and odd modes. It utilizes the fact that one of the two supermodes of the PCW coupler has a lower group velocity. The **coupling length** or the **switching length** ($\boldsymbol{L_{SW}}$) is defined as the length of the adjacent waveguides where the phase difference between these two modes equals an odd multiple of π. Thus, for a directional coupler based switching operation, the following condition must be satisfied [24]:

$$|\boldsymbol{\beta_e} - \boldsymbol{\beta_o}|\,\boldsymbol{L_{SW}} = (\boldsymbol{2n+1})\boldsymbol{\pi}, \mathrm{n} = 0, 1, 2\ldots\ldots$$

where, n is an integer number. Likewise, if the two modes have a phase difference which is an even multiple of π (i.e. 2mπ, with m being an integer) then the light is coupled back to the initial waveguide again.

The coupling length is a function of the splitting of the two modes. In the case of the proposed PCW coupler, the coupling length ($\boldsymbol{L_c}$) i.e., the distance required for complete power transfer between the waveguides (refer to Appendix), is given by Refs. [25–28]:

$$\boldsymbol{L_c} = \frac{\boldsymbol{\pi}}{\boldsymbol{2k}} = \frac{\boldsymbol{\pi}}{|\boldsymbol{\beta_2} - \boldsymbol{\beta_1}|} = \frac{\boldsymbol{\pi}}{\Delta\boldsymbol{\beta}} \tag{4}$$

Here, the propagation constants of the fundamental modes of the individual PC waveguides are taken as $\boldsymbol{\beta_1}$ and $\boldsymbol{\beta_2}$. Also, the coupling coefficient $\boldsymbol{k}$ is defined as, $\boldsymbol{k} = \frac{|\boldsymbol{\beta_2}-\boldsymbol{\beta_1}|}{\boldsymbol{2}}$.

4.2. Effect of temperature

The bandgap wavelength, corresponding to the band diagram in Fig. 3, is modulated using the thermo-optic change of refractive index in silicon. The change in refractive index leads to a change in the photonic band structure. Due to the considerable thermo-optic coefficient of Si, the refractive index of Si can be increased by 0.013 with a temperature change of 100 K based on Equation (1). As the temperature increases above the room temperature, the transmission spectrum of the PCW structure shifts to larger wavelengths. The band-edges corresponding to the PBG, along with the midgap wavelength, show a shift with an increase in temperature as reported in our earlier studies [1,4]. This temperature-induced shift in the bandgap also leads to changes in the coupling characteristics and parameters including the coupling length of the proposed PCW directional coupler.

On introducing a temperature variation ($\Delta\boldsymbol{T}$) (increasing $\boldsymbol{T}$ beyond the room temperature), Equation (1) can be used to determine the change in the refractive index ($\Delta\boldsymbol{n}$) of the Si rods as follows:

$$\Delta\boldsymbol{n} = \boldsymbol{13.182} \times \boldsymbol{10^{-5}} \times \Delta\boldsymbol{T} \tag{5}$$

Further, substituting the RI change from Equation (5) into the phase mismatch term in Equation (3), it is obtained that:

$$\Delta\boldsymbol{\beta} = \frac{\boldsymbol{2\pi}}{\boldsymbol{\lambda}}\left(\boldsymbol{13.182} \times \boldsymbol{10^{-5}} \times \Delta\boldsymbol{T}\right) \tag{6}$$

The difference in the propagation constants ($\Delta\boldsymbol{\beta}$) from Equation (6) is introduced in Equation (4) for theoretically computing the coupling length ($\boldsymbol{L_c}$) with variation in temperature.

$$\boldsymbol{L_c} = \frac{\boldsymbol{\lambda}}{\boldsymbol{2}\,(\Delta\boldsymbol{n})} = \frac{\boldsymbol{\lambda}}{\boldsymbol{2} \times \left(\boldsymbol{13.182} \times \boldsymbol{10^{-5}} \times \Delta\boldsymbol{T}\right)} \tag{7}$$

Here, the operating wavelength λ is taken as the midgap wavelength of the PC structure for simulations. This wavelength (~1.33 μm) is launched into the first waveguide (above the defect layer named WG-1 in Fig. 1). For the purpose of this study, Equation (7) represents the

condition required to switch light from one state to another as discussed in section 4.1. This condition is clearly temperature dependent. As indicated by CMT, the phase shifts between the odd and even modes at different frequencies are different since one of the modes exhibits a lower group velocity. Likewise, the phase shifts between the two modes will be different for different temperatures. This is because the odd and even modes, having different propagation constants and different group velocities, will experience the temperature induced change in refractive index differently. Hence, the effects of superposition of the even and odd modes are expected to be different at different temperatures, making the distance of optical power transfer different for different temperatures. This has been corroborated by computational results in the following section.

4.3. Coupling simulations and power transfer

For the wavelength domain simulations in COMSOL, to excite light propagation in WG-1, a point dipole source with only a z-component of electric field $\mathbf{E_Z}$ with an amplitude of 1 V/m is launched into WG-1 at a wavelength of 1.33 μm which is the midgap wavelength of the PC structure. This wavelength is chosen to minimize the dispersion and coupling losses in the proposed PCW structure.

Due to the second waveguide (below the defect layer named WG 2 in Fig. 1) being in close proximity to the first one, the launched field couples from WG-1 to WG-2. This is because there is an evanescent field extending outside each waveguide (WG). Since the two parallel guides are placed close enough, these parts of the evanescent fields must overlap spatially. The overlap of modes through the evanescent field induces a redistribution of energy between the WGs.

If the two WGs run parallel for a sufficient distance (which is the case here), the exchange of power can be highly significant. The field then starts coupling back to WG-1 and it keeps coupling back and forth over the interaction length of the two PCWs as shown in Fig. 4. It shows the electric field (ewfd) surface plot obtained using the wavelength domain computations run in COMSOL.

Thus, as per the coupled mode theory (CMT), there is a periodic exchange of power between the two PCWs (refer to Equations (***A19***) and (***A20***) in the Appendix). This power transfer is periodic with distance or the interaction length (along the direction of propagation). This periodic variation of field with the interaction length along the propagation direction (z) is depicted in Fig. 5. The electric field ($\mathbf{E_Z}$) values along the interaction length of the PCWs are obtained from the wavelength domain simulations and imported in a tabulated format for plotting. This graph, as shown in Fig. 5 for a sample value of temperature T = 310 K, is also plotted for different values of T above the room temperature (T > 310 K). The blue dotted line represents $\mathbf{E_Z}$ as a function of the interaction length in WG-1 in which the field is launched, hence the blue dotted line starts at the launch value $\mathbf{E_Z} = 1$ V/m. At this point the field value in WG-2, represented by the red dotted line, is negligible.

As power transfer takes place periodically between the two PCWs, the field values vary periodically with distance. The distance at which complete power transfer takes place from WG-1 to WG-2, i.e. the distance between the peak values of $\mathbf{E_Z}$ in the two PCWs, is taken as the value of the coupling length for that temperature.

From Fig. 5, which depicts the variation of the normalized $\mathbf{E_Z}$ value along the interaction length of the PCW/, the average value of $\boldsymbol{L_c}$ at 310 K is approximately 1689.545 nm. Similarly, the coupling length is calculated at different values of T, where T > 300 K.

The values of $\boldsymbol{L_c}$ have been obtained at different temperatures above the room temperature (~300 K) as described above. The variation of coupling length with temperature is depicted in Fig. 6. The coupling length shows a decreasing trend with increase in temperature as seen from Fig. 6. These simulation results are corroborated by Equation (7) and its interpretation discussed in section 4.2. As the mismatch between the propagation constants of the odd and even modes increase with temperature ($\boldsymbol{\Delta\beta}$ given in Equation (6)), the coupling length, $\boldsymbol{L_c} = \frac{\pi}{\Delta\beta}$, decreases, and a higher coupling coefficient is achieved. It can be inferred that the coupling efficiency increases above the room temperature with even a slight change in the refractive index of the Si rods. This study provides crucial perspectives on the effect of increase in temperature above the room temperature, which may occur due to the heating in such Si-based PCW devices and PICs. Additionally, it also offers important insights on enhancing Si-based PCWs for temperature-sensitive applications in PICs.

4.4. Comparative analysis

The proposed PCW directional coupler has been compared with previously published works, as presented in Table 2. Table 2 clearly shows that the proposed coupler shows a much shorter coupling length than the existing PhC-based structures, employing coupling mechanisms, without compromising on the functionality. A shorter coupling length minimizes lattice spacing and propagation losses, playing a crucial role in the development of compact and efficient devices optimized for integration. In designing PhC-based compact integrated modules, downscaling device dimensions can be of great use. Hence, the proposed design emerges as an efficient one in realizing couplers which are more compact than the existing ones. The design is flexible and can be optimized to have different operating wavelengths. It exhibits efficient light switching for a low refractive index change, which may result out of device heating above room temperatures, thus making it suitable for testing device performance above the room temperature.

5. Conclusion

The impact of temperature variation on the coupling parameters of a photonic crystal waveguide (PCW) coupler has been studied theoretically. The directional coupler has been implemented on a 2-D hexagonal lattice photonic crystal (PC) composed of Si rods in air. The thermo-optic

Table 2
Comparison with existing PhC-based structures.

Reference	Function	Structure	Coupling Length
Luis et al., 2021 [29]	Optical switch	Slow-light PCW	37 μm for 100 periods 74 μm for 200 periods
Bin Ni et al., 2019 [30]	Demultiplexer	Hybrid plasmonic waveguide	4.1 μm
L. Chiang et al., 2022 [31]	Thermo-optic switch	Slot waveguide directional coupler	90 μm
C. Zhang et al., 2020 [32]	Optical switch	Microring resonator	15.72 μm
R. Chen et al., 2022 [33]	Electrically controlled 2 × 2 programmable unit	Three-waveguide directional coupler on SOI	64 μm
X. Yang et al., 2021 [34]	2 × 2 non-volatile optical switch	Mach–Zehnder switch	9 μm (single-drive MZI) 3.88 μm (push-pull dual-drive MZI)
Chen et al., 2022 [35]	Tunable switch	GSST-based PhC directional coupler	2.6 μm
Xie et al., 2023 [36]	Light modulator	Polymer hybrid PCW device	NA
K. Miao et al., 2025 [37]	Polarization beam splitter	Twin butterfly porous cores in a photonic crystal fiber (PCF)	Lx = 12 μm and Ly = 25 μm (coupling lengths of the two orthogonal polarization modes)
C. Liu et al., 2025 [38]	Polarization beam splitter	Dual-core terahertz PCF	29.506 mm
Proposed work	Thermo-optically tuned coupler	Line-defect PCW	**1.6 μm**

effect in Si is leveraged to obtain variation in its refractive index with an increase in temperature above the room temperature (~300 K). As per the calculations using the Plane Wave Expansion (PWE) method, the proposed PC structure shows a photonic bandgap (PBG) for the TM mode and this PBG is found to shift with changes in temperature. The coupling length is computed at different temperatures by means of wavelength domain simulations done in COMSOL Multiphysics software. The coupling length shows a decreasing trend with an increase in temperature which is corroborated theoretically by the well-known coupled mode analysis of identical waveguides placed in close proximity to each other. The proposed coupler exhibits a much smaller coupling length, providing a compact yet efficient design which works for even a small change in refractive change over a considerable range of temperature. This investigation provides valuable insights into optimizing PCWs for temperature-sensitive applications in PICs**.**

CRediT authorship contribution statement

Varnam Sherawat: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yogita Kalra:** Visualization, Validation, Supervision, Resources, Methodology, Formal analysis. **Renuka Bokolia:** Validation, Supervision, Resources, Project administration, Investigation. **Ravindra Kumar Sinha:** Writing – review & editing, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors gratefully acknowledge the initiatives and support provided by the establishment of the TIFAC-Centre of Relevance and Excellence in Fiber Optics and Optical Communication at Delhi Technological University (Formerly Delhi College of Engineering) Delhi, through the “Mission Reach” program of Technology Vision-2020 of the Government of India. One of the authors, Varnam Sherawat (VS), acknowledges the PhD fellowship provided by Delhi Technological University, Delhi, India.

APPENDIX

The well-known coupled mode equations, used in the analysis of two optically interacting waveguides placed in the vicinity of each other, are given by Refs. [28,39]:

$$\frac{\partial a}{\partial z} = -i\beta_1 a - ik_{12} b \tag{A1}$$

$$\frac{\partial b}{\partial z} = -i\beta_2 b - ik_{21} a \tag{A2}$$

Where, β_1 and β_2 denote the propagation constants of the modes propagating along the z direction in the individual waveguides (WGs); k_{12} and k_{21} denote the coupling coefficients. The field amplitudes in the individual WGs, a and

b, are given by:

$$a = a_0 e^{-i\beta_1 z}$$

$$b = b_0 e^{-i\beta_2 z}$$

For a pair of identical WGs, $k_{12} = k_{21}$

For analysing the system of coupled waveguides, we consider the postulate that there exists a wave in the composite structure travelling with a phase constant β. The wave in the composite system is the superposition of the modes of the individual WGs. The field amplitudes in WG-1 and WG-2 are then given by:

$$a = a_0 e^{-i\beta z} \tag{A3}$$

$$b = b_0 e^{-i\beta z} \tag{A4}$$

Substituting $(A3)$ & $(A4)$ in the coupled mode equations $(A1)$ & $(A2)$, we get

$$-i\beta a_0 = -i\beta_1 a_0 - ik_{12} b_0$$

$$-i\beta b_0 = -i\beta_2 b_0 - ik_{21} a_0$$

Rearranging the above equations gives,

$$(\beta - \beta_1) a_0 - k_{12} b_0 = 0$$

$$(\beta - \beta_2) b_0 - k_{21} a_0 = 0$$

The above equations yield the following matrix,

$$\begin{bmatrix} \beta - \beta_1 & -k_{12} \\ -k_{21} & \beta - \beta_2 \end{bmatrix} \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

For a non-trivial solution, we have

$$\begin{vmatrix} \beta-\beta_1 & -k_{12} \\ -k_{21} & \beta-\beta_2 \end{vmatrix} = 0$$

$$\Rightarrow \beta^2 - (\beta_1+\beta_2)\beta + \beta_1\beta_2 - k^2 = 0 \tag{A5}$$

where, $k^2 = k_{21}k_{12}$

Now, (A5) is a quadratic equation in β and its two solutions are:

$$\beta = \frac{(\beta_1+\beta_2) \pm \sqrt{(\beta_1+\beta_2)^2 - 4(\beta_1\beta_2 - k^2)}}{2}$$

Simplifying the above equation leads to

$$\beta = \frac{1}{2}(\beta_1+\beta_2) \pm \sqrt{\frac{(\beta_1^2+\beta_2^2-2\beta_1\beta_2)+4k^2}{4}}$$

$$\Rightarrow \beta_{s,a} = \frac{1}{2}(\beta_1+\beta_2) \pm \sqrt{1/4(\beta_1-\beta_2)^2 + k^2} \tag{A6}$$

Thus, the composite system of waveguides has a set of two independent modes, one with the propagation constant β_s known as the *symmetric mode* (or the even mode), and the other with the propagation constant β_a known as the *anti-symmetric mode* (or the odd mode).

Hence, the general solutions to the coupled mode equations, given by (A3) and (A4), are modified to:

$$a(z) = a_s e^{-i\beta_s z} + a_a e^{-i\beta_a z} \tag{A7}$$

$$b(z) = \frac{\beta_s-\beta_1}{k_{12}} a_s e^{-i\beta_s z} + \frac{\beta_a-\beta_1}{k_{12}} a_a e^{-i\beta_a z} \tag{A8}$$

Assuming that at $z = 0$ unit power is launched into WG-1 and there is no power in WG-2.

So, the initial conditions at $z = 0$ yield:

$$a_s + a_a = 1 \tag{A9}$$

$$\frac{\beta_s-\beta_1}{k_{12}} a_s + \frac{\beta_a-\beta_1}{K_{12}} a_a = 0 \tag{A10}$$

Solving equations (A9) and (A10) gives,

$$a_a = \frac{-(\beta_1-\beta_s)}{\beta_s-\beta_a} \tag{A11}$$

$$a_s = \frac{\beta_1-\beta_a}{\beta_s-\beta_a} \tag{A12}$$

Substituting (A11) and (A12) into equations (A7) and (A8) gives the z dependent amplitudes in the two WGs:

$$a(z) = \frac{\beta_1-\beta_a}{\beta_s-\beta_a} e^{-i\beta_s z} + \left(\frac{\beta_s-\beta_1}{\beta_s-\beta_a}\right) e^{-i\beta_a z}$$

$$b(z) = \left(\frac{\beta_s-\beta_1}{k_{12}}\right)\left(\frac{\beta_1-\beta_a}{\beta_s-\beta_a}\right) e^{-i\beta_s z} + \left(\frac{\beta_a-\beta_1}{K_{12}}\right)\left(\frac{\beta_s-\beta_1}{\beta_s-\beta_a}\right) e^{-i\beta_a z}$$

Power in WG-1 will be given by:

$$|a(z)|^2 = 1 - \frac{k^2}{\frac{1}{4}\Delta\beta^2 + k^2} \sin^2\left\{\left(\sqrt{\frac{1}{4}\Delta\beta^2 + k^2}\right) z\right\} \tag{A13}$$

And power in WG-2 will be given by:

$$|b(z)|^2 = \frac{k^2}{\frac{1}{4}\Delta\beta^2 + k^2} sin^2\left\{\left(\sqrt{\frac{1}{4}\Delta\beta^2 + k^2}\right) z\right\} \tag{A14}$$

Equations (A13) and (A14) represent the power transfer equations in a directional coupler for the case of non-identical waveguides.

For the case of identical waveguides, we have $k_{12} = k_{21} = k$ and

$$\beta_1 = \beta_2 = \beta_0 \tag{A15}$$

Using equation (A6) the propagation constant for the symmetric mode becomes:

$$\beta_s = \beta_0 + k \tag{A16}$$

And for the anti-symmetric mode:

$$\beta_a = \beta_0 - k \tag{A17}$$

Substituting equations $(A15), (A16), (A17)$ in equations $(A11)$ and $(A12)$, we get

$$a_s = a_a = 1/2 \tag{A18}$$

Substituting these in $(A7)$, we get

$$a(z) = \frac{1}{2}e^{-i(\beta_0+k)z} + \frac{1}{2}e^{-i(\beta_0-k)z}$$

$$= \frac{1}{2}e^{-i\beta_0 z}\left\{e^{-ikz} + e^{ikz}\right\}$$

$$a(z) = e^{-i\beta_0 z}\, cos(kz)$$

Therefore, power in WG-1 is given as

$$|a(z)|^2 = cos^2(kz) \tag{A19}$$

Substituting $(A15), (A16)$ *and* $(A17)$ in $(A8)$, we obtain

$$b(z) = \frac{1}{2}e^{-i(\beta_0+k)z} - \frac{1}{2}e^{-i(\beta_0-k)z}$$

$$= -\frac{1}{2}e^{-i\beta_0 z}\left(e^{ikz} - e^{-ikz}\right)$$

$$= (-i)\, e^{-i\beta_0 z}\, sin(kz)$$

$$= e^{-i\pi/2}e^{-i\beta_0 z}\, sin(kz)$$

$$b(z) = e^{-i\left(\beta_0 z + \frac{\pi}{2}\right)}\, sin(kz)$$

Hence, the power in WG-2 is given by

$$|b(z)|^2 = sin^2(kz) \tag{A20}$$

Equations $(A19)$ and $(A20)$ represent the power transfer equations for the system of identical waveguides. From these equations, it is evident that complete power transfer from WG-1 to WG-2 takes place at $z = \frac{\pi}{2k}$. This distance is defined as the *coupling length* for the case of identical waveguides.

Data availability

Data will be made available on request.

References

[1] V. Sherawat, R. Bokolia, R.K. Sinha, Impact of thermal and refractive index tuning on the bandgap and band-edges of a silicon photonic crystal waveguide with sensing applications, Opt. Commun. 518 (2022) 128348, https://doi.org/10.1016/j.optcom.2022.128348.

[2] T. Karrock, M. Gerken, Pressure sensor based on flexible photonic crystal membrane, Biomed. Opt. Express 6 (2015) 4901, https://doi.org/10.1364/BOE.6.004901.

[3] D. Yang, H. Tian, N. Wu, Y. Yang, Y. Ji, Nanoscale torsion-free photonic crystal pressure sensor with ultra-high sensitivity based on side-coupled piston-type microcavity, Sens. Actuat. A. Phys. 199 (2013) 30–36, https://doi.org/10.1016/j.sna.2013.04.022.

[4] V. Sherawat, R. Bokolia, R.K. Sinha, Studies on temperature-dependent bandgap and gap-to-midgap ratio in diamond lattice photonic crystal with biosensing applications, Opt. Mater. 145 (2023) 114470, https://doi.org/10.1016/j.optmat.2023.114470.

[5] H.E. Ruda, N. Matsuura, Properties and applications of photonic crystals, in: Opt. Prop. Mater. Their Appl., Wiley, 2019, pp. 251–268, https://doi.org/10.1002/9781119506003.ch9.

[6] E. Dimopoulos, A. Sakanas, A. Marchevsky, M. Xiong, Y. Yu, E. Semenova, J. Mørk, K. Yvind, Electrically-Driven photonic crystal lasers with ultra-low threshold, Laser Photon. Rev. 16 (2022), https://doi.org/10.1002/lpor.202200109.

[7] B. Vakili, S. Bahadori-Haghighi, R. Ghayour, All-optical switching using a new photonic crystal directional coupler, Adv. Electromagn. 4 (2015) 63–67, https://doi.org/10.7716/aem.v4i1.307.

[8] J. Broeng, T. Søndergaard, S.E. Barkou, P.M. Barbeito, A. Bjarklev, Waveguidance by the photonic bandgap effect in optical fibres, J. Opt. Pure Appl. Opt. 1 (1999) 477–482, https://doi.org/10.1088/1464-4258/1/4/311.

[9] T. Baba, Slow light in photonic crystals, Nat. Photonics 2 (2008) 465–473, https://doi.org/10.1038/nphoton.2008.146.

[10] H.S. Dutta, A.K. Goyal, V. Srivastava, S. Pal, Coupling light in photonic crystal waveguides: a review, Photon. Nanostruct: Fundam. Appl. 20 (2016) 41–58, https://doi.org/10.1016/j.photonics.2016.04.001.

[11] K. Okamoto, Coupled mode theory, in: Fundam. Opt. Waveguides, Elsevier, 2022, pp. 169–217, https://doi.org/10.1016/B978-0-12-815601-8.50004-5.

[12] P.T. Kristensen, J.R. de Lasson, M. Heuck, N. Gregersen, J. Mork, On the theory of coupled modes in optical cavity-waveguide structures, J. Lightwave Technol. 35 (2017) 4247–4259, https://doi.org/10.1109/JLT.2017.2714263.

[13] W. Huang, Y. Zhang, Y. Pei, M. Wang, F. Shi, K. Li, Effects of propagation phase on the coupling of plasmonic optical modes, Chin. Opt Lett. 21 (2023) 010003, https://doi.org/10.3788/COL202321.010003.

[14] F. Diebel, D. Leykam, M. Boguslawski, P. Rose, C. Denz, A.S. Desyatnikov, All-optical switching in optically induced nonlinear waveguide couplers, Appl. Phys. Lett. 104 (2014), https://doi.org/10.1063/1.4886414.

[15] R. Soref, B. Bennett, Electrooptical effects in silicon, IEEE J. Quant. Electron. 23 (1987) 123–129, https://doi.org/10.1109/JQE.1987.1073206.

[16] S. Kuchinsky, V.Y. Golyatin, A.Y. Kutikov, T.R. Pearsall, D. Nedelikovic, Coupling between photonic crystal waveguides, IEEE J. Quant. Electron. 38 (2002) 1349–1352, https://doi.org/10.1109/JQE.2002.802954.

[17] C.M. de Sterke, D.G. Salinas, J.E. Sipe, Coupled-mode theory for light propagation through deep nonlinear gratings, Phys. Rev. E. 54 (1996) 1969–1989, https://doi.org/10.1103/PhysRevE.54.1969.

[18] P.E. Barclay, K. Srinivasan, O. Painter, Design of photonic crystal waveguides for evanescent coupling to optical fiber tapers and integration with high-Q cavities, J. Opt. Soc. Am. B 20 (2003) 2274, https://doi.org/10.1364/JOSAB.20.002274.

[19] M.K. Chin, S.T. Ho, Design and modeling of waveguide-coupled single-mode microring resonators, J. Lightwave Technol. 16 (1998) 1433–1446, https://doi.org/10.1109/50.704609.

[20] C. Manolatou, M.J. Khan, S. Fan, P.R. Villeneuve, H.A. Haus, J.D. Joannopoulos, Coupling of modes analysis of resonant channel add-drop filters, IEEE J. Quant. Electron. 35 (1999) 1322–1331, https://doi.org/10.1109/3.784592.

[21] D.L. Lee, Electromagnetic Principles of Integrated Optics, Wiley, New York, 1986.

[22] P. Orlandi, F. Morichetti, M.J. Strain, M. Sorel, A. Melloni, P. Bassi, Tunable silicon photonics directional coupler driven by a transverse temperature gradient, Opt. Lett. 38 (2013) 863, https://doi.org/10.1364/ol.38.000863.

[23] D.M. Beggs, T.P. White, L. O'Faolain, T.F. Krauss, Ultracompact and low-power optical switch based on silicon photonic crystals, Opt. Lett. 33 (2008) 147, https://doi.org/10.1364/ol.33.000147.

[24] N. Yamamoto, T. Ogawa, K. Komori, Photonic crystal directional coupler switch with small switching length and wide bandwidth, Opt. Express 14 (2006) 1223, https://doi.org/10.1364/OE.14.001223.

[25] L.M. Ali, Coupling length and coupling loss in AlGaAs photonic crystal waveguides, J. Opt. 46 (2017) 187–190, https://doi.org/10.1007/s12596-016-0367-9.

[26] L.B. Soldano, E.C.M. Pennings, Optical multi-mode interference devices based on self-imaging: principles and applications, J. Lightwave Technol. 13 (1995) 615–627, https://doi.org/10.1109/50.372474.
[27] Y. Nagpal, R.K. Sinha, Modeling of photonic band gap waveguide couplers, Microw. Opt. Technol. Lett. 43 (2004) 47–50, https://doi.org/10.1002/mop.20371.
[28] A. Yariv, Coupled-mode theory for guided-wave optics, IEEE J. Quant. Electron. QE-9 (1973).
[29] L. Torrijos-Moran, A. Brimont, A. Griol, P. Sanchis, J. Garcia-Ruperez, Ultra-compact optical switches using slow light bimodal silicon waveguides, J. Lightwave Technol. 39 (2021) 3495–3501, https://doi.org/10.1109/JLT.2021.3066479.
[30] B. Ni, J. Xiao, Compact and broadband silicon-based mode-division (de) multiplexer using an asymmetrical directional coupler, Opt. Commun. 451 (2019) 141–146, https://doi.org/10.1016/j.optcom.2019.06.043.
[31] L.-Y. Chiang, C.-T. Wang, S. Pappert, P.K.L. Yu, Silicon–organic hybrid thermo-optic switch based on a slot waveguide directional coupler, Opt. Lett. 47 (2022) 3940, https://doi.org/10.1364/OL.467858.
[32] C. Zhang, M. Zhang, Y. Xie, Y. Shi, R. Kumar, R.R. Panepucci, D. Dai, Wavelength-selective 2 × 2 optical switch based on a Ge 2 Sb 2 Te 5 -assisted microring, Photonics Res. 8 (2020) 1171, https://doi.org/10.1364/PRJ.393513.
[33] R. Chen, Z. Fang, J.E. Fröch, P. Xu, J. Zheng, A. Majumdar, Broadband nonvolatile electrically controlled programmable units in silicon photonics, ACS Photonics 9 (2022) 2142–2150, https://doi.org/10.1021/acsphotonics.2c00452.
[34] X. Yang, M.S. Nisar, W. Yuan, F. Zheng, L. Lu, J. Chen, L. Zhou, Phase change material enabled 2 × 2 silicon nonvolatile optical switch, Opt. Lett. 46 (2021) 4224, https://doi.org/10.1364/OL.435552.
[35] J. He, J. Yang, H. Ma, X. Jiang, H. Yuan, Y. Yu, Design of a multi-functional integrated optical switch based on phase change materials, Photonics 9 (2022) 320, https://doi.org/10.3390/photonics9050320.
[36] Y. Xie, L. Chen, H. Li, Y. Yi, Polymer and hybrid optical devices manipulated by the thermo-optic effect, Polymers 15 (2023) 3721, https://doi.org/10.3390/polym15183721.
[37] K. Miao, S. Xu, Design of a compact terahertz polarization beam splitter with ultralow loss based on twin butterfly porous core photonic crystal fiber, J. Nanophotonics 19 (2025), https://doi.org/10.1117/1.JNP.19.016001.
[38] C. Liu, Y. Sun, Y. Zeng, W. Liu, J. Lv, L. Yang, J. Wang, Q. Liu, P.K. Chu, An ultra-high extinction ratio broadband photonic crystal fiber splitter designed for the terahertz band, Opt. Fiber Technol. 90 (2025) 104136, https://doi.org/10.1016/j.yofte.2025.104136.
[39] A. Ghatak, K. Thyagarajan, Introduction to Fiber Optics, Cambridge University Press, Cambridge, 1998, https://doi.org/10.1017/CBO9781139174770.

Towards Creating an Ecosystem of Paperless Society: Unveiling Issues and Challenges in Adoption of DigiLocker

Rajan Yadav*

ABSTRACT

DigiLocker, a free document storage and sharing cloud-based platform, is an e-governance project of the Indian government under its Digital India initiative. Though the government has been working to make this initiative successful, a considerable portion of the population is unaware of the service. The few aware of this initiative have yet to opt for a DigiLocker. The present study attempts to identify the issues and challenges that affect the adoption and use of these services designed to create a paperless society. The proposed framework of the study is based on the theoretical underpinning of the Technology Adoption Model (TAM), the Unified Theory of Acceptance and Use of Technology (UTAUT) and the Unified Model of Electronic Government Adoption (UMEGA). The study analysed the data using PLS-SEM, including the moderation impact of select variables. The study identifies various factors that act as motivators and deterrents in adopting and using DigiLocker and provides strong empirical support to frame suitable strategic interventions to sensitize people to improve further usage and adoption.

Keywords: *DigiLocker, technology adoption, e-governance, PLS-SEM*

INTRODUCTION

The Digital India Programme of the Government of India envisions making India a digitally empowered society and knowledge economy with an emphasis on digital empowerment of citizens by improving e-governance, services-on-demand, and digital infrastructure. Sustained efforts are being made to create collaborative linkages on various citizen-centric projects to transform the traditional service delivery process into a seamlessly integrated process that is available on a real-time basis to all citizens in a transparent, convenient, and secure manner. Over the years, various State and Central Government initiatives have been undertaken to navigate into an era of e-governance. Persistent efforts have been made at government and grassroots levels to improve public service delivery and simplify access to it.

The revolution of information technology has penetrated almost all domains. The increasing use of cloud platforms has provided individuals with ease of access. DigiLocker is one such scheme launched by the government of India, which provides free 1 GB cloud storage to store documents in digital form for the citizens of India. This initiative is a part of the government's Digital India initiative. Documents such as PAN cards, Passports, driving licenses, etc, can be stored in DigiLocker and carried easily by individuals. Therefore, such an initiative by the government of India has provided ease to storing all such documents in the digital locker, which is a cloud-based space for keeping documents in digital form. This initiative is widespread through newspapers and electronic media. The Central Board of Secondary Education has opened 80 lakh digital

* *Professor, Delhi Technological University, Delhi, rajanyadav@dtu.ac.in*

lockers for the students of 10th and 12th class. However, only 2,06,57,736 registered users and 116 issuer organizations of DigiLocker exist to date. Maharashtra has the highest number of registered users of DigiLocker, whereas Delhi ranks 11th, with only 2,73,658 registered users. States like Sikkim, Mizoram and Chhattisgarh have the lowest registrations to date. Therefore, there is a need to study the reasons for not adopting a free service which ensures the safe storage of documents. The present study is therefore aimed to understand the level of awareness of individuals for DigiLocker and to explore the various factors which affect the technology adoption process.

2. Objectives of the study

- To identify the motivators and inhibitors of low adoption and usage of DigiLocker.
- To analyse the impact of social, demographic, economic and geographic variation in the adoption of DigiLocker.
- To propose an empirical, evidence-based strategic framework to improve the adoption and usage of DigiLocker.
- To suggest a few measures to policymakers for widespread awareness of DigiLocker.

3. Literature Review

User Acceptance of Information Technology (UAIT) is a prerequisite for effectively implementing any information technology-based project (AlAwadhi et al., A., 2008). UAIT is the "initial decision made by the individual to interact with the technology" (Venkatesh et al., 2004).

UAIT is followed by adoption, defined as the "direct experience with the technology after an individual has decided to accept it" (Venkatesh et al., 2004).

Most of the research in the information technology acceptance field spans around the path-breaking work of Davis (1989). He proposed the TAM, the most popular model for explaining the intention and behaviour of individuals for technology acceptance. TAM theorizes that two precursors form behavioural intention: perceived usefulness and ease of use (Venkatesh & Davis, 2000). Perceived usefulness is defined as the extent to which a user perceives that technology adoption will enhance his/her performance. On the other hand, perceived ease of use revolves around minimizing efforts for using a technology. It was also proposed that perceived ease of use affects perceived usefulness since the easier a technology is, the more useful it can be (Venkatesh, 2000). The model was parsimonious and has been applied in several studies over the last three decades. It has been used to study faculty members' intention to use learning management systems (LMS) (Fathema et al., 2015), consumer acceptance of e-commerce "" and acceptance of E-learning in developing countries (Tarhini et al., 2017).

Another model that is significant in understanding technology adoption was given by Venkatesh, Morris, Davis, and Davis (2003), named the UTAUT model. It consists of four variables: effort expectancy, performance expectancy, social influence and facilitating conditions. It also captures the moderating influence of gender, age, experience, and voluntariness of use (Im et al., 2011). The UTAUT model rationalises that effort expectancy is significant in users' adoption of information technology. The model explains the 70 per cent variance, which is much higher than TAM. This model is said to be more synthesized and complete since it has merged eight different models discussed in information system literature. These models (TRA, TAM, TPB, the Motivational Model (MM) (Davis et al., 1992), the combined TAM and TPB (C-TAM-TPB) (Taylor et al., 1995), the Model of PC Utilization (MPCU) (Thompson et al., 1991) DOI and Social Cognitive Theory (SCT)) had their origins in sociology and psychology. UTAUT explains how acceptance of information technology differs across various individuals. The moderators used in the model also emphasize the differences in intention to use technology in terms of different age groups and gender.

Though digital services are provided in various countries, there are challenges in usage and adoption. Developing countries, particularly, face a lack of infrastructure, awareness and technical skills required to

adopt ICT (Sivathanu, 2018). The DigiLocker initiative results from cloud computing technology (Khandekar & Devadiga, 2016); it enables the virtual storage of various pictures and documents. This initiative is under the Digital India government scheme. It has helped to provide a digital space wherein citizens of India can store documents issued by the government and other public authorities. A top priority project of the Hon. Narendra Modi Administration, the completed Digital India initiative has specific plans aimed at restructuring the National Informatics Centre. (Petare et al., 2015).

3.1. Conceptual Model

Figure 1 below provides the proposed relationships among the dependent and independent variables. It also shows the moderating variables which have been tested for their contingent effect on the relationships among the constructs. Each relationship (arrow) in the conceptual model represents a hypothesis tested using PLS-SEM.

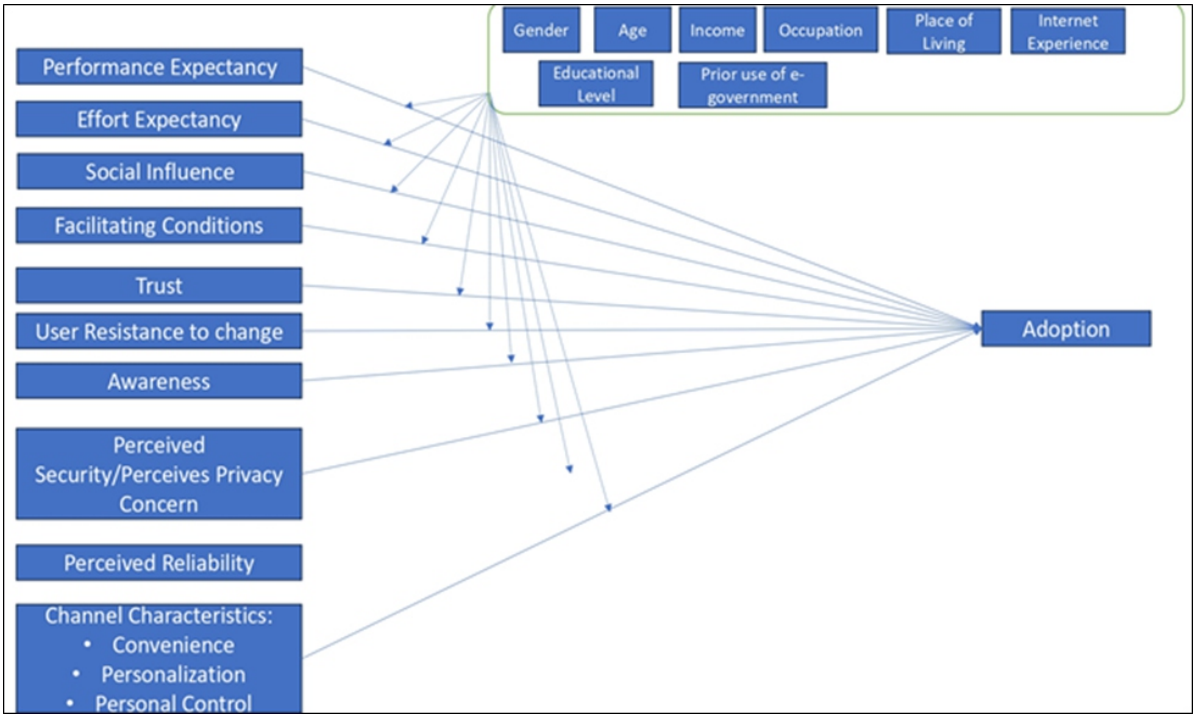


Figure 1 : Conceptual Model

4. Research Methodology

The target population for our study included Indian citizens residing within the country, aged 15 and above, who regularly used essential IT-enabled services. A Google survey form was administered across various social media platforms to collect data to ensure broad distribution and reach within the target audience. The study employs a cross-sectional descriptive research design. We have captured data at a specific time, providing a snapshot of the population's perceptions and behaviours.

This study utilised a non-probability convenience sampling technique to reach respondents easily accessible to the researchers, resulting in a final sample size of 145 participants. We analysed data using Partial Least Squares-Structural Equation Modelling (PLS-SEM), a method suitable for complex analyses with relatively smaller minor samples. Smart-PLS 4 software was employed for analysis, enabling the research team to interpret relationships and draw meaningful insights from the collected data.

5. Data Analysis

5.1. Measurement Instrument

Items were taken from extant research to maintain the validity of the content of the model. These items were modified to adopt the context of the adoption of DigiLocker. The table below displays the constructed items as well as their sources.

Table 1: Construct Items

S. No.	Factors	Citation
1.	Performance Expectancy	(AlHadid et al., 2021) (Sharma et al., 2018)
2.	Effort Expectancy	(AlHadid et al., 2021) (Sharma et al., 2018)
3.	Social Influence	(Ahmad & Khalid, 2017) (AlHadid et al., 2021) (Sharma et al., 2018)
4.	Facilitating Conditions	(AlHadid et al., 2021) (Sharma et al., 2018)
5.	Trust	(Shahzad et al., 2020) (Ahmad & Khalid, 2017) (Venkatesh et al., 2016) (Sharma et al., 2018)
6.	User Resistance to change	(Talukdar et al., 2020)
7.	Perceived Awareness	(Shareef et al., 2010) (Shahzad et al., 2020)
8.	Perceived Security	(Shareef et al., 2010)
9.	Perceived Reliability	(Shahzad et al., 2020)
10.	Channel Characteristics	
	Convenience	(Shahzad et al., 2020) (Venkatesh et al., 2016)
	Personalization Personal/Active Control	(Shahzad et al., 2020) (Venkatesh et al., 2016) (Chen et al., 2016)
11.	Prior use of e-government services	(Hou et al., 2019)
	Intention to use	(Shahzad et al., 2020)

5.2. Model assessment in Smart PLS 4

Each item's construct reliability and validity were assessed to test the measurement model. Cronbach's α, Composite Reliability (CR) and Average Variance Extracted (AVE) were computed to assess the reliability and convergent validity **(Table 2)**. We calculated the Variance Inflation Factor (VIF) to check the multicollinearity between variables. The threshold limit of Cronbach's α is 0 to 1, and values above 0.7 suggest an acceptable level of reliability. In our study, Cronbach's α varies from 0.879 to 0.962, under the threshold limit. The threshold limit of composite reliability is that the value should be greater than 0.7.

In our study, Composite Reliability varies from 0.919 to 0.969, which is under the threshold limit. We computed the AVE test to assess the convergent validity. The threshold limit for AVE is that the value should be greater than 0.5. In our study, AVE ranges from 0.714 to 0.887, under the threshold limit.

The study assessed the multi-collinearity among the constructs in the model using VIF. Many research frameworks based on OLS Regression observe multi-collinearity among two or more constructs demonstrating a linear relationship. In such cases, the values of different path coefficients may be consistent. However, the results may not be robust and reliable due to high VIF values, which indicate high muti-collinearity. A VIF value above 4 is acceptable and an indicator of low multi-collinearity. The study found the VIF values in all the relationships as per the thumb rule i,e. < 4 permitting to proceed further with the analysis.

Our study's VIF values range from 1 to 8.919, indicating varying degrees of multicollinearity between variables. In our study, some values are greater than 5 because the sample size we chose is small, and the variables are homogeneous. Several variables demonstrate high levels of multicollinearity, with PE2 showing the highest VIF at 8.919, followed closely by T5 at 7.987, PE1 at 7.7, T6 at 7.148, FC1 at 7.021, and EE4 at 7.548. A moderate level of correlation is present in several variables, including CC1, CC2, and CC3 (all around 6), A2 (5.381), PE3 (6.855), SI5 (6.034), FC2 (6.211), and T4 (5.411). All interaction terms of construct factors with daily internet usage and education level show VIF values 1, indicating no multicollinearity issues in these relationships and no correlation.

We computed the Heterotrait-Monotrait (HTMT) ratio to assess the discriminant validity. The threshold limit for the HTMT ratio is that values should be below 0.90. In our study, the HTMT values vary from 0.007 to 0.938 **(Table 3)**. To evaluate the model's fit and determine how well the model explains the observed data, we computed the R2 value. If R2 value is greater than 0.50, it indicates strong explanatory power. In our study, the R2 value is 0.801 **(Table 4)**.

Table 2: Composite Reliability

	Cronbach's α	CR (rho_c)	AVE
A	0.936	0.959	0.887
CC	0.948	0.957	0.714
EE	0.962	0.969	0.84
FC	0.879	0.919	0.742
PA	0.923	0.942	0.765
PE	0.949	0.96	0.799
PR	0.934	0.958	0.884
PS	0.939	0.956	0.846
SI	0.934	0.948	0.754
T	0.951	0.961	0.804
UR	0.901	0.93	0.77

Table 3: HTMT Ratios

| | A | CC | Daily interEE | Education FC | PA | PE | PR | PS | SI | T | UR | Education | Daily interEducation | Daily interEducation | Education | Education | Education | Daily interEducation | Daily interEducation | Daily interEducation

αTable 4: R-Squared Values

	R-square
A	0.801

1.1. Hypothesis testing

To check the significant relationship between variables, we run bootstrapping with 5000 samples. We look at the path coefficient (i.e. P- value) to check for a significant relation between variables. To have a significant relation between values, the p-value should be less than 0.005.

This is illustrated in **Table 5** and **Figure 2**; findings suggest that significant factors help in the Adoption of DigiLocker. We opt for education level and daily internet usage as a key moderator influencing the adoption of DigiLocker.

From the findings, we discovered that effort expectancy, facilitating conditions and performance expectancy significantly impact the adoption of DigiLocker. EE -> A (p value 0.019), **FC -> A** (p value 0) and **PE -> A** (p value 0.003). Furthermore, we also discovered that education level and daily internet usage have not moderately influenced the adoption of DigiLocker.

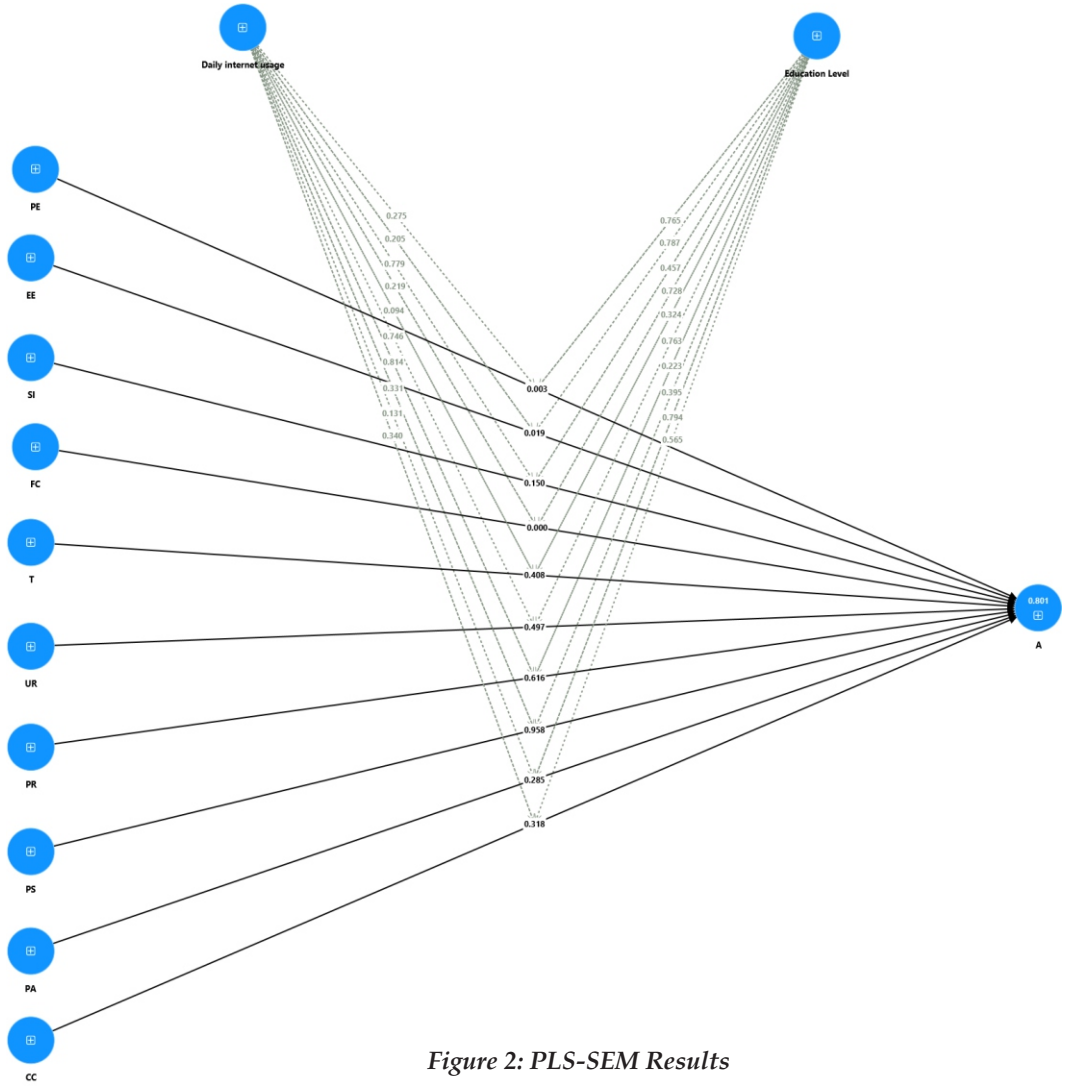


Figure 2: PLS-SEM Results

Table 5: Structural model outcomes

	Original sample (O)	Sample mean (M)	Standard deviation (STDEV)	T statistics (|O/STDEV|)	P values
CC -> A	0.148	0.133	0.149	0.999	0.318
Daily internet usage -> A	-0.048	-0.045	0.061	0.776	0.438
EE -> A	-0.308	-0.3	0.131	2.355	0.019
Education Level -> A	-0.027	-0.024	0.068	0.399	0.69
FC -> A	0.594	0.606	0.133	4.478	0
PA -> A	-0.177	-0.188	0.165	1.07	0.285
PE -> A	0.376	0.373	0.126	2.989	0.003
PR -> A	0.093	0.092	0.185	0.502	0.616
PS -> A	0.01	0.011	0.187	0.052	0.958
SI -> A	0.245	0.255	0.17	1.441	0.15
T -> A	-0.135	-0.116	0.163	0.827	0.408
UR -> A	0.124	0.1	0.183	0.679	0.497
Education Level x FC -> A	0.051	0.03	0.146	0.348	0.728
Daily Internet usage x PE -> A	0.141	0.125	0.129	1.092	0.275
Daily Internet usage x T -> A	-0.307	-0.288	0.184	1.674	0.094
Education Level x UR -> A	0.056	0.103	0.186	0.302	0.763
Daily Internet usage x CC -> A	0.16	0.167	0.168	0.955	0.34
Daily Internet usage x PR -> A	-0.05	-0.035	0.213	0.236	0.814
Education Level x EE -> A	0.038	0.03	0.141	0.27	0.787
Education Level x T -> A	-0.192	-0.145	0.195	0.986	0.324
Daily Internet usage x FC -> A	0.171	0.132	0.139	1.229	0.219
Daily Internet usage x EE -> A	-0.169	-0.142	0.134	1.266	0.205
Education Level x PR -> A	-0.211	-0.233	0.173	1.219	0.223
Education Level x SI -> A	0.101	0.136	0.136	0.743	0.457
Education Level x PA -> A	0.045	-0.015	0.172	0.261	0.794
Education Level x PE -> A	-0.033	-0.033	0.11	0.299	0.765
Daily internet usage x UR -> A	0.057	0.053	0.175	0.324	0.746
Daily internet usage x SI -> A	-0.064	-0.035	0.228	0.281	0.779
Education Level x CC -> A	0.088	0.075	0.153	0.576	0.565
Daily internet usage x PA -> A	-0.259	-0.271	0.171	1.511	0.131
Daily internet usage x PS -> A	0.245	0.215	0.252	0.972	0.331
Education Level x PS -> A	0.194	0.187	0.228	0.85	0.395
a					

6. Managerial Implications

The study is highly relevant for policymakers since the initiative for DigiLocker is based on having more paperless transactions. However, the response to DigiLocker so far is lukewarm. Therefore, all stakeholders in this project must understand the problems individual citizens face while adopting this application, their threats and challenges. The study's findings provide input to policymakers for designing awareness campaigns to spread this initiative to various parts of the country. They also put forward the factors that affect the technology adoption process in a developing country like India. A study of such magnitude is highly relevant for society because it will serve as a barometer to understand the underlying reasons for non-acceptance of a technology-based platform despite being offered free of cost. The present study has also proposed a practical framework to harness the utility of DigiLocker. Hence, in the long run, such findings will help to create better policies for the masses in terms of improvement in IT infrastructure and other vital areas.

7. Acknowledgement: The author acknowledges with thanks Indian Council of Social Science Research (ICSSR) for providing the financial support in writing this research paper under major research project.

BIBLIOGRAPHY:

- *AlAwadhi, S., & Morris, A. (2008, January). The use of the UTAUT model in the adoption of e-government services in Kuwait. In Proceedings of the 41st annual Hawaii International Conference on System Sciences (HICSS 2008) (pp. 219-219). Ieee.*
- *Ahmad, S. Z., & Khalid, K. (2017). The adoption of M-government services from the user's perspectives: Empirical evidence from the United Arab Emirates. International Journal of Information Management, 37(5), 367–379.*
- *AlHadid, I., Abu-Taieh, E., Alkhawaldeh, R. S., Khwaldeh, S., Masa'deh, R., Kaabneh, K., & Alrowwad, A. (2022). Predictors for E-Government Adoption of SANAD App Services Integrating UTAUT, TPB, TAM, Trust, and Perceived Risk. International Journal of Environmental Research and Public Health, 19(14), 8281.*
- *Anon. 2003. "Consumer Acceptance of Electronic Commerce: Integrating Trust and Risk with the Technology Acceptance Model." International Journal of Electronic Commerce 7(3):101–34.*
- *Chen, Z.-J., Vogel, D., & Wang, Z.-H. (2016). How to satisfy citizens? Using mobile government to reengineer fair government processes. Decision Support Systems, 82, 47–57. https://doi.org/10.1016/j.dss.2015.11.005*
- *Davis, F. D. (1989). Perceived usefulness, perceived ease of use and user acceptance of information technology. MIS Quarterly, 13(3), 319–340.*
- *Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and intrinsic motivation to use computers in the workplace 1. Journal of applied social psychology, 22(14), 1111-1132.*
- *Fathema, N., Shannon, D., & Ross, M. (2015). Expanding the Technology Acceptance Model (TAM) to examine faculty use of Learning Management Systems (LMSs) in higher education institutions. Journal of Online Learning & Teaching, 11(2).*
- *Hou, J. (Jove), Arpan, L., Wu, Y., Feiock, R., Ozguven, E., & Arghandeh, R. (2020). The Road toward Smart Cities: A Study of Citizens' Acceptance of Mobile Applications for City Services. Energies, 13(10), 2496*
- *Im, I., Hong, S., & Kang, M. S. (2011). An international comparison of technology adoption: Testing the UTAUT model. Information & management, 48(1), 1-8.*
- *Khandekar, C., & Devadiga, N. (2016). VoteItRight: Digital Empowerment of Citizens. International Journal of Computer Applications, 154(5).*
- *Pavlou, P. A. (2003). Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. International journal of electronic commerce, 7(3), 101-134.*
- *Sivathanu, B. (2018). An Empirical Study of Cloud-Based E-Governance Services Adoption in India. International Journal of Electronic Government Research (IJEGR), 14(1), 86-107.*
- *Shareef, M. A., Kumar, V., Kumar, U., & Dwivedi, Y. K. (2011). e-Government Adoption Model (GAM): Differing service maturity levels. Government Information Quarterly, 28(1), 17–35.*
- *Sharma, S. K., Al-Badi, A., Rana, N. P., & Al-Azizi, L. (2018). Mobile applications in government services (mG-App) from user's perspectives: A predictive modelling approach. Government Information Quarterly, 35(4), 557–568.*

- *Petare, P., Mohite, P., & Joshi, M. (2015). DigiLocker (Digital Locker-Ambitious Aspect of Digital India Programme.). Ge-International Journal Of Management Research, 3, 2321-1709.*
- *Tarhini, A., Hone, K., Liu, X., & Tarhini, T. (2017). Examining the moderating effect of individual-level cultural values on users' acceptance of E-learning in developing countries: a structural equation modelling of an extended technology acceptance model. Interactive Learning Environments, 25(3), 306-328.*
- *Taylor, S., & Todd, P. A. (1995). Understanding information technology usage: A test of competing models. Information systems research, 6(2), 144-176.*
- *Thompson, R. L., Higgins, C. A., & Howell, J. M. (1991). Personal computing: toward a conceptual model of utilization. MIS quarterly, 125-143.*
- *Venkatesh, V. (2000). Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology acceptance model. Information systems research, 11(4), 342-365.*
- *Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. Management science, 46(2), 186-204.*
- *Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS quarterly, 425-478.*
- *Venkatesh, V., Morris, M. G., Sykes, T. A., & Ackerman, P. L. (2004). Individual reactions to new technologies in the workplace: The role of gender as a psychological construct. Journal of Applied Social Psychology, 34(3), 445-467.*

arXiv:2504.16788v1 [cs.CV] 23 Apr 2025

Towards Explainable AI: Multi-Modal Transformer for Video-based Image Description Generation

Lakshita Agarwal[1] and Bindu Verma[2*]

[1]Department of Information Technology, Delhi Technological University, Delhi, 110042, India.

[2*]Assistant Professor, Department of Information Technology, Delhi Technological University, Delhi, 110042, India.

*Corresponding author(s). E-mail(s): bindu.cvision@gmail.com;
Contributing authors: lakshitaa3@gmail.com;

Abstract

Understanding and analyzing video actions are essential for producing insightful and contextualized descriptions, especially for video-based applications like intelligent monitoring and autonomous systems. The proposed work introduces a novel framework for generating natural language descriptions from video datasets by combining textual and visual modalities. The suggested architecture makes use of ResNet50 to extract visual features from video frames that are taken from the Microsoft Research Video Description Corpus (MSVD), and Berkeley DeepDrive eXplanation (BDD-X) datasets. The extracted visual characteristics are converted into patch embeddings and then run through an encoder-decoder model based on Generative Pre-trained Transformer-2 (GPT-2). In order to align textual and visual representations and guarantee high-quality description production, the system uses multi-head self-attention and cross-attention techniques. The model's efficacy is demonstrated by performance evaluation using BLEU (1-4), CIDEr, METEOR, and ROUGE-L. The suggested framework outperforms traditional methods with BLEU-4 scores of 0.755 (BDD-X) and 0.778 (MSVD), CIDEr scores of 1.235 (BDD-X) and 1.315 (MSVD), METEOR scores of 0.312 (BDD-X) and 0.329 (MSVD), and ROUGE-L scores of 0.782 (BDD-X) and 0.795 (MSVD). By producing human-like, contextually relevant descriptions, strengthening interpretability, and improving real-world applications, this research advances explainable AI.

Keywords: Video-based Description Generation, Transformer Models, Explainable AI, Autonomous Systems, Natural Language Processing

1 Introduction

In computer vision and natural language processing, creating informative textual descriptions of events that occur in video sequences is a fundamental problem known as "video description generation" [1]. In contrast to static image captioning, video-based description generation involves an awareness of object interactions, temporal dependencies, and contextual linkages in order to provide narratives that are both contextually rich and coherent. Applications like intelligent surveillance, multimedia retrieval, autonomous driving, and human-computer interaction greatly benefit from these capabilities. Capturing dynamic scene

changes, identifying pertinent objects, and producing descriptions that appropriately reflect the changing context are the challenges [2] [3]. Conventional methods for comprehending videos, such as object detection and action recognition, mostly concentrate on figuring out "what" is happening in a scene. Explaining "why" things happen is more difficult, though, particularly in complex, multi-agent systems where behaviour is influenced by environmental suggestions [4]. Known as video captioning or video-based description generation, the process combines textual and visual modalities to generate interpretable and informative natural language descriptions [5] [6]. The ability to produce context-aware and semantically relevant video descriptions has greatly improved with recent developments in deep learning, especially transformer-based systems.

This work presents a novel transformer-based architecture for generating video descriptions by effectively combining text and image modalities. Using a GPT-2 encoder-decoder with multi-head attention and ResNet50 for visual feature extraction, the model is trained on diverse datasets, including MSVD and BDD-X. Training is optimized with gradient accumulation and mixed precision to boost efficiency. Evaluated using BLEU (1–4), CIDEr, METEOR, and ROUGE-L, the model produces fluent, contextually relevant descriptions aligned with human annotations, contributing to advancements in video explainability and AI-driven understanding.

The major contributions of the proposed work are summarized below:

- To provide natural language descriptions of video sequences, the following work proposes a transformer-based architecture that integrates a GPT-2-based language model with the visual features retrieved by ResNet50.
- The method combines video datasets from various sources (MSVD, and BDD-X), allowing the model to generalise across domains. High-quality training samples are guaranteed via sophisticated preprocessing.
- Gradient accumulation and mixed precision training are used to optimise the model, increasing computing efficiency without sacrificing output quality.
- Fluency, contextual relevance, and coherence are ensured by thoroughly evaluating the generated descriptions using BLEU (1–4), CIDEr, METEOR, and ROUGE-L standards.
- Although this method is extremely beneficial for intelligent transportation and autonomous driving, it can also be used in other fields, including robotics, assistive technology, and surveillance.

The remaining sections of the paper are organized as follows: Section 2 examines relevant studies in multimodal learning and video description generation. The suggested methodology, including feature extraction, model architecture, and dataset preprocessing, is described in Section 3. Experimental results and assessment results are presented in Section 4. Lastly, conclusions and possible future study directions are covered in Section 5.

2 Related Work

In computer vision and natural language processing, video-based description creation is an essential task that allows systems to produce textual summaries of visual content. The creation of video descriptions must consider temporal relationships, object interactions, and dynamic scene changes across frames, in contrast to static image captioning. Early video description generation approaches included recurrent neural networks (RNNs) or long short-term memory (LSTM) networks for text synthesis and convolutional neural networks (CNNs) for feature extraction.

A Sequence-to-Sequence (Seq2Seq) model trained on MSVD [7] was presented by Venugopalan et al. [8] to produce video descriptions. However, because of the limits of LSTMs, these models have trouble handling contextual inconsistencies and long-range dependencies. Later, to make generated descriptions more relevant, attention techniques were added [9]. Transformer-based models have led to a considerable improvement in video captioning systems. End-to-end dense Video Captioning was proposed by Zhou et al. [10], who used spatiotemporal attention mechanisms to grasp multiple frames. Compared to conventional RNN-based models, MART (Memory-Augmented Recurrent Transformer), which was introduced by Lei et al. [11], performs more

effectively at capturing long-term video dependencies. More recently, multimodal pretraining for video-language understanding has been investigated by VideoBERT [12] and ClipBERT [13]. These methods achieve state-of-the-art outcomes on datasets such as MSVD [7] and MSR-VTT [14] by aligning textual and visual representations. Datasets such as Berkeley DeepDrive eXplanation (BDD-X) [15] have been useful for producing explainable AI-driven annotations for vehicle-based video description generation. Studies like Shoman et al. [4] integrated explainability processes into vision-language models and concentrated on justification-based video descriptions for autonomous driving. An attention-guided transformer was developed by Cui et al. [16] to improve contextual reasoning in captioning models based on BDD-X.

Recent advancements in multimodal learning combine text, audio, and visual data. Hori et al.[17] introduced an audio-visual attention model using speech and scene context, while Yu et al.[18] enhanced captioning on datasets like Flickr8k and MSVD using spatiotemporal object graphs. These methods improve contextual awareness, particularly in dynamic scenarios like vehicle interactions. Despite progress, generating accurate, human-like video descriptions remains challenging. The proposed approach addresses this by combining a GPT-2 encoder-decoder with ResNet50-based feature extraction, outperforming prior methods on BLEU (1–4), CIDEr, METEOR, and ROUGE-L. It refines multimodal embeddings for context-aware captions and boosts efficiency with gradient accumulation and mixed precision training. By bridging explainable vehicular (BDD-X) and general (MSVD) video captioning, the model delivers interpretable, context-rich descriptions suitable for real-world use.

3 Overview of the Proposed Work

The proposed study introduces a transformer-based, context-aware reasoning framework for generating logical and relevant video descriptions. Leveraging the BDD-X and MSVD datasets, the system learns from both action-justified driving videos and diverse scene representations. It employs an optimized GPT-2 encoder-decoder model with multi-head attention and ResNet50-based visual feature extraction to enhance contextual understanding. Training is guided by action-justification pairs from BDD-X and captioned sequences from MSVD, promoting coherent, human-like output. Gradient accumulation and mixed precision training boost computational efficiency without compromising accuracy. Evaluations using BLEU (1–4), CIDEr, METEOR, and ROUGE-L show superior performance over traditional methods. The basic architecture of the proposed framework is shown in Figure 1, which further demonstrates the pipeline for data pre-processing, model training, and description generation.

3.1 Data Pre-processing

The input data includes video frames from the BDD-X and MSVD datasets. Video frames are extracted, pertinent images are filtered, and then the frames are transformed into structured input for the model as part of the pre-processing pipeline. ResNet50, which converts spatial information from images into high-dimensional feature vectors, is used for feature extraction. After that, a linear projection and position encoding method are used to patch and embed these feature vectors. The GPT-2 encoder-decoder module then receives the processed embeddings and aligns the word/-token embeddings with the appropriate visual features. To further improve model resilience, tokenization, data augmentation, and train-test splits are used.

3.2 Model Architecture

The proposed method combines a GPT-2 encoder-decoder model to generate context-aware descriptions of vehicle actions with ResNet50 to extract image features. The model is fine-tuned on the BDD-X and MSVD datasets to become optimal in dynamic scenarios for generating meaningful descriptions. The training process employs gradient accumulation and mixed precision training to achieve optimal computing efficiency.

ResNet50 acts as a feature extractor by passing each input frame $x \in \mathbb{R}^{3\times H\times W}$ through a series of convolutional, batch normalization, ReLU, and residual layers. The output feature vector f is

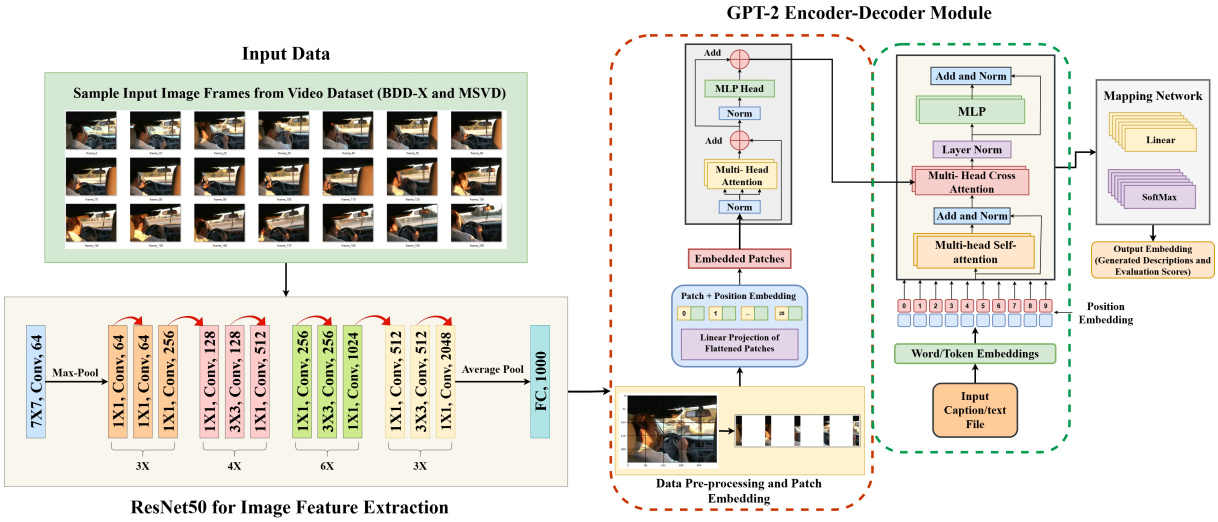


Fig. 1: Framework for the Proposed Model (ResNet50-GPT2): The system incorporates ResNet50 for image feature extraction and a GPT-2 encoder-decoder model to generate context-aware video-based image descriptions.

computed as:

$$f = \text{ResNet50}(x) = \text{AvgPool}(F_{\text{res}}(x)) \quad (1)$$

where, F_{res} denotes the output of the final convolutional block, and AvgPool is the global average pooling operation applied to extract the final feature representation $f \in \mathbb{R}^{2048}$.

Using position embeddings $P \in \mathbb{R}^{N \times d}$ and a learnable linear projection $W_p \in \mathbb{R}^{2048 \times d}$, the visual characteristics retrieved by ResNet50 are transformed into embedded visual tokens:

$$E_v = fW_p + P \quad (2)$$

The GPT-2 model uses these embedded patches as input tokens to produce descriptions that are logical and sensitive to context. The model utilizes self-attention mechanisms to enhance contextual understanding, formulated as:

$$\text{Self-Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (3)$$

where d_k is the key vector dimension, and Q, K, V represent the query, key, and value matrices.

To maintain the semantic relevance of the generated descriptions to video-based activities, the multi-head cross-attention method aligns image embeddings with textual descriptions. The decoder uses attention-weighted contextual embeddings to improve the textual output:

$$\text{Decoder-Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (4)$$

The final word prediction probability is computed as:

$$y_t = \text{Softmax}(W_o h_t + b_o) \quad (5)$$

where W_o and b_o are the output weight matrix and bias, respectively, and h_t represents the hidden state at time step t.

To enhance generalization, the model employs a combined loss function incorporating cross-entropy loss and L2 regularization:

$$\mathcal{L} = -\sum_{t=1}^{T} \log P(y_t \mid y_{<t}) + \lambda \|\theta\|^2 \quad (6)$$

where λ is the regularization parameter and $P(y_t \mid y_{<t})$ denotes the probability of predicting y_t given prior words.

In BLEU-4, CIDEr, METEOR, and ROUGE-L metrics, the proposed framework outperforms traditional methods, ensuring that the generated descriptions are sensible, logical, and context-aware. These metrics are utilized to analyze the

contextual validity and linguistic quality of the descriptions generated by the GPT-2 model for unseen videos upon training. The method achieves performance improvement over traditional methods by contrasting the generated descriptions with ground-truth annotations. Ultimately, the framework generates natural language narratives that effectively describe dynamic visual scenes, promoting explainability in a range of applications.

The following summarizes Algorithm 1 for the proposed model and its evaluation:

Algorithm 1 Context-Aware Description Generation

Input: Video frames (BDD-X, MSVD), image-caption pairs (Flickr8k)
Output: Generated descriptions, Evaluation Scores
1. Preprocessing and Embedding
Extract and embed image frames using patch encoding
Tokenize and format textual descriptions
2. Dataset Split
Create balanced train-test sets with captions and action-justification pairs
3. Feature Extraction
Visual_features $\leftarrow$ ResNet50(I)
Project features and tokenize text using GPT-2
4. Model Training
Initialize GPT-2 with multi-head attention and cross-modal fusion
Train with AdamW, gradient accumulation, and mixed precision
5. Inference and Evaluation
for each test sample **do**
    Generate and decode descriptions
end for
Compute BLEU-4, CIDEr, METEOR, ROUGE-L
6. Output
Present evaluation scores and visualize results

4 Experimental Analysis

The proposed model was implemented in PyTorch and evaluated on Google Colab Pro+ using 52 GB RAM, an NVIDIA A100 GPU (40 GB VRAM), and a virtual Intel Xeon CPU. It generates context-aware descriptions using a GPT-2 encoder-decoder (approximately 66.7M parameters) and ResNet50 for visual feature extraction. Input frames were resized to `(batch_size, 224, 224, 3)`. Training was conducted over 75 epochs with the AdamW optimizer (learning rate 1×10^{-5}, weight decay 0.01), using a batch size of 32 and gradient accumulation. Mixed precision training and gradient clipping were applied to improve computational efficiency and training stability. The model was trained using the MSVD dataset[1], which contains short descriptive captions for diverse video clips, and the BDD-X dataset[2], which includes over 26K annotated actions across 8.4M frames and 6,970 videos. During preprocessing, linguistic descriptions were tokenized and structured into action-justification pairs, while image frames were embedded via ResNet50 and projected linearly into patch embeddings. These embeddings were then fed into the multi-head self-attention module of the GPT-2 encoder-decoder to align image and text contexts effectively. Performance was evaluated on a 20% test split using BLEU-4, CIDEr, METEOR, and ROUGE-L. The model consistently outperformed conventional captioning techniques, demonstrating its ability to generate logical, contextually accurate descriptions across diverse visual inputs. The architecture's use of self- and cross-attention mechanisms ensured semantic accuracy and narrative coherence, reinforcing its applicability in real-world visual captioning tasks. The following Table 1 depicts the overall architectural structure of the proposed framework.

4.1 Ablation Study:

Utilising the BDD-X and the MSVD datasets, an ablation research was carried out to evaluate the contribution of various components in the suggested model. The research methodically eliminated or altered important architectural components, such as multi-head attention, gradient accumulation and mixed precision training optimisations, and visual feature extraction (ResNet-50), in order to assess each component's effect

[1] https://www.kaggle.com/datasets/vtrnanh/msvd-dataset-corpus
[2] https://github.com/JinkyuKimUCB/BDD-X-dataset

Table 1: Experimental Setup and Performance Metrics

Component	Details
Framework Used	PyTorch
Image Feature Extractor	ResNet-50
Model	GPT-2 Transformer-based Encoder-Decoder
Input Size	(224, 224, 3)
Batch size	32
Number of Epochs	75
Learning Rate	1×10^{-5}
Optimizer	AdamW
Training Strategy	Gradient Accumulation, Mixed-Precision Training
Datasets Used	BDD-X, MSVD, Filtered Flickr8k
Evaluation Metrics	BLEU 1-4, CIDEr, METEOR, ROUGE-L

Table 2: Ablation Study: Comparing the Impact of Model Components on Different Datasets

Dataset	Model Variant	BLEU-1	BLEU-2	BLEU-3	BLEU-4	CIDEr	METEOR	ROUGE-L
BDD-X	Baseline GPT-2 Model	0.693	0.668	0.638	0.601	0.921	0.234	0.640
	GPT-2 + Single-Head Attention	0.731	0.705	0.673	0.634	1.032	0.257	0.672
	GPT-2 + Without Gradient Accumulation	0.785	0.758	0.723	0.681	1.142	0.278	0.710
	ResNet50-GPT2 (Proposed)	**0.860**	**0.835**	**0.800**	**0.755**	**1.235**	**0.312**	**0.782**
MSVD	Baseline GPT-2 Model	0.710	0.685	0.654	0.612	0.980	0.242	0.651
	GPT-2 + Single-Head Attention	0.742	0.715	0.683	0.641	1.105	0.265	0.685
	GPT-2 + Without Gradient Accumulation	0.798	0.772	0.740	0.698	1.231	0.293	0.724
	ResNet50-GPT2 (Proposed)	**0.880**	**0.855**	**0.823**	**0.778**	**1.315**	**0.329**	**0.795**

on model performance. Model modifications were compared using the BLEU-1 to BLEU-4, CIDEr, METEOR, and ROUGE-L metrics, which provided a thorough assessment of the effects of each component on the quality of description production. As can be seen from Table 2, the results demonstrate the importance of each element in producing textual descriptions for video frames that are both logical and contextually rich. The model's performance on the MSVD and BDD-X datasets highlights the importance of each architectural component. The most significant impact came from removing ResNet-50, leading to major performance drops—BLEU-4 and CIDEr declined by 15.4% and 25.4% on BDD-X, and 21.3% and 25.5% on MSVD—emphasizing the need for strong visual features. Replacing multi-head attention with a single-head variant caused moderate losses, notably a 16.0% drop in CIDEr on MSVD, showing its role in capturing fine-grained text-visual interactions. Omitting gradient accumulation reduced training efficiency and led to average BLEU-4 and CIDEr decreases of 9.8% and 7.5%, confirming its value for stable optimization. Figure 2 illustrates these results. The proposed framework obtained the best scores on all evaluation measures, outperforming the ablated versions on all the datasets consistently. The improvements are especially clear in MSVD and BDD-X, where temporal dependency modeling and interpreting intricate visual scenes are critical. This better performance is a result of the synergy among multiple important elements: ResNet50 for feature extraction of high-level spatial features from images, GPT-2 as a strong language model with the ability to produce fluent and coherent text sequences, multi-head attention for modeling fine-grained alignments between visual and linguistic modalities, and gradient accumulation with mixed-precision training for efficient and stable optimization. The use of GPT-2, specifically, adds to the model's capacity to generate contextually accurate and linguistically advanced descriptions. Collectively, these features make the proposed framework capable of generating high-quality, context-aware descriptions, with each module contributing significantly to the overall performance improvement.

4.2 Results and Analysis:

To assess the efficacy of the suggested architecture, we carried out a comprehensive performance analysis on two benchmark datasets: MSVD and

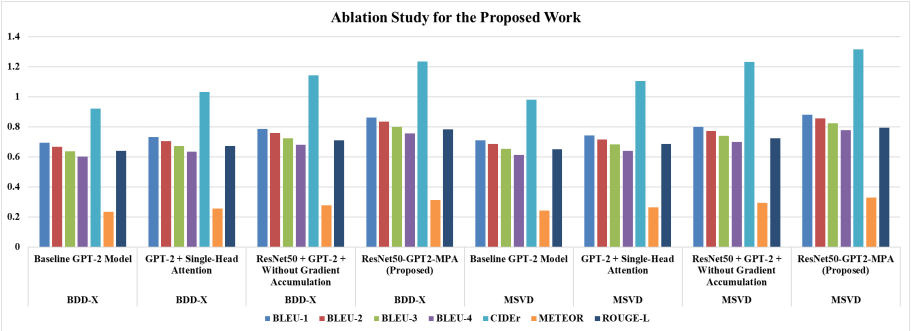


Fig. 2: Demonstration of Ablation Study for the Proposed Work

Table 3: Results and Analysis of the Proposed Framework

Dataset	BLEU-1	BLEU-2	BLEU-3	BLEU-4	CIDEr	METEOR	ROUGE-L
BDD-X	0.860	0.835	0.800	0.755	1.235	0.312	0.782
MSVD	0.880	0.855	0.823	0.778	1.315	0.329	0.795

BDD-X. The suggested framework was tested against many model ablations, such as excluding gradient accumulation, using single-head attention, and removing ResNet-50. The evaluation's findings are compiled in Table 3. The outcomes shown in the table demonstrate how well our suggested methodology performs across datasets, consistently achieving top scores in all evaluation metrics. While high BLEU scores reflect fluency and grammatical accuracy, the superior performance in CIDEr and METEOR highlights strong semantic alignment and high-quality descriptions. With a BLEU-4 score of 0.778 and a CIDEr score of 1.315, the MSVD dataset showcases the model's strength in producing diverse and precise captions for short video content. Similarly, the BDD-X dataset results—0.755 BLEU-4 and 1.235 CIDEr—illustrate the model's robustness in capturing intricate driving scenes and generating contextually appropriate, descriptive narratives. Consistently high ROUGE-L scores (0.795 for MSVD and 0.782 for BDD-X) further validate the framework's ability to generate outputs that closely align with human references. Notably, elevated CIDEr values emphasize the model's effectiveness in capturing contextual and domain-specific nuances, while higher METEOR scores reflect improved alignment and paraphrasing capabilities. Figure 3 demonstrates the graphical representation of the results obtained for the proposed framework. In general, the findings demonstrate that the suggested approach greatly improves the ability of video-based descriptions, resulting in syntactically fluid, semantically rich, and contextually relevant descriptions.

4.3 Comparison With State-of-the-Art Methods:

We compare the performance of the suggested model with current state-of-the-art (SOTA) captioning methods on two benchmark datasets—BDD-X and MSVD—in order to assess its efficacy. A thorough evaluation employing several evaluation metrics, such as BLEU-1 to BLEU-4, CIDEr, METEOR, and ROUGE-L, is given by the results in Table 4. On both datasets, the proposed framework significantly outperforms earlier approaches, proving its capacity to produce more semantically precise and contextually rich descriptions for behaviours involving vehicles. Specifically, the suggested model leads in terms of BLEU-4 and CIDEr scores, demonstrating enhanced description coherence and relevance. The table's comparative analysis demonstrates how the BDD-X and MSVD datasets have advanced regarding vehicle action comprehension and general video-based description generation. Previous models like OSCAR [21], ViT-GPT2 [20], X-LAN [19], and AoANet [22] showed consistent improvements in descriptive quality on the BDD-X dataset, with AoANet receiving the highest

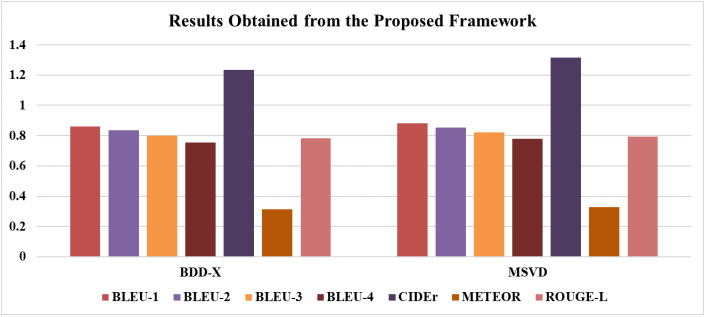


Fig. 3: Graphical Representation of the Results Obtained

Table 4: Comparison of State-of-the-Art Models for Vehicle Action Understanding

Dataset	Model	BLEU-1	BLEU-2	BLEU-3	BLEU-4	CIDEr	METEOR	ROUGE-L
BDD-X	GPT-2 (Baseline)	0.312	0.285	0.210	0.152	0.653	0.185	0.356
	X-LAN [19]	0.451	0.412	0.356	0.293	0.998	0.265	0.541
	ViT-GPT2 [20]	0.482	0.445	0.398	0.324	1.086	0.258	0.523
	OSCAR [21]	0.510	0.472	0.423	0.358	1.124	0.276	0.548
	AoANet [22]	0.532	0.495	0.442	0.372	1.189	0.289	0.573
	ResNet50-GPT2 (Proposed)	**0.860**	**0.835**	**0.800**	**0.755**	**1.235**	**0.312**	**0.782**
MSVD	GPT-2 (Baseline)	0.334	0.305	0.243	0.178	0.705	0.192	0.372
	Show, Attend and Tell [23]	0.480	0.443	0.395	0.304	0.943	0.252	0.520
	M^2 Transformer [24]	0.520	0.486	0.438	0.391	1.270	0.292	0.586
	Dense Video Captioning [25]	0.498	0.463	0.412	0.357	1.135	0.268	0.552
	GRU-EVE [26]	0.535	0.502	0.456	0.370	1.246	0.285	0.594
	ResNet50-GPT2 (Proposed)	**0.880**	**0.855**	**0.823**	**0.778**	**1.315**	**0.329**	**0.795**

scores. However, with a BLEU-4 score of 0.755 and a CIDEr score of 1.235, the proposed model outperforms all of these approaches, significantly advancing over prior research. Similarly, on the MSVD dataset, transformer-based models such as M^2 Transformer [24] and GRU-EVE [26] performed well, with GRU-EVE achieving a BLEU-4 score of 0.370 and CIDEr of 1.246. The proposed model demonstrates clear improvements, especially in BLEU-4 (0.778) and CIDEr (1.315), showcasing its ability to generate contextually rich and semantically aligned descriptions. Figure 4 depicts a graphical representation of SOTA methods. Overall, the suggested approach continuously outperforms comparison on all datasets, setting new benchmarks and proving its ability to capture fine-grained contextual information for interpreting video-based actions.

5 Conclusion and Future Direction

The proposed approach successfully combines ResNet50 for feature extraction with an improved GPT-2 model, allowing for the production of contextually aware, high-quality descriptions of video footage from the BDD-X and MSVD datasets. Through the use of transformer-based language modelling, the method guarantees descriptions that are consistent with ground-truth annotations and logical. In comparison to traditional approaches, the combination of ResNet50 for visual feature extraction and GPT-2 for text generation, as well as dataset augmentation, gradient accumulation, and mixed precision training, improves training efficiency and performance. Although transformer-based architectures are effective when used for large datasets, the study emphasizes that domain-specific fine-tuning is necessary to produce accurate and comprehensible descriptions. Future steps will involve implementing cross-modal attention mechanisms, sophisticated vision-based encoders, and optimized preprocessing methods to further improve model adaptability. The generalization and compatibility across other domains would be enhanced by expanding datasets to encompass a wider

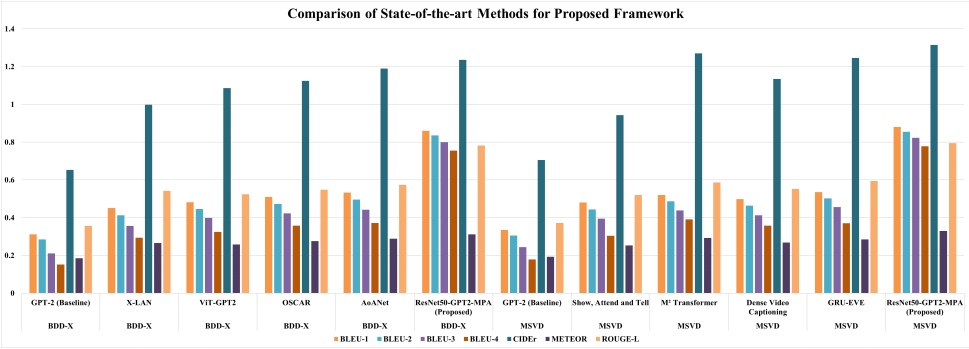


Fig. 4: Graphical Representation of the SOTA Methods

range of real-world circumstances. By enhancing interpretability, transparency, and decision-making clarity, this research advances explainable AI and increases the adaptability and dependability of automated video comprehension for a range of applications.

Acknowledgment

The authors sincerely thank the Object-Oriented Programming Laboratory, Department of IT, DTU, Delhi, India, for providing the necessary resources to complete the research.

Statements and Declarations

Data Availability Statement: The paper contains links to all datasets.
Competing Interests: The authors declare they have no competing financial interests.
Conflict of Interest: The authors declare no conflict of interest.

References

[1] N. Aafaq, A. Mian, W. Liu, S.Z. Gilani, M. Shah, Video description: A survey of methods, datasets, and evaluation metrics. ACM Computing Surveys (CSUR) **52**(6), 1–37 (2019)

[2] T. Li, H. Wang, X. Li, W. Liao, T. He, P. Peng, Generative planning with 3d-vision language pre-training for end-to-end autonomous driving. arXiv preprint arXiv:2501.08861 (2025)

[3] L. Agarwal, B. Verma, From methods to datasets: A survey on image-caption generators. Multimedia Tools and Applications **83**(9), 28077–28123 (2024)

[4] M. Shoman, D. Wang, A. Aboah, M. Abdel-Aty, *Enhancing traffic safety with parallel dense video captioning for end-to-end event analysis*, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2024), pp. 7125–7133

[5] Y. Li, C. Wu, L. Li, Y. Liu, J. Zhu, Caption generation from road images for traffic scene modeling. IEEE Transactions on intelligent transportation systems **23**(7), 7805–7816 (2021)

[6] C. Gou, Y. Zhou, D. Li, Driver attention prediction based on convolution and transformers. The Journal of Supercomputing **78**(6), 8268–8284 (2022)

[7] D. Chen, W.B. Dolan, *Collecting highly parallel data for paraphrase evaluation*, in *Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies* (2011), pp. 190–200

[8] S. Venugopalan, M. Rohrbach, J. Donahue, R. Mooney, T. Darrell, K. Saenko, *Sequence to sequence-video to text*, in *Proceedings of the IEEE international conference on computer vision* (2015), pp. 4534–4542

[9] L. Yao, A. Torabi, K. Cho, N. Ballas, C. Pal, H. Larochelle, A. Courville, *Describing videos by exploiting temporal structure*, in *Proceedings of the IEEE international conference on computer vision* (2015), pp. 4507–4515

[10] L. Zhou, Y. Zhou, J.J. Corso, R. Socher, C. Xiong, *End-to-end dense video captioning with masked transformer*, in *Proceedings of*

the IEEE conference on computer vision and pattern recognition (2018), pp. 8739–8748

[11] J. Lei, L. Wang, Y. Shen, D. Yu, T.L. Berg, M. Bansal, Mart: Memory-augmented recurrent transformer for coherent video paragraph captioning. arXiv preprint arXiv:2005.05402 (2020)

[12] C. Sun, A. Myers, C. Vondrick, K. Murphy, C. Schmid, *Videobert: A joint model for video and language representation learning*, in *Proceedings of the IEEE/CVF international conference on computer vision* (2019), pp. 7464–7473

[13] J. Lei, L. Li, L. Zhou, Z. Gan, T.L. Berg, M. Bansal, J. Liu, *Less is more: Clipbert for video-and-language learning via sparse sampling*, in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2021), pp. 7331–7341

[14] J. Xu, T. Mei, T. Yao, Y. Rui, *Msr-vtt: A large video description dataset for bridging video and language*, in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2016), pp. 5288–5296

[15] J. Kim, A. Rohrbach, T. Darrell, J. Canny, Z. Akata, *Textual explanations for self-driving vehicles*, in *Proceedings of the European conference on computer vision (ECCV)* (2018), pp. 563–578

[16] C. Cui, Y. Ma, X. Cao, W. Ye, Y. Zhou, K. Liang, J. Chen, J. Lu, Z. Yang, K.D. Liao, et al., *A survey on multimodal large language models for autonomous driving*, in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (2024), pp. 958–979

[17] C. Hori, T. Hori, T.Y. Lee, Z. Zhang, B. Harsham, J.R. Hershey, T.K. Marks, K. Sumi, *Attention-based multimodal fusion for video description*, in *Proceedings of the IEEE international conference on computer vision* (2017), pp. 4193–4202

[18] J. Yu, J. Li, Z. Yu, Q. Huang, Multimodal transformer with multi-view visual representation for image captioning. IEEE transactions on circuits and systems for video technology **30**(12), 4467–4480 (2019)

[19] Y. Pan, T. Yao, Y. Li, T. Mei, *X-linear attention networks for image captioning*, in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2020), pp. 10971–10980

[20] I. Vasireddy, G. HimaBindu, B. Ratnamala, Transformative fusion: Vision transformers and gpt-2 unleashing new frontiers in image captioning within image processing. International Journal of Innovative Research in Engineering & Management **10**(6), 55–59 (2023)

[21] X. Li, X. Yin, C. Li, P. Zhang, X. Hu, L. Zhang, L. Wang, H. Hu, L. Dong, F. Wei, et al., *Oscar: Object-semantics aligned pre-training for vision-language tasks*, in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXX 16* (Springer, 2020), pp. 121–137

[22] L. Huang, W. Wang, J. Chen, X.Y. Wei, *Attention on attention for image captioning*, in *Proceedings of the IEEE/CVF international conference on computer vision* (2019), pp. 4634–4643

[23] O. Vinyals, A. Toshev, S. Bengio, D. Erhan, *Show and tell: A neural image caption generator*, in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2015), pp. 3156–3164

[24] M. Cornia, M. Stefanini, L. Baraldi, R. Cucchiara, *Meshed-memory transformer for image captioning*, in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2020), pp. 10578–10587

[25] R. Krishna, K. Hata, F. Ren, L. Fei-Fei, J. Carlos Niebles, *Dense-captioning events in videos*, in *Proceedings of the IEEE international conference on computer vision* (2017), pp. 706–715

[26] N. Aafaq, N. Akhtar, W. Liu, S.Z. Gilani, A. Mian, *Spatio-temporal dynamics and semantic attribute enriched visual encoding for video captioning*, in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2019), pp. 12487–12496

Tri-FusionNet: Enhancing Image Description Generation with Transformer-based Fusion Network and Dual Attention Mechanism

Lakshita Agarwal*, Bindu Verma† * Department of Information Technology, Delhi Technological University, Delhi, India - 110042 † Assistant Professor, Department of Information Technology, Delhi Technological University, Delhi, India - 110042

arXiv:2504.16761v1 [cs.CV] 23 Apr 2025

***Abstract*—Image description generation is essential for accessibility and AI understanding of visual content. Recent advancements in deep learning have significantly improved natural language processing and computer vision. In this work, we propose Tri-FusionNet, a novel image description generation model that integrates transformer modules: a Vision Transformer (ViT) encoder module with dual-attention mechanism, a Robustly Optimized BERT Approach (RoBERTa) decoder module, and a Contrastive Language-Image Pre-Training (CLIP) integrating module. The ViT encoder, enhanced with dual attention, focuses on relevant spatial regions and linguistic context, improving image feature extraction. The RoBERTa decoder is employed to generate precise textual descriptions. CLIP's integrating module aligns visual and textual data through contrastive learning, ensuring effective combination of both modalities. This fusion of ViT, RoBERTa, and CLIP, along with dual attention, enables the model to produce more accurate, contextually rich, and flexible descriptions. The proposed framework demonstrated competitive performance on the Flickr30k and Flickr8k datasets, with BLEU scores ranging from 0.767 to 0.456 and 0.784 to 0.479, CIDEr scores of 1.679 and 1.483, METEOR scores of 0.478 and 0.358, and ROUGE-L scores of 0.567 and 0.789, respectively. On MS-COCO, the framework obtained BLEU scores of 0.893 (B-1), 0.821 (B-2), 0.794 (B-3), and 0.725 (B-4). The results demonstrate the effectiveness of Tri-FusionNet in generating high-quality image descriptions.**

***Index Terms*—Image Description Generation, Natural Language Processing, Computer Vision, Transformers, Dual Attention Mechanism, Artificial Intelligence.**

I. Introduction

In order to improve accessibility and comprehension, image description generation, which is essential to computer vision and artificial intelligence— involves producing textual descriptions of images. These descriptions, which make use of transformer architectures and deep learning, seek to accurately capture the context and content of images [1]. By combining computer vision and natural language processing, this multimodal method allows machines to understand and communicate with visual content in a way that is similar to that of humans. The goal of the research is to find image-processing algorithms that efficiently link textual and visual elements while preserving linguistic fluency. Applications include social media image annotation, real-time navigation systems, visual question answering, helping the blind, and self-driving automobiles. After breakthroughs in image classification models, adversarial examples are being developed for more complex tasks like image description generation [2].

Corresponding author: Dr. Bindu Verma (email: bindu.cvision@gmail.com).

Transformer models excel in generating image descriptions using self-attention but struggle with complex relationships. Single transformers may fail to capture both global and local details, while dual models often struggle to integrate visual and textual information effectively. This can hinder fine-grained visual understanding and complicate non-autoregressive techniques [3]. To address these issues, we propose a model combining Vision Transformer (ViT) with dual attention, RoBERTa, and a CLIP integration module. ViT transforms images into embedding sequences, capturing global and local features through transformer layers and pre-training on large datasets for image classification. RoBERTa, based on BERT, enhances textual context understanding by leveraging extensive text pre-training [4]. CLIP, developed by OpenAI, enables vision-language learning by aligning image and text representations for multimodal tasks [5].

With dual attention, the suggested Tri-FusionNet architecture combines three transformers to improve image descriptions by making them more contextually relevant and advanced. This fusion includes a CLIP integrating model for textual representation generation, a RoBERTa decoder for category derivation and a Vision Transformer (ViT) for feature capture. By combining ViT, RoBERTa and CLIP with dual attention, one may enhance the quality of descriptions and build a robust and flexible system for producing visual descriptions.

The major contribution of the proposed framework is highlighted below:

- This work presents a novel image description-generating model called Tri-FusionNet, which represents a major breakthrough in multi-modal approaches to generating precise and contextually rich descriptions. This method improves the model's fine-tuning capabilities by combining several components, which increases the efficacy of creating image descriptions.
- The combination of ViT encoder with dual-attention, RoBERTa decoder, and CLIP integration module is employed to create an efficient model that maximizes the interaction between different modalities, resulting in more accurate and contextually appropriate image descriptions.
- The proposed framework addresses the challenges posed by different benchmark datasets by developing a model

that can adapt to the characteristics of each dataset as well as potentially setting new state-of-the-art results on these datasets.

Organization of the paper is as follows: Section II explores prior research and previous works. Section III provides a comprehensive presentation of the suggested framework. Section IV presents and analyses the experimental design, benchmark dataset information, results and ablation study. Section V provides a conclusion as well as a discussion of the ability of the study and its future directions.

II. Related Work

Deep learning frameworks have been investigated for autonomous image description systems. These frameworks are extensively used in the fields of computer vision and human-computer interaction. The transformer exhibits potential for multi-modal tasks such as visual description creation, as it is skilled at collecting long-range dependencies and modeling sequential data. Li et al. [6] developed the EnTangled Attention (ETA) Transformer, exhibiting state-of-the-art performance on the MSCOCO dataset, whereas Cornia et al. [7] presented M2, a Meshed Transformer with Memory, for image description generation. Furthermore, a Multi-modal Transformer (MT) model was introduced by Yu et al. [8], that stores intra-modal and inter-modal interactions in a single attention block, enabling precise description development and sophisticated multi-modal reasoning.

Radford et al. [9] demonstrated that pre-training a model to predict image-caption pairs from a dataset of 400 million (image, text) pairs enables zero-shot transfer to various downstream tasks, achieving competitive performance without task-specific training. MiniGPT-v2 is a unified model designed to handle diverse vision-language tasks like image description, visual question answering, and visual grounding, using unique task identifiers to improve learning efficiency and performance on various benchmarks [10]. In order to improve image captioning performance on the MS COCO and Flickr8k datasets, Yang et al. [11] suggested SAMT-Generator, a multi-stage transformer feature enhancement network with second attention and Maxout decoding. A novel approach of the S^2 transformer was developed that integrated Spatial-aware Pseudo-supervised and Scale-wise Reinforcement modules, achieving state-of-the-art results on the MSCOCO benchmark [12]. Yang et al. [13] presented a fully Transformer-based image captioning model called CA-Captioner, which applies HAPE positional encoding, LSM, RNorm function, and LFE, thereby achieving improved performance on the MS COCO, Flickr8k, and Flickr30k datasets, especially in terms of BLEU4 and CIDEr metrics.

On the Flickr30k and MSCOCO datasets, Parvin et al. [14] proposed a double-attention framework that outperformed state-of-the-art models. By utilising previous activations to improve semantic attention and performance, PMA-Net [15] integrates prototypical memory vectors into Transformer-based image captioning, attaining a 3.7 CIDEr boost on the COCO dataset. By treating heterogeneous encodings (such as visual and textual) as augmented views and using a shared encoder with contrastive loss to improve representation quality and a hierarchical decoder to adaptively weigh views, HAAV [16] presents a novel method for image captioning. It achieves notable CIDEr improvements on MS-COCO (+5.6%) and Flickr30k (+16.8% with SSL) datasets. DualVision Transformer (Dual-ViT) was introduced by Yao et al. [17], providing higher accuracy with effective token vector compression. Spatial Pyramid Transformer (SPT) was mainly used for adaptive semantic interaction across grid resolutions, preserving spatial and fine-grained information, achieving state-of-the-art performance on MS-COCO [18] for image description generation.

The challenges of creating image descriptions remain, despite the progress made in medical imaging and NLP with models such as ETA Transformer, Multi-modal Transformer (MT) and Dual-Level Collaborative Transformer (DLCT). These challenges include the semantic gap between language and vision, as well as the absence of global information that is critical for scene comprehension. The suggested Tri-FusionNet combines dual attention, RoBERTa and CLIP transformers with Vision Transformer (ViT) to address these issues. By using ViT for image embedding, RoBERTa for textual interpretation, and CLIP for visual-textual fusion, it is possible to improve understanding of visual content.

III. Proposed Framework

Recently, image description creation has gained a lot of attention as a challenging topic combining computer vision and natural language processing (NLP). Although the current models have come a long way, there is still an opportunity for development in producing precise and contextually appropriate descriptions. This work presents Tri-FusionNet, which combines a dual attention mechanism with Vision Transformer (ViT), CLIP, and RoBERTa decoder to improve spatial and channel-wise information extraction from images. Better descriptive words are generated over a range of images, advancing image description generation. The model is optimized with Adam optimizer and Cross-Entropy loss over epochs. Its performance is assessed using metrics including BLEU, CIDEr, ROUGE-L, and METEOR scores on benchmark datasets. Figure 1 represents the framework of the proposed approach.

A. Overview of the Proposed Architecture:

1) Data Pre-processing: The data preprocessing step in the proposed Tri-FusionNet framework involves normalizing the input images' pixel values to enhance the model's performance. The mean and standard deviation of the pixel values throughout the dataset are computed to accomplish this. The distribution of the pixel values is then altered to have a standard deviation (or variance) of one and a mean of zero. This normalization process reduces the effect of varied lighting or contrasts between images by standardizing the range of pixel values. The model can learn more from the data by focusing the values around zero and scaling them to have unit variance, guaranteeing that each feature contributes equally to the training process. Figure 1(a) demonstrates the Data Pre-Processing step. Next, to prepare the data for the Vision Transformer (ViT) encoder module in image description generation,

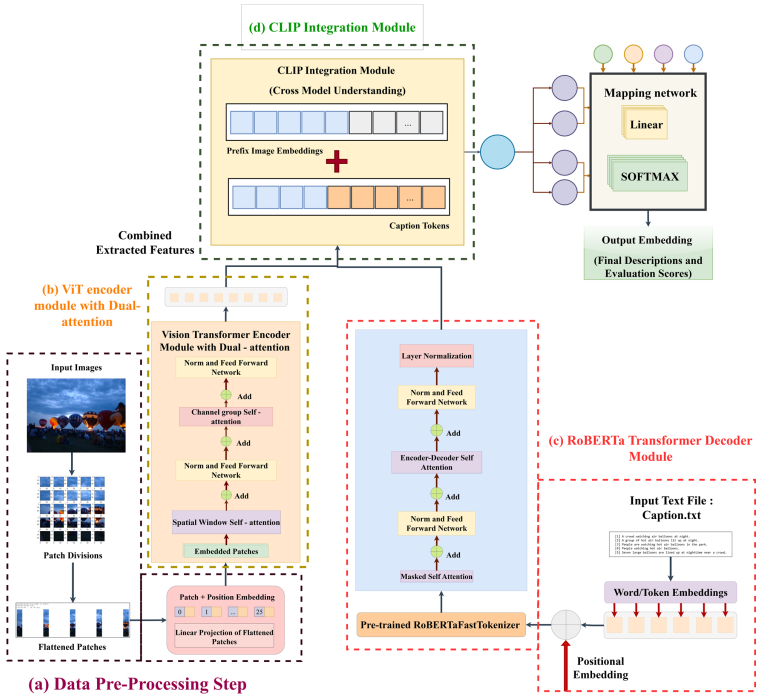


Fig. 1. Structural representation of Tri-FusionNet framework: The architecture consists of three phases: firstly, high-level visual features are first extracted from pre-processed images using a Vision transformer encoder with dual-attention mechanism; next, words from the input caption file are tokenized by a RoBERTa decoder; and last, the combined data is fed into a CLIP-integrating module to create image descriptions using dense network layers.

input images are divided into fixed-size patches. To construct embeddings, each patch is flattened and projected linearly. The process of patch extraction and positional encoding is thoroughly explained in the `supplementary.tex` file. The ViT encoder receives sequences of tokenized patches that have been enhanced with positional data. Through this procedure, the model is able to efficiently collect the image's spatial and visual properties, which sets the stage for the next tasks, like labeling the image.

2) Vision Transformer Encoder Module with Dual Attention Mechanism: The input image is divided into fixed-size patches using the Vision Transformer's encoder-based architecture, which is then processed through transformer layers and linearly embedded. Its ability to retrieve relevant data from both spatial and channel-wise dimensions is enhanced by the dual-attention process, which makes accurate descriptions possible. In the vision transformer encoder with a dual attention mechanism, we adopt a hierarchical layout. The encoder combines spatial window attention with channel-group attention and is structured into four phases: (a) insertion of the patch embedding layer, (b) application of spatial window attention, (c) utilization of channel-group attention, and (d) incorporation of the jointly extracted features. Figure 1(b) demonstrates the Vision encoder module with dual-attention mechanism. In the initialization phase, the input is in the form of patches along with positional embedding obtained from the data pre-processing stage. Assuming a visual feature R with dimensions $R^{(P\times C)}$, where P represents the total number of patches and C is the total number of channels, the standard global self-attention is represented by Equation (1):

$$S_A(Q,K,V) = \text{Concat}(\text{head}1,\ldots,\text{head}N_h), \quad (1)$$

where,

$$\text{head}_i = \text{Attention}(Q_i,K_i,V_i) = \text{Softmax}\left(\frac{Q_iK_i^T}{\sqrt{C_h}}\right)V_i \quad (2)$$

here, $\text{Q}_i = X_i \times (W_i)^Q$, $\text{K}_i = X_i \times (W_i)^K$ and $\text{V}_i = X_i \times (W_i)^V$. Consider $R^{P\times C_h}$ dimensional visual features with N_h heads, where X_i represents the ith head of the input feature and W_i denotes the projection weights for the ith head in the context of Q, K, V, with $C = C_h \times N_h$.

The model concurrently arranges spatial window attention and channel group attention in the vision transformer with a dual attention mechanism to obtain both local and global data but with a linear complication to the spatial dimension. The spatial window attention algorithm calculates self-attention within local windows, which are positioned to equally and non-overlapping segments in the field of vision. Assuming the presence of N_w distinct windows, each comprised of P_w patches, the total number of patches, denoted as P, can be expressed as $P = P_w \times N_w$. The representation of spatial window attention is given by Equation (3):

$$\text{Attention}w(Q,K,V) = (\text{Attention}(Q_i,K_i,V_i))^{N_w}\ i = 0 \quad (3)$$

where each $Q_i, K_i, V_i \in \mathbb{R}^{P_w\times C_h}$ are local window queries, keys, and values, respectively, and P is the spatial size presenting the linear complexity.

An alternative viewpoint on self-attention is provided by channel-wise attention, which ensures thorough spatial domain coverage by concentrating on tokens at the patch level as

opposed to pixels. Each transposed token, with the number of heads set to 1, engages with global data in a linear spatial complexity along the channel dimension. $C = N_g \times C_g$ is the outcome of letting C_g represent the number of channels in each group and N_g the number of groups. As a result, Equation (4) defines global channel group attention, which allows tokens at the image level to interact across channels.

$$\text{Attention}c(Q, K, V) = \left((\text{Attentiongroup}(Q_i, K_i, V_i))^T \right)_{i=0}^{N_g} \tag{4}$$

where, every $Q_i, K_i, V_i \in \mathbb{R}^{P \times C_g}$ represents grouped channel-wise image-level queries, keys and values.

The final encoder output of the Vision Transformer using the dual attention mechanism can be represented by Equation (5):

$$E_{\text{output}} = \text{Concat}(\text{Attention}_w(Q, K, V), \text{Attention}_c(Q, K, V)) \tag{5}$$

$\text{Attention}_w(Q, K, V)$ represents the spatial window attention and $\text{Attention}_c(Q, K, V)$ represents the channel group attention, as defined in Equations (3) and (4), respectively. In order to achieve linear complexity in both dimensions for computational efficiency, the final encoder output combines channel group attention and spatial window attention. In contrast to spatial-wise global attention, channel attention functions globally, aggregating information rather than locally. It is a complement to spatial window attention. By distributing weights according to how relevant they are to the task, the model uses its attention mechanism to prioritize particular image patches. By using these weights, heat maps are produced that show the areas of the image where the model focuses. On the heat map, places that are important for accurate and contextually rich descriptions are indicated by intense areas.

3) RoBERTa Decoder: The study presents a novel approach to image description generation that combines natural language processing (NLP) and computer vision. A Vision Transformer (ViT) with dual attention is employed to extract visual information. Simultaneously, textual data from a "caption.txt" file is used during training to build word embeddings and guide the model in learning semantic relationships between images and their descriptions. This process helps optimize the model and align visual and textual features effectively. During inference, the model does not require an input description or text file. Instead, it generates a new description based solely on the visual features extracted from the input image. This distinction ensures the model's flexibility and generalization, enabling it to produce unique and contextually rich descriptions for each image. These embeddings are seamlessly integrated into a decoder model based on RoBERTa, an enhanced version of BERT, along with ViT-extracted image features. Figure 1(c) demonstrates the step involved in the RoBERTa decoder module.

The decoder uses RoBERTa's pre-trained language understanding abilities to contextualize the data from the text and image embedding, allowing it to produce descriptions that generate better explanations and are relevant to the context. To decode and generate image descriptions, the RoBERTa decoder module is obtained from the following Equation (6).

$$D_{\text{Output}} = \text{RoBERTaDecoder}\left(W_{\text{emb}}, C\right), \tag{6}$$

where, W_{emb} represents the word embeddings generated from the caption data, C is the additional contextual information derived from the original caption file, and D_{Output} is the output generated by the RoBERTa decoder, which consists of the final description of the image. Through multiple layers of feed-forward neural networks and self-attention, multiple word embeddings are processed by the RoBERTa decoder, which enables it to efficiently contextualize the textual information.

4) CLIP Integration Module: The final step of the proposed approach involves generating image embeddings using a Vision Transformer (ViT) with a dual attention mechanism and text embeddings using the RoBERTa decoder. This integration aims to combine the ViT encoder, RoBERTa decoder, and CLIP for effective image description generation. During the embedding generation process, the ViT encoder is used to obtain image embeddings from the image dataset, while the RoBERTa decoder generates text embeddings from the caption data. The joint representation of the image and its corresponding text is then obtained by aligning these embeddings in the subsequent phase. Specifically, CLIP's encoder is used to align the visual and textual embeddings through contrastive learning, ensuring that the visual and textual features are mapped into a shared space for accurate description generation. The embeddings obtained from both the ViT encoder and the RoBERTa decoder are concatenated to form a unified representation, as shown in Equation (7):

$$CLIP_{\text{integration}} = \text{Concatenate}(E_{\text{output}}, D_{\text{output}}), \tag{7}$$

where E_{output} is the image embedding from the ViT encoder, D_{output} is the text embedding from the RoBERTa decoder, and $CLIP_{\text{integration}}$ represents the combined embedding used by the CLIP model. This integrated representation helps generate coherent and accurate image descriptions by aligning the visual and textual information effectively. The model uses a contrastive loss function in the embedding space to distinguish between positive and negative pairs, ensuring the alignment between visual and textual features. The final integration module, after encoding and decoding the image-description pair, generates the output descriptions. The architectural explanation of the CLIP integration module is explained in the `supplementary.txt` file.

The reference file for descriptions is crucial in assessing how well the Tri-FusionNet model generates descriptions. Visual features extracted from images are projected into a higher-dimensional space using a linear layer, allowing the model to capture intricate relationships among input features. These features are then normalized into a probability distribution across the lexicon of possible words or tokens using a Softmax layer. During training and evaluation, the model generates predicted descriptions, which are compared to reference descriptions using evaluation metrics such as CIDEr, ROUGE-L, BLEU 1-4, and METEOR to assess the model's ability to capture and translate context and image semantics into accurate descriptions.

The Tri-FusionNet model operates through a synergistic pipeline. The Vision Transformer (ViT) encoder with dual attention, RoBERTa decoder, and CLIP integration module

collaborate to generate accurate and contextually rich image captions.

- To handle input images, the Vision Transformer (ViT) encoder flattens, divides them into patches, and embeds positional encodings. Its dual attention method combines Spatial Window Self-Attention, which concentrates on local spatial relationships to maintain contextual importance, with Channel Group Self-Attention, which catches fine-grained features inside image channels. As a result, both local details and global context are successfully represented in the rich visual feature embeddings.
- To process the input text file with ground truth descriptions, the RoBERTa decoder tokenises the text, embeds the tokens, then decodes them into contextual embeddings. It uses its pre-trained language understanding to ensure fluency and coherence and Masked Self-Attention to concentrate on sequentially relevant tokens. This produces textual embeddings that effectively express the descriptions' semantic meaning and linguistic structure.
- Using a contrastive learning technique, the CLIP integration module maps both modalities into a common latent space by aligning textual embeddings from the RoBERTa decoder and visual features from the ViT encoder. By facilitating cross-modal fusion and preserving semantic coherence, this alignment makes it possible to produce linguistically correct and visually justified descriptions.

The model optimizes for metrics such as BLEU, CIDEr, METEOR, and ROUGE-L and provides accurate and contextually rich descriptions through parallel processing, dual attention, and contrastive learning. By utilizing each module's unique capabilities to convert raw image and text inputs, the approach guarantees high-quality description generation. The detailed algorithm of the proposed model is given in the `supplementary.tex` file.

IV. Experimental Analysis

The proposed model was evaluated using Google Colab Pro+ with 52 GB RAM, an NVIDIA A100 GPU (40 GB VRAM), and a virtual Intel Xeon CPU. Implemented in Keras and TensorFlow 2.12, Tri-FusionNet integrates three transformer modules, increasing computational demands. Optimization techniques like parallelization, mixed precision training, and resource management reduce memory usage and ensure efficiency. Despite its complexity, model parallelism enables viability on limited hardware. Scalability tests on larger datasets showed a linear increase in training time, demonstrating efficient scaling. Optimized attention mechanisms and parameter sharing enhance computational efficiency compared to state-of-the-art models, balancing performance with resource constraints. The model was tested on MSCOCO[1], Flickr8k[2], and Flickr30k[3], with efficiency, cost, and scalability compared in Table I.

Because of its transformer and CLIP components, the proposed Tri-FusionNet model has a moderate processing cost but excels at handling huge datasets. On the other hand, BLIP Transformer [20] provides a fair trade-off between cost and performance, whereas models such as NIC [19] and M2 Transformer [7] are more resource-efficient but less scalable.

TABLE I
Comparison of Tri-FusionNet with state-of-the-art models on scalability and computational cost.

Model	Architecture	Scalability	Computational Cost	Efficiency
NIC [19]	CNN with Attention Mechanism	Limited scalability	Low cost and less resource-intensive	Good for smaller tasks but lacks richness
BLIP Transformer [20]	Vision-Language Pretraining	Scales well with efficient pretraining	Moderate cost and compact architecture	Balanced performance and cost
M2 Transformer [7]	Memory-Augmented Transformer	Effective for mid-sized datasets	Moderate cost and memory-efficient	Competitive with moderate resources
Tri-FusionNet	**ViT, RoBERTa, CLIP**	**Handles large datasets**	**Moderate cost and memory efficient**	**Superior image-text alignment**

The implementation details of the proposed framework is demonstrated thoroughly in `supplementary.tex` file.

A. Ablation Study:

An ablation study was carried out in order to fully comprehend the contributions of each element in the suggested image description generation model. The objective of this study is to assess the impact of the proposed Tri-FusionNet model, which integrates the CLIP model, RoBERTa and Vision Transformer (ViT) with dual attention, on the overall performance. Through heat map analysis, the model can learn more about how it interprets visual cues and focuses on various areas of the image. This comprehension can enhance the model's architecture, boost its functionality, and make the model's decision-making process more interpretable, as shown in Figure 2.

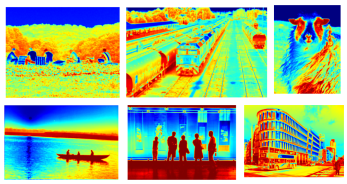


Fig. 2. Example of obtained heat maps based on dual-attention mechanism.

A thorough study of the outcomes within the framework of the work proposed for MSCOCO dataset is shown in Table II. It presents various results of evaluation metrics for different models applied to the input image dataset.

The Tri-FusionNet model sets a new benchmark by significantly outperforming baseline models among those evaluated on the MSCOCO dataset for image description generation, including the CNN-RNN Model, CNN-LSTM Model, ViT with Attention, ViT with Dual Attention, Vanilla ViT, ViT only, RoBERTa Transformer, CLIP Transformer, and ViT with RoBERTa Model. It achieves a 1.88 CIDEr score increase and notable BLEU values (0.893, 0.821, 0.794, and 0.725). Tri-FusionNet also demonstrates better precision, recall, and alignment with reference phrases, along with enhanced n-gram

[1]https://github.com/cocodataset/cocoapi
[2]https://www.kaggle.com/adityajn105/flickr8k
[3]https://www.kaggle.com/hsankesara/flickr-image-dataset

TABLE II
PERFORMANCE METRICS OF IMAGE DESCRIPTION GENERATION MODELS IN THE MSCOCO DATASET

Model	B-1	B-2	B-3	B-4	C	M	R-L
CNN-RNN Model	0.54	0.45	0.33	0.21	1.02	0.31	0.27
CNN-LSTM Model	0.55	0.42	0.30	0.28	1.20	0.48	0.35
ViT with Attention	0.57	0.45	0.34	0.29	1.05	0.40	0.39
ViT with Dual Attention	0.61	0.50	0.38	0.27	1.52	0.43	0.31
Vanilla ViT	0.72	0.61	0.50	0.38	1.45	0.52	0.45
ViT Only	0.62	0.56	0.42	0.34	0.785	0.44	0.39
RoBERTa Transformer	0.73	0.68	0.51	0.43	1.54	0.62	0.50
CLIP Transformer	0.66	0.54	0.43	0.32	1.09	0.51	0.38
ViT and RoBERTa	0.74	0.63	0.55	0.46	1.80	0.68	0.56
ViT+ RoBERTa + CLIP	0.78	0.69	0.65	0.58	1.83	0.73	0.62
Tri-FusionNet (Proposed)	**0.893**	**0.821**	**0.794**	**0.725**	**1.88**	**0.78**	**0.689**

overlap, as shown by its ROUGE-L score of 0.689 and METEOR score of 0.78. When compared to the ViT + RoBERTa + CLIP model, which achieves scores of 0.78 for METEOR, 0.69 for ROUGE-L, 0.65 for BLEU-4, 0.58 for BLEU-3, 1.83 for CIDEr, 0.73 for BLEU-2, and 0.62 for BLEU-1, Tri-FusionNet's performance highlights the effectiveness of its integrated attention mechanisms. The dual-attention mechanism proves especially useful, while the Vanilla ViT model demonstrates the strength of vision transformers. The CLIP Transformer outperforms the RoBERTa Transformer, despite the latter's slight under-performance.

Table III displays a comprehensive analysis of the results obtained from the Flickr8k dataset in the context of the proposed work. It presents various evaluation metrics for different models applied to the input image. The Flickr30k

TABLE III
PERFORMANCE METRICS OF IMAGE DESCRIPTION GENERATION MODELS IN THE FLICKR30K DATASET

Model	B-1	B-2	B-3	B-4	C	M	R-L
CNN-RNN Model	0.45	0.35	0.30	0.20	0.900	0.231	0.197
CNN-LSTM Model	0.43	0.32	0.25	0.18	0.901	0.28	0.25
ViT with Attention	0.47	0.33	0.23	0.17	0.907	0.34	0.331
ViT with Dual Attention	0.43	0.36	0.275	0.242	1.02	0.373	0.21
Vanilla ViT	0.682	0.513	0.420	0.382	0.355	0.412	0.353
ViT Only	0.614	0.586	0.542	0.443	0.855	0.544	0.393
RoBERTa Transformer	0.653	0.589	0.431	0.367	1.14	0.512	0.460
CLIP Transformer	0.516	0.424	0.343	0.232	1.256	0.534	0.489
ViT and RoBERTa	0.724	0.613	0.455	0.346	1.250	0.368	0.256
ViT+ RoBERTa + CLIP	0.741	0.621	0.573	0.428	1.092	0.389	0.432
Tri-FusionNet (Proposed)	**0.767**	**0.654**	**0.647**	**0.456**	**1.679**	**0.478**	**0.567**

dataset is used to evaluate several image description generation models, including ViT with RoBERTa, CNN-LSTM, Vanilla ViT, ViT with Attention, ViT with Dual Attention, ViT only, and RoBERTa Transformer. Despite its simplicity, the CNN-RNN Model provides competitive metrics and BLEU scores (0.20 to 0.45) as a baseline. The CNN-LSTM Model shows minor improvements, particularly in METEOR and ROUGE-L ratings. As demonstrated by higher METEOR, ROUGE-L, and BLEU-1 to BLEU-3 scores, the ViT with Attention and ViT with Dual Attention models benefit from attention mechanisms that improve language detail capture. While the RoBERTa Transformer performs exceptionally well in BLEU-2 and BLEU-3, the Vanilla ViT variant also performs robustly. The ViT + RoBERTa + CLIP model achieves scores of 0.741 for METEOR, 0.621 for ROUGE-L, 0.573 for BLEU-4, 0.428 for BLEU-3, 1.092 for CIDEr, 0.389 for BLEU-2, and 0.432 for BLEU-1. The Tri-FusionNet model produces accurate and varied visual descriptions, outperforming other models across BLEU-1 to BLEU-4 and achieving a BLEU-4 score of 0.456.

Table IV displays a comprehensive analysis of the results obtained from the Flickr8k dataset in the context of the proposed work. It presents various evaluation metrics for different models applied to the input image. Using CNN-LSTM and

TABLE IV
PERFORMANCE METRICS OF IMAGE DESCRIPTION GENERATION MODELS IN THE FLICKR8K DATASET

Model	B-1	B-2	B-3	B-4	C	M	R-L
CNN-RNN Model	0.467	0.435	0.370	0.323	0.670	0.131	0.137
CNN-LSTM Model	0.433	0.232	0.225	0.118	0.301	0.128	0.235
ViT with Attention	0.547	0.343	0.323	0.217	0.607	0.234	0.131
ViT with Dual Attention	0.543	0.436	0.344	0.246	1.002	0.253	0.121
Vanilla ViT	0.632	0.511	0.320	0.242	0.255	0.316	0.153
ViT Only	0.542	0.466	0.342	0.266	0.675	0.344	0.343
RoBERTa Transformer	0.553	0.459	0.413	0.337	0.983	0.312	0.156
CLIP Transformer	0.525	0.324	0.244	0.132	1.065	0.234	0.289
ViT and RoBERTa	0.745	0.513	0.456	0.312	1.189	0.298	0.457
ViT+ RoBERTa + CLIP	0.756	0.652	0.518	0.460	1.231	0.321	0.528
Tri-FusionNet (Proposed)	**0.784**	**0.678**	**0.538**	**0.479**	**1.381**	**0.389**	**0.654**

CNN-RNN as baseline models, the Flickr8k dataset compares several image description models. While Vanilla ViT performs well in BLEU-1 to BLEU-2, attention mechanisms in ViT with Attention and ViT with Dual Attention models enhance visual descriptions. In BLEU-3 and CIDEr, ViT only, the RoBERTa Transformer outperforms the CLIP Transformer, which lags slightly behind. The ViT + RoBERTa + CLIP model achieves scores of 0.756 for METEOR, 0.652 for ROUGE-L, 0.518 for BLEU-4, 0.460 for BLEU-3, 1.231 for CIDEr, 0.321 for BLEU-2, and 0.528 for BLEU-1. ViT and RoBERTa work together to create a balanced model. Nonetheless, the suggested Tri-FusionNet surpasses all other models, attaining the highest ratings in CIDEr, ROUGE-L, and BLEU-1 to BLEU-4 measures, demonstrating its efficiency in producing precise and diverse image descriptions.

By combining language and vision transformer components, the Tri-FusionNet consistently outperforms other models in the MSCOCO, Flickr30k and Flickr8k datasets, demonstrating its superior architecture for image description generation and producing better image descriptions. It also yields higher scores in BLEU-1 to BLEU-4 metrics and CIDEr, ROUGE-L and other metrics.

B. Results and Analysis:

The approach was evaluated on three benchmark dataset: MSCOCO, Flickr30k and Flickr8k for the proposed architecture. The following Table V represents the results obtained for the MSCOCO dataset.

Using four sets of test images from the MSCOCO dataset, the table presents a thorough comparison of ground-truth descriptions and predicted descriptions generated by a model. This is an effective way of assessing and determining how well the model can provide meaningful and accurate descriptions of images.

TABLE V
QUANTITATIVE RESULTS OBTAINED ON MSCOCO DATASET

Test Image	Ground Truth	Predicted Description
	Two men playing frisbee on grass surrounded by trees.	Two men playing frisbee on grass.
	Trains on railway tracks, with trees and blue sky.	Trains on tracks with trees.
	Vegetables including carrot, radish and turnip on a table.	Green and red vegetables on a table.
	A red bus in front of a white building with blue sky.	A white building with blue sky.

TABLE VI
QUANTITATIVE RESULTS OBTAINED ON FLICKR30K DATASET

Test Image		
Ground Truth Descriptions	1. A white woman standing in a grocery store not-so-candidly posing for the camera while examining the items on a shelf 2. The lady with the shopping cart is surrounded by toys galore and various children's bicycles. 3. A woman in a green winter coat stands with a cart in the middle of a department store isle 4. a woman in a green jacket is in the toy aisle with a shopping cart and her purse 5. The woman in the green coat is pushing a cart through the toy aisle	1. Six men are sitting or laying on a patch of earth in a wooded area. 2. A group of workers sitting in a field take a break from work. 3. A group of men are sitting in the farm fields taking a break. 4. Six men sit in a field of crops containing wooden crates. 5. Pickers working out on a farm
Predicted Description	A woman in a green coat shops in the toy aisle with a cart.	Six men take a break in a field.
Test Image		
Ground Truth Descriptions	1. A gray bird stands majestically on a beach while waves roll in. 2. A white crane stands tall as it looks out upon the ocean. 3. A tall bird is standing on the sand beside the ocean. 4. A large bird stands in the water on the beach. 5. A water bird standing at the ocean 's edge.	1. A group of people stares at a wall that is filled with drawings in a building. 2. There are 5 people here looking at some pictures on the wall. 3. Five people are taking in an exhibit of Japanese art. 4. People watching at the arts in exhibition. 5. Five people looking at artwork.
Predicted Description	A bird stands tall on the beach.	Five people looking art in a building.

Table VI aims to showcase how well the model aligns for Flickr30k dataset. Four sets of test images are listed in the table, together with the accompanying ground truth descriptions and the model's predicted descriptions for each. This table is as an assessment tool, demonstrating the model's capacity to produce accurate and relevant descriptions for a range of images found in the Flickr30k dataset.

Table VII is designed to demonstrate the alignment of the model with the actual content of the images, as indicated by the provided ground truth descriptions for the Flickr8k dataset. Four sets of test images are included in the table, each with the corresponding ground truth descriptions and the predicted descriptions from the model. This table serves as a means of assessment, showcasing the model's capacity to provide accurate and perceptive descriptions for a range of images from the Flickr8k dataset.

TABLE VII
QUANTITATIVE RESULTS OBTAINED ON FLICKR8K DATASET

Test Image		
Ground Truth Descriptions	1. A winter landscape with four people walking in the snow. 2. Beautiful snowy landscape with people treading through the snow. 3. Cross-country skiers are traveling towards the mountains at sunset. 4. Four people walking across thick snow during a sunset. 5. The sun is almost behind the snowy mountains.	1. A beautiful sunset with three people in a boat on the lake. 2. As the sunsets 3 people are on a small boat enjoying the view. 3. Three people are in a canoe on a calm lake with the sun reflecting yellow. 4. Three people are on a boat in the middle of the water while the sun is in the back. 5. Three people in a boat float on the water at sunset.
Predicted Description	Four people walk through a snowy mountain.	Three people enjoying a beautiful sunset from a boat.
Test Image		
Ground Truth Descriptions	1. "a crowd wearing red , cheers on the red football team." 2. Football players in red congratulate each other as crowds in red cheer behind. 3. The Oklahoma Sooners football team discuss their game while fans cheer. 4. Two football players talk during a game. 5. Two Oklahoma Sooner football players talk on the sideline.	1. The two small dogs run through the grass. 2. Two fluffy white dogs running in green grass. 3. Two small dogs run through the grass. 4. Two small dogs that look almost identical are playing in the grass. 5. Two yellow dogs run together in green grass.
Predicted Description	A crowd in red cheers on the football team.	Two small dogs run through the green grass.

When compared to reference descriptions, the model produces predicted descriptions during the training and evaluation phases. The model's efficacy in capturing visual semantics and context is evaluated by comparing the predicted and real descriptions using metrics such as BLEU 1-4, CIDEr, METEOR and ROUGE-L. Table VIII provides an overview of a framework that has been suggested for the development of image descriptions from the three datasets.

TABLE VIII
OVERALL RESULTS OBTAINED FROM THE PROPOSED MODEL

Dataset	B-1	B-2	B-3	B-4	C	M	R-L
MSCOCO	0.893	0.821	0.794	0.725	1.88	0.78	0.689
Flickr30k	0.767	0.654	0.647	0.456	1.679	0.478	0.567
Flickr8k	0.784	0.678	0.538	0.479	1.381	0.389	0.654

The performance metrics of the proposed model are shown in the table for the MSCOCO, Flickr30k and Flickr8k datasets. These metrics include BLEU-1 to BLEU-4, CIDEr, METEOR and ROUGE-L scores. The suggested framework obtained BLEU scores for MSCOCO of 0.893 (B-1), 0.821 (B-2), 0.794 (B-3) and 0.725 (B-4) and for CIDEr, METEOR and ROUGE-L, 1.483, 0.358 and 0.789, respectively. With BLEU scores ranging from 0.767 (B-1) to 0.456 (B-4), bolstered by a CIDEr score of 1.679 and METEOR (0.478) and ROUGE-L (0.567) ratings suggesting sufficient matching and overlap, competitive performance was observed on the Flickr30k dataset. A CIDEr score of 1.381, favorable matches, and overlap was shown by BLEU scores on the Flickr8k dataset, which varied from 0.784 (B-1) to 0.479 (B-4). METEOR (0.389) and ROUGE-L (0.654)

scores also showed favorable matches and overlap. Across a range of benchmark datasets, the proposed approach performs well overall in producing image descriptions. The graphical depiction of the outcomes is shown in Figure 3. These scores

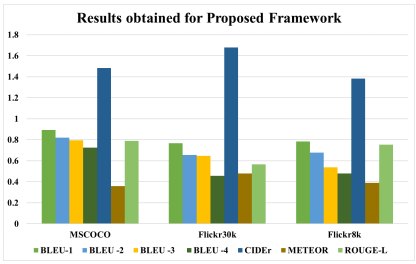


Fig. 3. Graphical representation of the results obtained from the datasets for the proposed Tri-FusionNet framework.

serve as evaluations of the model's performance in generating descriptions that align well with the reference descriptions. Higher scores indicate a higher degree of similarity and quality in the generated descriptions.

Table IX presents a qualitative comparison between the ground truth descriptions and the descriptions generated by the proposed model, highlighting successful and unsuccessful predictions and error analysis.

TABLE IX
QUALITATIVE ANALYSIS OF GENERATED IMAGE DESCRIPTIONS

Test Image	Ground Truth Description	Generated Description	Remarks/Error Analysis
	A white dog is holding a purple frisbee in its mouth on the green grass.	White dog holding purple frisbee on grass.	**Success:** Correct identification of objects and actions.
	A gold and black motorcycle parked on a paved surface road.	A car parked on road.	**Failure:** Incorrect object recognition.
	Three ducks stand by a calm pond with a wooden fence in front of them.	Three ducks stand by pond.	**Success:** Correct identification of objects and actions.
	A red and white vintage plane on display in a museum.	Spaceship preparing to launch.	**Failure:** Incorrect object recognition.

The following table compares the generated descriptions of four sample images with ground truth annotations to evaluate the performance of an image captioning model. Successful examples include correctly identifying "three ducks on a peaceful pond" and "a white dog clutching a purple frisbee on grass." These cases demonstrate the model's ability to recognize objects and actions in straightforward scenarios. However, the model exhibits notable failures, such as mistaking a gold and black motorcycle for "a car parked on a road" and, in row 4, misidentifying a red and white vintage plane as "a spaceship preparing to launch." These errors highlight the model's difficulty in accurately identifying complex objects or understanding the context of the scene. This analysis provides valuable insights into the model's strengths in simple situations and its limitations when faced with more intricate or ambiguous visuals, offering guidance for improving the model's training and recognition capabilities.

C. Comparison with Other State-of-the-art Methods:

A comprehensive generalisation analysis highlights the performance of the proposed model outside of the training set by assessing its capacity to produce image descriptions over a wide range of visual domains and datasets. A comparison study of image description generation models using the MSCOCO dataset is presented in Table X, which evaluates model performance using metrics such as BLEU, METEOR, CIDEr and Rouge-L.

TABLE X
COMPARATIVE ANALYSIS OF IMAGE DESCRIPTION GENERATION MODELS ON MSCOCO DATASET

Model	BLEU	METEOR	CIDEr	Rouge-L
Topic-based multi-channel attention (TMA) model [21]	0.658	-	0.800	0.534
Transformer based local graph semantic attention (TLGSA) [22]	0.724	-	1.003	0.534
Dynamic-balanced double-attention [23]	0.741	0.254	1.107	0.537
Meshed Transformer [7]	0.816	0.294	1.293	0.592
Global Enhanced Transformer [24]	0.816	0.284	1.301	0.591
Multimodal Transformer [8]	0.817	0.294	1.30	0.596
S^2 Transformer [12]	0.811	0.296	1.335	0.591
PMA-Net [15]	0.847	0.305	1.414	0.613
HAAV [16]	0.810	0.302	1.415	-
SPT Transformer [18]	0.812	0.296	1.344	0.592
Geometry attention transformer [25]	0.811	0.384	1.27	0.591
Vision-enhanced and Consensus-aware Transformer [26]	0.822	0.296	1.345	0.596
X-transformer + Faster RCNN [27]	0.821	0.296	1.334	0.598
Local-global guidance for transformer [28]	0.861	0.392	1.329	0.651
SAMT [11]	0.774	0.284	1.205	0.572
Tri-FusionNet (Proposed)	**0.893**	**0.780**	**1.880**	**0.689**

The performance of different models on the MSCOCO dataset indicates strong progress in image description generation. Models like Meshed Transformer [7], Global Enhanced Transformer [24], and Multimodal Transformer [8] show strong overall performance in terms of high scores obtained in BLEU, METEOR, CIDEr, and Rouge-L metrics. In contrast, models like Topic-based multi-channel attention (TMA) [21], Transformer-based local graph semantic attention (TLGSA) [22], and Dynamic-balanced double-attention [23] have relatively weaker performance, especially in terms of BLEU and CIDEr scores. Geometry Attention Transformer [25] and S^2 Transformer [12] obtain the best CIDEr scores, while Local-global guidance for transformer [28] reaches the best Rouge-L score of 0.651. Among the best-performing models, PMA-Net [15] and HAAV [16] obtain CIDEr scores higher than 1.4. In addition, PMA-Net obtains the best BLEU score of 0.847 among all the models. However, Tri-FusionNet breaks all other records as the new baseline, obtaining state-of-the-art performance scores for both BLEU 0.893, METEOR 0.780, CIDEr 1.880, and also Rouge-L with a 0.689 score in all the tested metrics, which highlights its capacity in producing better descriptions with semantic accuracy of an image for all the existing models.

To further validate the performance of the proposed Tri-FusionNet model, we evaluated it on the online MSCOCO test server, which is widely used as a benchmark for image description generation. The comparative results are presented in Table XI, highlighting the superior performance of Tri-

FusionNet in all the metrics, demonstrating its efficiency in generating contextually rich and accurate image descriptions.

TABLE XI
COMPARATIVE ANALYSIS OF IMAGE DESCRIPTION GENERATION MODELS ON ONLINE MSCOCO TEST SERVER

Model	BLEU	METEOR	Rouge-L	CIDEr
NIC [19]	0.714	0.302	0.521	0.945
m-RNN [29]	0.753	0.301	0.593	0.926
ReviewNet [30]	0.810	0.301	0.609	0.967
SCN [31]	0.828	0.309	0.619	1.013
Adaptive [32]	0.834	0.311	0.628	1.051
Att2all [33]	0.859	0.312	0.635	1.157
GateCap [34]	0.863	0.319	0.644	1.190
LSTM-A [35]	0.862	0.312	0.635	1.170
Up-Down [36]	0.877	0.322	0.648	1.192
RFNet [37]	0.877	0.327	0.657	1.240
GCN-LSTM [38]	-	0.330	0.660	1.259
SGAE [39]	0.882	0.327	0.661	1.252
AoANet [40]	0.880	0.338	0.667	1.282
Tri-FusionNet (Proposed)	**0.885**	**0.750**	**0.678**	**1.580**

Table XII compares various image description generation models using the Flickr30k dataset. The analysis includes metrics for evaluating the models' performance, such as BLEU, METEOR, CIDEr, and Rouge-L.

TABLE XII
COMPARATIVE ANALYSIS OF IMAGE DESCRIPTION GENERATION MODELS ON FLICKR30K DATASETS

Model	BLEU	METEOR	CIDEr	Rouge-L
Topic-based multi-channel attention (TMA) model [21]	0.650	-	0.334	0.436
Transformer based local graph semantic attention (TLGSA) [22]	0.643	-	0.450	0.489
Dynamic-balanced double-attention [23]	0.678	0.209	0.517	0.500
HAAV [16]	0.743	0.251	0.856	-
Multimodal Transformer [8]	0.744	0.236	-	-
Local-global guidance for transformer [28]	0.758	0.263	0.708	0.560
X-transformer + Faster RCNN [27]	0.753	0.253	0.707	0.543
Tri-FusionNet (Proposed)	**0.767**	**0.478**	**1.679**	**0.567**

Models like Transformer-based local graph semantic attention (TLGSA) [22], Topic-based multi-channel attention (TMA) [21], and Dynamic-balanced double-attention [23] show lower performance, particularly in BLEU and CIDEr, indicating limitations in generating high-quality descriptions. While the Multimodal Transformer [8] performs well in BLEU and METEOR, it lacks comprehensive metric coverage. Local-global guidance for Transformer [28] and X-transformer + Faster RCNN [27] demonstrate strong BLEU and CIDEr scores but fall slightly behind in Rouge-L. HAAV [16] achieves notable performance, especially with a CIDEr score of 0.856. However, the proposed Tri-FusionNet outperforms all models, achieving the highest scores across all metrics: 0.767 for BLEU, 0.478 for METEOR, 1.679 for CIDEr, and 0.567 for Rouge-L, establishing it as the most effective model for generating accurate and contextually rich descriptions on the Flickr30k dataset.

Table XIII presents a comparative analysis of various image description generation models on the Flickr8k dataset. It offers evaluation metrics for evaluating how well various models perform, including BLEU, METEOR, CIDEr and Rouge-L.

Models like TLGSA based on transformer [22], Vision encoder-decoder [41], Topic-based multi-channel attention [21], and Optimal transformers with beam search [42] have performance relatively lower and are bad at BLEU and CIDEr score, due to the lesser ability for producing high quality and relevant descriptions. Specifically, even though SAMT [11] has scores of BLEU and METEOR comparatively better values, it also lacks other competitive values of CIDEr and Rouge-L. On the other hand, proposed Tri-FusionNet obtains excellent results with setting new state-of-the-art for all metrics: BLEU-0.784, METEOR-0.389, CIDEr-1.381, and Rouge-L - 0.654. Such results emphasize that the proposed Tri-FusionNet model is more likely to produce exceptional, meaningful, and expressive image descriptions compared to existing models for Flickr8k.

TABLE XIII
COMPARATIVE ANALYSIS OF IMAGE DESCRIPTION GENERATION MODELS ON FLICKR8K DATASET

Model	BLEU	METEOR	CIDEr	Rouge-L
Topic-based multi-channel attention (TMA) model [21]	0.630	-	0.472	0.465
Transformer based local graph semantic attention (TLGSA) [22]	0.659	-	0.471	0.565
Vision encoder decoder [41]	0.395	0.177	0.380	0.297
Optimal transformers with Beam Search [42]	0.634	0.1987	0.520	-
SAMT [11]	0.682	0.212	-	0.448
Tri-FusionNet (Proposed)	**0.784**	**0.389**	**1.381**	**0.654**

V. CONCLUSION AND FUTURE DIRECTION

In this paper, a new model, Tri-FusionNet, for generating image descriptions that combine the CLIP transformer, RoBERTa, and the Vision Transformer with dual-attention processes, was presented. The proposed model outperforms state-of-the-art models as a result of extensive trials on the MSCOCO, Flickr30k, and Flickr8k datasets. It achieved notable gains across important evaluation criteria, such as BLEU, METEOR, CIDEr, and ROUGE. The model is successful because it can capture both local and global image data, use the CLIP transformer to effectively align textual and visual modalities, and employ RoBERTa for improved language understanding. The suggested model has useful applications in real-time image captioning systems, including autonomous car systems for scene detection, assistive technologies for the blind, and content management systems for automatic image tagging. Deploying the model still presents difficulties, though, such as managing edge devices' demanding processing requirements, ensuring the system is resilient to threats and noisy inputs, and resolving ethical challenges like bias in generated captions. Furthermore, the model's performance may be impacted by the unpredictability introduced by real-world settings, such as occlusions, dim lighting, computational cost, and other social circumstances. To enable more extensive real-world applications, future research will concentrate on improving fine-tuning procedures, tackling deployment issues, and extending the model's capabilities through sophisticated multimodal fusion techniques.

Supplementary Material The supplementary material contains additional experimental details, architectures and algorithm of the proposed work. The file `supplementary.txt` is available alongside this manuscript in the submission system.

Acknowledgment The Malware and Cyber-Forensics Lab at

Delhi Technological University, New Delhi, India, is acknowledged by the authors for providing the resources required to complete the research.

Data Availability Statement All the dataset links are provided in the paper.

Conflict of Interest The authors declare that they have no conflict of interest.

REFERENCES

[1] T. Ghandi, H. Pourreza, and H. Mahyar, "Deep learning approaches on image captioning: A review," *ACM Computing Surveys*, vol. 56, no. 3, pp. 1–39, 2023.

[2] L. Agarwal and B. Verma, "From methods to datasets: A survey on image-caption generators," *Multimedia Tools and Applications*, vol. 83, no. 9, pp. 28 077–28 123, 2024.

[3] K. Han, Y. Wang, H. Chen, X. Chen, J. Guo, Z. Liu, Y. Tang, A. Xiao, C. Xu, Y. Xu *et al.*, "A survey on vision transformer," *IEEE transactions on pattern analysis and machine intelligence*, vol. 45, no. 1, pp. 87–110, 2022.

[4] V. Atliha and D. Šešok, "Text augmentation using bert for image captioning," *Applied Sciences*, vol. 10, no. 17, p. 5978, 2020.

[5] X. Pan, T. Ye, D. Han, S. Song, and G. Huang, "Contrastive language-image pre-training with knowledge graphs," *Advances in Neural Information Processing Systems*, vol. 35, pp. 22 895–22 910, 2022.

[6] G. Li, L. Zhu, P. Liu, and Y. Yang, "Entangled transformer for image captioning," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 8928–8937.

[7] M. Cornia, M. Stefanini, L. Baraldi, and R. Cucchiara, "Meshed-memory transformer for image captioning," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 10 578–10 587.

[8] J. Yu, J. Li, Z. Yu, and Q. Huang, "Multimodal transformer with multi-view visual representation for image captioning," *IEEE transactions on circuits and systems for video technology*, vol. 30, no. 12, pp. 4467–4480, 2019.

[9] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark *et al.*, "Learning transferable visual models from natural language supervision," in *International conference on machine learning*. PMLR, 2021, pp. 8748–8763.

[10] J. Chen, D. Zhu, X. Shen, X. Li, Z. Liu, P. Zhang, R. Krishnamoorthi, V. Chandra, Y. Xiong, and M. Elhoseiny, "Minigpt-v2: large language model as a unified interface for vision-language multi-task learning," *arXiv preprint arXiv:2310.09478*, 2023.

[11] X. Yang, Y. Yang, S. Ma, Z. Li, W. Dong, and M. Woźniak, "Samt-generator: A second-attention for image captioning based on multi-stage transformer network," *Neurocomputing*, vol. 593, p. 127823, 2024.

[12] P. Zeng, H. Zhang, J. Song, and L. Gao, "S2 transformer for image captioning." in *IJCAI*, 2022, pp. 1608–1614.

[13] X. Yang, Y. Yang, J. Wu, W. Sun, S. Ma, and Z. Hou, "Ca-captioner: A novel concentrated attention for image captioning," *Expert Systems with Applications*, vol. 250, p. 123847, 2024.

[14] H. Parvin, A. R. Naghsh-Nilchi, and H. M. Mohammadi, "Image captioning using transformer-based double attention network," *Engineering Applications of Artificial Intelligence*, vol. 125, p. 106545, 2023.

[15] M. Barraco, S. Sarto, M. Cornia, L. Baraldi, and R. Cucchiara, "With a little help from your own past: Prototypical memory networks for image captioning," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 3021–3031.

[16] C.-W. Kuo and Z. Kira, "Haav: Hierarchical aggregation of augmented views for image captioning," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2023, pp. 11 039–11 049.

[17] T. Yao, Y. Li, Y. Pan, Y. Wang, X.-P. Zhang, and T. Mei, "Dual vision transformer," *IEEE transactions on pattern analysis and machine intelligence*, 2023.

[18] H. Zhang, P. Zeng, L. Gao, X. Lyu, J. Song, and H. T. Shen, "Spt: Spatial pyramid transformer for image captioning," *IEEE Transactions on Circuits and Systems for Video Technology*, 2023.

[19] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan, "Show and tell: A neural image caption generator," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3156–3164.

[20] J. Li, D. Li, C. Xiong, and S. Hoi, "Blip: Bootstrapping language-image pre-training for unified vision-language understanding and generation," in *International conference on machine learning*. PMLR, 2022, pp. 12 888–12 900.

[21] K. Qian and L. Tian, "A topic-based multi-channel attention model under hybrid mode for image caption," *Neural Computing and Applications*, vol. 34, no. 3, pp. 2207–2216, 2022.

[22] K. Qian, Y. Pan, H. Xu, and L. Tian, "Transformer model incorporating local graph semantic attention for image caption," *The Visual Computer*, pp. 1–12, 2023.

[23] C. Wang and X. Gu, "Dynamic-balanced double-attention fusion for image captioning," *Engineering Applications of Artificial Intelligence*, vol. 114, p. 105194, 2022.

[24] J. Ji, Y. Luo, X. Sun, F. Chen, G. Luo, Y. Wu, Y. Gao, and R. Ji, "Improving image captioning by leveraging intra-and inter-layer global representation in transformer network," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 35, no. 2, 2021, pp. 1655–1663.

[25] C. Wang, Y. Shen, and L. Ji, "Geometry attention transformer with position-aware lstms for image captioning," *Expert systems with applications*, vol. 201, p. 117174, 2022.

[26] S. Cao, G. An, Z. Zheng, and Z. Wang, "Vision-enhanced and consensus-aware transformer for image captioning," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 10, pp. 7005–7018, 2022.

[27] L. Lou, K. Lu, and J. Xue, "A novel cross-fusion method of different types of features for image captioning," in *2023 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2023, pp. 1–8.

[28] H. Parvin, A. R. Naghsh-Nilchi, and H. M. Mohammadi, "Transformer-based local-global guidance for image captioning," *Expert Systems with Applications*, vol. 223, p. 119774, 2023.

[29] J. Mao, "Deep captioning with multimodal recurrent neural networks (m-rnn)," *arXiv preprint arXiv:1412.6632*, 2014.

[30] Z. Yang, Y. Yuan, Y. Wu, W. W. Cohen, and R. R. Salakhutdinov, "Review networks for caption generation," *Advances in neural information processing systems*, vol. 29, 2016.

[31] Z. Gan, C. Gan, X. He, Y. Pu, K. Tran, J. Gao, L. Carin, and L. Deng, "Semantic compositional networks for visual captioning," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 5630–5639.

[32] J. Lu, C. Xiong, D. Parikh, and R. Socher, "Knowing when to look: Adaptive attention via a visual sentinel for image captioning," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 375–383.

[33] S. J. Rennie, E. Marcheret, Y. Mroueh, J. Ross, and V. Goel, "Self-critical sequence training for image captioning," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7008–7024.

[34] S. Wang, L. Lan, X. Zhang, and Z. Luo, "Gatecap: Gated spatial and semantic attention model for image captioning," *Multimedia Tools and Applications*, vol. 79, pp. 11 531–11 549, 2020.

[35] T. Yao, Y. Pan, Y. Li, Z. Qiu, and T. Mei, "Boosting image captioning with attributes," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 4894–4902.

[36] P. Anderson, X. He, C. Buehler, D. Teney, M. Johnson, S. Gould, and L. Zhang, "Bottom-up and top-down attention for image captioning and visual question answering," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 6077–6086.

[37] W. Jiang, L. Ma, Y.-G. Jiang, W. Liu, and T. Zhang, "Recurrent fusion network for image captioning," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 499–515.

[38] T. Yao, Y. Pan, Y. Li, and T. Mei, "Exploring visual relationship for image captioning," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 684–699.

[39] X. Yang, K. Tang, H. Zhang, and J. Cai, "Auto-encoding scene graphs for image captioning," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 10 685–10 694.

[40] L. Huang, W. Wang, J. Chen, and X.-Y. Wei, "Attention on attention for image captioning," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 4634–4643.

[41] A. Abdelaal, N. F. ELshafey, N. W. Abdalah, N. H. Shaaban, S. A. Okasha, T. Yasser, M. Fathi, K. M. Fouad, and I. Abdelbaky, "Image captioning using vision encoder decoder model," in *2024 International Conference on Machine Intelligence and Smart Innovation (ICMISI)*. IEEE, 2024, pp. 101–106.

[42] A. Shetty, Y. Kale, Y. Patil, R. Patil, and S. Sharma, "Optimal transformers based image captioning using beam search," *Multimedia Tools and Applications*, vol. 83, no. 16, pp. 47 963–47 977, 2024.

frontiers | Frontiers in Sustainable Cities

TYPE Specialty Grand Challenge
PUBLISHED 30 April 2025
DOI 10.3389/frsc.2025.1595280

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
James Evans,
The University of Manchester,
United Kingdom

*CORRESPONDENCE
Prashant Kumar
✉ p.kumar@surrey.ac.uk

RECEIVED 17 March 2025
ACCEPTED 31 March 2025
PUBLISHED 30 April 2025

CITATION
Kumar P, Sahani J, Corada Perez K, Ahlawat A, Andrade MdF, Athanassiadou M, Cao S-J, Collins L, Dey S, Di Sabatino S, Halios CH, Harris F, Horton C, Inostroza L, Jones L, Rodding Kjeldsen T, McCallan B, McNabola A, Kumar Mishra R, Kumar Mishra S, Sperandio Nascimento EG, Owen G, Ravindra K, Reis NC, Soebarto V, Wenk J, Wood HS and Yao R (2025) Urban greening for climate resilient and sustainable cities: grand challenges and opportunities.
Front. Sustain. Cities 7:1595280.
doi: 10.3389/frsc.2025.1595280

COPYRIGHT
© 2025 Kumar, Sahani, Corada Perez, Ahlawat, Andrade, Athanassiadou, Cao, Collins, Dey, Di Sabatino, Halios, Harris, Horton, Inostroza, Jones, Rodding Kjeldsen, McCallan, McNabola, Kumar Mishra, Kumar Mishra, Sperandio Nascimento, Owen, Ravindra, Reis, Soebarto, Wenk, Wood and Yao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Urban greening for climate resilient and sustainable cities: grand challenges and opportunities

Prashant Kumar[1,2,3]*, Jeetendra Sahani[1], Karina Corada Perez[4], Ajit Ahlawat[5,6], Maria de Fatima Andrade[7], Maria Athanassiadou[8], Shi-Jie Cao[1,3], Lisa Collins[9], Sagnik Dey[10,11], Silvana Di Sabatino[12], Christos H. Halios[13], Felicity Harris[14], Colin Horton[15], Luis Inostroza[16,17], Laurence Jones[18,19], Thomas Rodding Kjeldsen[20], Ben McCallan[21,22], Aonghus McNabola[23,24], Rajeev Kumar Mishra[25], Sumit Kumar Mishra[26,27,28], Erick G. Sperandio Nascimento[1,29,30], Gwilym Owen[31], Khaiwal Ravindra[32], Neyval C. Reis[33], Veronica Soebarto[34], Jannis Wenk[35,36], Hannah Sloan Wood[37] and Runming Yao[13,38,39]

[1]Faculty of Engineering and Physical Sciences, Global Centre for Clean Air Research (GCARE), School of Engineering, Civil and Environmental Engineering, University of Surrey, Guildford, United Kingdom, [2]Institute for Sustainability, University of Surrey, Guildford, United Kingdom, [3]School of Architecture, Southeast University, Nanjing, China, [4]Sustainability Research Institute, University of East London, London, United Kingdom, [5]Department of Geoscience and Remote Sensing, Delft University of Technology (TU Delft), Delft, Netherlands, [6]Leibniz Institute for Tropospheric Research (TROPOS), Leipzig, Germany, [7]Atmospheric Sciences Department, University of São Paulo, São Paulo, Brazil, [8]Urban Climate Applications, Met Office, Exeter, United Kingdom, [9]School of Biosciences, University of Surrey, Guildford, United Kingdom, [10]Centre for Atmospheric Sciences, Indian Institute of Technology Delhi, New Delhi, India, [11]Department of Health, Policy and Management, Korea University, Seoul, Republic of Korea, [12]Department of Physics and Astronomy "Augusto Righi", University of Bologna, Bologna, Italy, [13]School of Built Environment, University of Reading, Reading, United Kingdom, [14]Portsmouth City Council, Portsmouth, United Kingdom, [15]Rugby Borough Council, Rugby, United Kingdom, [16]Faculty of Regional Development and International Studies, Mendel University in Brno, Brno, Czechia, [17]Universidad Autónoma de Chile, Santiago de Chile, Chile, [18]UK Centre for Ecology & Hydrology, Environment Centre Wales, Bangor, United Kingdom, [19]Department of Geography and Environmental Science, Liverpool Hope University, Liverpool, United Kingdom, [20]Department of Architecture and Civil Engineering, University of Bath, Bath, United Kingdom, [21]Surrey County Council, Kingston upon Thames, United Kingdom, [22]Zero Carbon Guildford, Guildford, United Kingdom, [23]School of Engineering, RMIT University, Melbourne, VIC, Australia, [24]School of Engineering, Trinity College Dublin, Dublin, Ireland, [25]Department of Environmental Engineering, Delhi Technological University, New Delhi, India, [26]CSIR National Physical Laboratory, New Delhi, India, [27]Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, Uttar Pradesh, India, [28]Department of Chemistry, University of British Columbia, Vancouver, BC, Canada, [29]Faculty of Engineering and Physical Sciences, Surrey Institute for People-Centred Artificial Intelligence, University of Surrey, Guildford, United Kingdom, [30]Stricto Sensu Department, SENAI CIMATEC University, Salvador, Bahia, Brazil, [31]Resilience Unit, Cardiff Council, Cardiff, United Kingdom, [32]Department of Community Medicine and School of Public Health, Post Graduate Institute of Medical Education and Research (PGIMER), Chandigarh, India, [33]Department of Environmental Engineering, Federal University of Espírito Santo, Vitória, ES, Brazil, [34]School of Architecture and Civil Engineering, The University of Adelaide, Adelaide, SA, Australia, [35]Department G—Qualitative Hydrology, Federal Institute of Hydrology (BfG), Koblenz, Germany, [36]Department of Chemical Engineering, University of Bath, Bath, United Kingdom, [37]Independent Consultant, Copenhagen, Denmark, [38]Joint International Research Laboratory of Green Buildings and Built Environments, Ministry of Education, Chongqing University, Chongqing, China, [39]National Centre for International Research of Low-carbon and Green Buildings (Ministry of Science and Technology), Chongqing University, Chongqing, China

KEYWORDS

urban greening, climate adaptation, UN SDGs, green infrastructure, strategic urban planning

Highlights

- Urban greening can improve air quality, health, biodiversity, and climate resilience.
- Collaboration among communities and stakeholders is key to urban greening success.
- Community engagement ensures acceptance and long-term success of green infrastructure.
- Barriers like funding, regulations, and resistance impede green infrastructure projects.
- Accurate data and planning are essential for maximizing green infrastructure benefits.

Urban greening plays a crucial role in enhancing climate resilience, environmental quality, public health, and societal wellbeing. Policy makers at all levels are increasingly embracing greening and other nature-based solutions; however, successful implementation of these approaches requires a multidisciplinary strategy involving collaboration, community engagement, and adaptive interventions. This paper synthesizes key insights from an expert panel, comprising representatives of government agencies, research institutions, private sector, and local authorities, forming an international panel of experts, convened by the RECLAIM Network Plus (https://www.reclaim-network.org). Funded by the United Kingdom Research and Innovation (UKRI), this network project brings together over 650 members from 40 plus countries. It provides national leadership in urban greening, serving as a "one-stop-shop" for towns and cities to access green and blue infrastructure support, resources, and peer connections. The paper highlights grand challenges, priorities, successful case studies and opportunities in urban greening initiatives. The importance of strategic planning is also addressed, together with technological advancements and access to suitable data for maximizing the impact of green infrastructure as well as innovative funding models, such as corporate social responsibility and green finance. This work highlights the importance of an integrated, inclusive, and forward-thinking approach to urban greening for more resilient, sustainable, and equitable cities. Further, it emphasizes the need for cross-sector collaboration among local authorities, researchers, and businesses, as well as community involvement in successful planning for both short- and long-term outcomes. Ultimately, urban greening strategies must be informed by future climate scenarios while prioritizing equity and social justice to ensure adaptation options benefitting all communities.

1 The role of urban greening in climate resilience and liveability

Urban greening, encompassing nature-based solutions (NbS) and green-blue-gray infrastructure (GBGI), has emerged as a vital strategy for building climate-resilient and livable cities, addressing biodiversity loss, and supporting Sustainable Development Goals (SDGs) (Wang et al., 2023; Kumar, 2021; Jones et al., 2022a). As urban populations continue to grow and climate change intensifies, urban communities worldwide are facing environmental challenges, including extreme heat, increased flooding, deteriorating air quality, noise pollution and biodiversity loss. Effective integration of urban greening—incorporating street trees, parks, urban forests, wetlands, green roofs and green walls, and permeable surfaces (Kumar et al., 2024a,b)—contributes to enhancing environmental sustainability and public wellbeing (Astell-Burt and Feng, 2019). One of the most critical roles of urban greening is its contribution to climate resilience. Cities experience the urban heat island (UHI) effect, where built-up areas absorb and retain heat, leading to higher temperatures compared to surrounding rural areas (Oke, 1982; Yang et al., 2024). Urban greening helps mitigate this phenomenon by providing shade, cooling through evapotranspiration, and reducing heat absorption. Green spaces, such as parks, tree-lined streets, and gardens, can significantly lower local temperatures, making cities more comfortable spaces, while also reducing energy demand and mitigating heat emissions from cooling systems (Yang et al., 2024; Kumar et al., 2024b). Unlike gray infrastructure, NbS do not further exacerbate the root causes of climate change through reliance on emissions-intensive products such as concrete. The cooling effect of vegetation becomes even more crucial during heatwaves, which are becoming more frequent due to climate change. By incorporating greening, cities can help mitigate heat-related illnesses, reduce strain on healthcare systems, and help prevent heat-induced impacts on worker productivity (Jones et al., 2024). Beyond temperature regulation, integrating blue (water bodies) and green (vegetation) elements with gray infrastructure enables cities to develop more holistic and sustainable flood management strategies compared to the traditional drainage systems. While impervious surfaces in cities prevent rainwater from naturally soaking into the ground, leading to increased runoff and urban flooding (Redfern et al., 2016), GBGI and Sustainable Drainage Systems (SuDS) offer effective solutions (Vesuviano et al., 2025). By mimicking natural water cycles, GBGI and SuDS not only reduce flood risk but also improve water quality, replenish groundwater reserves, and prevent polluted stormwater overflow that can contaminate natural waterways and compromise vital ecosystems (Monachese et al., 2025). GBGI and SuDS also serve functional purposes besides significantly enhancing the aesthetic quality of urban environments. The presence of green spaces and water features has been demonstrated to contribute to recreational opportunities, aesthetic enjoyment, and the overall wellbeing of urban residents (Johnson and Geisendorf, 2022). Furthermore, these systems can serve as educational opportunities concerning water management and environmental stewardship (Well and Ludwig, 2022). Finally, integrating NbS, such as reforestation for slope stabilization, into urban planning enhances climate resilience by mitigating landslide risks. Integrating NbS into conventional risk reduction infrastructure can significantly enhance public wellbeing and promote long-term sustainability (Alcántara-Ayala, 2025; Devanand et al., 2023).

Urban greening offers significant advantages in improving air quality (Abhijith et al., 2017; Islam A. et al., 2024; Kaur et al., 2025). Trees and vegetation function as natural air purifiers, efficiently capturing airborne pollutants such as particulate matter (Kumar et al., 2024a). In heavily polluted urban areas, a well-designed urban greening strategy can lead to significant health benefits, including reduced respiratory illnesses, improved public health and lower healthcare costs (Mueller et al., 2022; Jones et al., 2019; Nemitz et al., 2020). Beyond improving local air quality, urban greening

reduces atmospheric concentrations of carbon dioxide through carbon sequestration (Page et al., 2021). However, some challenges remain in balancing NbS implementation to avoid hindering the vertical dispersion of air pollution emitted at ground level, such as the case of avenue planted trees in deep street canyons (Buccolieri et al., 2009; Abhijith et al., 2017; Kumar et al., 2019a).

Urban greening initiatives contribute to biodiversity conservation while concurrently enhancing environmental stewardship among urban populations (Roslund et al., 2021; Prescott et al., 2016). The integration of fragmented urban green spaces like urban forests, wetlands, and community gardens provides crucial habitats for diverse wildlife, promotes ecological balance, and ensures essential services such as pollination and pest control (Bowler et al., 2025; Dainese et al., 2019). While the Intergovernmental Panel on Climate Change recommends 30%–50% natural reserves globally to facilitate the natural recovery of terrestrial and marine ecosystems (IPCC, 2023), urban greening could partially contribute to this aim by connecting fragmented spaces in urban areas, thus supporting broader conservation efforts.

Beyond its environmental benefits, urban greening plays a vital role in improving social wellbeing (Kumar et al., 2019b). Green spaces provide residents with accessible opportunities for recreation, relaxation, and exercise in nature, leading to improved physical and mental health. Access to these areas has been shown to reduce stress, anxiety, and depression (Cohen-Cline et al., 2015), while encouraging physical activity and social interaction (Pretty and Barton, 2020; Pretty et al., 2003). Urban parks, community gardens, and tree-lined streets foster social cohesion by providing gathering spaces for diverse communities. Specifically, in high-density urban areas, small interventions such as green walls, green gates or pocket parks can substantially improve residents' quality of life (Abhijith et al., 2025; Kumar et al., 2024a). Despite the many beneficial aspects of NbS, several challenges remain in their implementation, as discussed in Sections 2–4.

2 Evolution of urban greening: past, present, and future

2.1 The past: an evolution of urban greening

Urban development has undergone significant transformations over the centuries, driven by industrialization, technological advancements, and changing social and economic conditions, among other factors. In Europe, the Roman concept of "*rus in urbe*," or countryside created within a building or city, was an early integration of natural landscapes into urban environments. In South Asia, the ancient towns in the Indus Valley in the Indian subcontinent exemplify early urban planning incorporating nature (Gregory, 2008). This includes traditional water harvesting structures and sustainable water management (Sharma, 2006). Mesopotamian and Persian civilizations notably exemplified by Persian gardens from sixth century BC Achaemenid dynasty, integrated green spaces into their landscapes for aesthetic, climatic and spiritual purposes, often representing paradise (UNESCO World Heritage Centre, 2025). Early advancements in urban sanitation infrastructure were excavated at Mohenjo-daro (2600–1800 BC) in present-day Pakistan (Jansen, 1985). Pre-industrial green spaces in Europe were often tied to aristocratic estates, religious institutions, or communal lands, such as Paris' Luxembourg Gardens established in the early seventeenth century and London's 142-hectare Hyde Park, established in the sixteenth century, which was initially developed for private hunting purposes (Self, 2014) before it was opened to the public in the following century (Porter, 1995). The gardens of Schönbrunn Palace is another example of an early urban green space, dating back to the fourteenth century, originally planned as a hunting ground, and later opened for public usage in the eighteenth century (Schönbrunn, 2025).

Rapid urbanization during the Industrial Revolution led to the widespread replacement of natural landscapes with gray infrastructure: roads, buildings, and engineered drainage systems. While this contributed to economic development, it also resulted in significant environmental challenges, including air and water pollution, UHI, and the loss of green spaces. Some cities developed public parks as an attempt to address overcrowding and poor living conditions, including New York's Central Park and Sefton Park in Liverpool. There are also many historic examples of the recognition of the benefits of green pockets for the health and wellbeing of the local community, in smaller cities and towns. One such example is the 4 acres of St Thomas Leisure Ground in Exeter, opened on 30th March 1891 for Queen Victoria's Jubilee, at the instigation of the Local Board of Health (Devon Gardens Trust, 2025).

In the early twentieth century, the Garden City movement advocated for the integration of green elements within planned urban developments, such as green belts and parks (Howard, 1902). At the same time, many Western cities underwent a transition toward a car-centric urban development model as automobile ownership rates increased. In the United States, car registrations tripled from 8 million to 26 million between 1920 and 1930 (Gavanas, 2024). This resulted in urban sprawl, with many cities expanding their road networks significantly by the 1950s (Barrington-Leigh and Millard-Ball, 2015). Initiatives to rebalance rapid urbanization with ecological stewardship followed (Mumford, 1938). The 1960s and 1970s brought a renewed focus on urban greening, driven by environmental concerns, the degradation of urban environments and legislative pressure, including the establishment in the USA of the U.S. Environmental Protection Agency in 1970 and the passage of the Clean Air Act and Clean Water Act. A significant shift in urban planning philosophy manifested advocating for the integration of natural systems and a closer collaboration between urban designers and scientists (McHarg, 1969). The eco-urbanism movement of the 1980s put a renewed focus on nature as greening, and the 1990s to early 2000s brought growing concern over human and urban pressures on ecosystems, with the emerging concepts of green infrastructure and ecosystem services (Di Marino et al., 2023).

Today, over three-quarters of global cities have adopted urban greening policies, integrating sustainability with urban growth (Khamdamov and Usmanov, 2024). As cities worldwide grapple with the challenges of climate change, pollution, and declining biodiversity, urban greening has emerged as a key strategy for building urban resilience and enhancing liveability. Studies show

that greater green space availability is associated with better health outcomes such as lower mortality rates and improved mental health and sleep quality (Galitskaya et al., 2024). A green coverage between 10% and 20% has shown measurable health advantages (Nieuwenhuijsen et al., 2017). Neighborhoods maintaining at least 30% tree canopy coverage demonstrate significant improvements in mental, physical, and social health, as emphasized by the "3–30–300 rule" (Konijnendijk, 2023). The 3–30–300 rule offers clear urban greening guidelines, setting three minimum standards: (i) every citizen should be able to see at least three trees from their home, (ii) their neighborhood should have 30% tree canopy cover; and (iii) they should not live more than 300 m away from the nearest public park or green space.

2.2 The present: a new understanding of urban greening

In recent decades, urban greening has gained significant recognition for its importance for people's wellbeing, improving quality of life, and building sustainable cities (Di Marino et al., 2023). As topics such as climate resilience, environmental justice, and community engagement have become more prominent within built environment discourse, urban greening has emerged as a core element of sustainable and resilient urban planning. Urban planners and policymakers have embraced this concept, incorporating GBGI and NbS into their strategies, with GBGI exemplifying a multifunctional approach by integrating vegetation, water bodies, and built environment elements to create sustainable urban spaces.

Today, over 130 countries have incorporated NbS actions in their nationally determined SDG contributions, complementing technology-based solutions, including in urban contexts (IUCN, 2016). In relation to this, the International Union for Conservation of Nature and Natural Resources's Global Standard for NbS consolidates ecosystem-based approaches aligning urban greening strategies with SDGs (IUCN, 2020).

Various urban greening initiatives worldwide are aiming to address climate risks, SDGs, improving public health, and enhancing ecological sustainability. Examples are presented in chronological order of references:

- **Copenhagen's Cloudburst Management Plan**: A city-wide strategy designed to manage extreme rainfall events through the use of green roofs, permeable surfaces, and green corridors, which also enhances urban aesthetics and air quality (The City of Copenhagen, 2012).
- **UK's 25-Year Environment Plan, A Green Future**: This plan promotes biodiversity net gain (BNG) and SuDS to integrate water-sensitive urban design into city landscapes [Department for Environment, Food and Rural Affairs (DEFRA), 2018].
- **Indian Government's Sustainable Development Goals and Urban Local Bodies**: This guidance proposes key indices to promote and monitor sustainable urban stormwater management practices, aiming to enhance flood resilience and ensure effective drainage solutions in new developments (Government of India, 2018).
- **Welsh Government's National Standards for SuDS (2018)**: These standards require new developments of more than one dwelling or where an area covered by construction work equals or exceeds 100 m^2 to include SuDS features that comply with national standards (Welsh Government, 2018).
- **Natural England's Green Infrastructure Framework (2023)**: This tool highlights the institutional push for increasing green cover in urban residential areas, with an ambitious target of achieving 40% green cover in cities (Natural England, 2023).
- **South Australia's Long-Term Vision**: The State Government has created a vision for "a resilient and liveable Adelaide for all: cooler, leafier, and more biodiverse," setting priority actions for greener infill development and improving greening equity (Government of South Australia, 2024).
- **Singapore's City in Nature Plan**: This initiative aims to increase urban greening while promoting biodiversity conservation (Government of Singapore, 2025).
- **UKRI's Tomorrow's Engineering Research Challenges**: This strategic initiative by the umbrella UK science funding body prioritizes green materials and environmentally conscious engineering practices, aiming for engineering advancements that positively contribute to the planet (EPSRC, 2025).

Collectively, such initiatives highlight a fundamental shift in urban planning, where green spaces are not seen as mere decorative elements but as crucial infrastructure for resilience and wellbeing. Cities worldwide are adopting innovative solutions to address urban challenges, including the integration of green roofs and vertical gardens, in the context of urban resilience and adaptation planning (Saqib et al., 2024).

2.3 Future directions: technology-driven and inclusive urban greening

Urban greening is expected to become increasingly data-driven, leveraging technologies such as Geographic Information Systems (GIS), artificial intelligence (AI), and remote sensing to optimize the design and management of green infrastructure (Stessens et al., 2021; Rahmati, 2024). These tools will enable built environment professionals (e.g., landscape architects, developers, city planners) to more effectively assess environmental needs, target adaptation options, model the impact of greening interventions, and monitor the effectiveness of urban green spaces in a holistic way. Future urban greening will integrate Internet of Things (IoT), AI, and real-time analytics for smart water management, optimizing flow, and drainage efficiency (Langenheim et al., 2022). Digital twin technology and AI-driven planning will enable real-time simulations and optimized GBGI and SuDS (Dembski et al., 2020; Xu et al., 2024). Advanced permeable pavements and bioengineered materials will enhance infiltration and pollutant filtration, leveraging nanotechnology for contaminant absorption (Verma et al., 2024). In addition to technologically driven future approaches, there is also an increasing emphasis on citizen/community engagement in urban planning processes, however addressing structural barriers and ensuring meaningful participation remain ongoing challenges (Sondou et al., 2025).

Green finance, corporate sustainability and decision-making processes that are informed by a comprehensive economic valuation of the net benefits of urban greening (including its environmental, social, and health benefits), are also expected to play an important role in shaping the next phase of urban greening (Williams et al., 2025; Tefera et al., 2023). Policies promoting investment in green bonds, environmental, social, and governance (ESG) frameworks, and public–private partnerships are expected to accelerate large-scale greening projects. Additionally, smaller-scale community-led urban greening initiatives will be vital to ensuring that cities remain inclusive, mitigating the risk of "green gentrification," where the introduction of urban green spaces inadvertently displaces low-income residents due to rising property values.

Finally, as climate change intensifies, urban greening strategies will need to be resilient and adaptive. Cities will need to focus on climate-resilient landscaping considering their micro-climatic conditions, for example, selecting appropriate drought-resistant plant species, water-sensitive urban designs, and ecological corridors that promote biodiversity. Urban agriculture, vertical gardens or farms, and rooftop farms are also expected to gain prominence, addressing both food security and environmental sustainability in growing urban populations (Drottberger et al., 2023).

3 Best practice in urban greening

Urban areas worldwide have implemented innovative greening projects that serve as effective models for sustainable development. Inspiring case studies showcasing tangible benefits and strategies have emerged from urban greening initiatives, with platforms like the Urban Nature Atlas compiling best practices globally, as does the NbS Initiative's approach for rural areas (Urban Nature Atlas, 2025; Nature-based Solutions Initiative, 2025; Debele et al., 2023). Analyzing these case studies is crucial for understanding effective strategies, benefits, and lessons learned in urban greening, informing future urban planning and sustainability efforts.

3.1 Selected case studies

The following curated selection of 10 case studies, presented in alphabetical order of the name of city, has been drawn from authors' recommendations and discussions at the RECLAIM Network Plus Conference (http://www.reclaim-network.org):

1. Adelaide, South Australia has unveiled a street tree planting plan to increase canopy cover from 33% to 40% by 2035, complemented by green infrastructure initiatives like water-sensitive urban design (City of Adelaide., 2024). Similarly, Melbourne, Victoria, has implemented the Urban Forest Strategy 2021–2032, aiming for a "resilient, healthy, and diverse" urban environment. This strategy targets increasing tree canopy cover from 22% to 40% by 2040, while also focusing on urban forest diversity, vegetation health, soil moisture, water quality, and biodiversity enhancement (City of Melbourne., 2021).
2. Brazil's Ministry of the Environment initiated the Cidades+Verdes Program (França and Ramos de Almeida, 2021) to enhance urban green areas, aiming to improve quality of life for 85% of Brazilians living in cities. The Urban Environmental Registry (CAU) was created to map and share information about these areas, allowing municipalities to register urban green spaces and citizens to access and evaluate them. Urban greening initiatives in Brazil largely focus on rain gardens to address frequent storms and flooding (ICLEI, 2025). Examples include: (1) Belo Horizonte, Minas Gerais, implemented three rain gardens to improve urban drainage; (2) Niterói, Rio de Janeiro: The Parque Orla Piratininga Alfredo Sirkis project, a 680,000 m^2 linear park along the P iratininga Lagoon, featuring 35,290 m^2 of filtering gardens to protect and restore local ecosystems; (3) São Paulo, São Paulo, installed rain gardens across the city to mitigate flooding during the rainy summer season; and (4) Curitiba, Paraná, launched the city's first Urban Farm in 2020, dedicating 4,000 m^2 to organic food production, with plans for an additional 11,000 m^2 unit. Other regions have also implemented green infrastructure, primarily for flood prevention. These efforts highlight Brazil's commitment to nature-based solutions, enhancing resilience and urban quality of life.
3. Cardiff, Wales, United Kingdom has launched an ambitious "*One Planet Strategy*" (Cardiff Council, 2020). This 10-year afforestation programme aims to support Cardiff to be a carbon neutral city by 2030, increasing green cover and promoting community engagement in sustainable development. Since its inception, the programme has planted over 82,000 trees across the city, emphasizing local community participation and encouraging residents to take ownership of green spaces. Cardiff Council, with Welsh Government funding, is developing flood risk reduction plans for Whitchurch Brook, a historically heavily modified local watercourse. This project combines green infrastructure with SuDS techniques to control surface water flow while improving public space areas. The project integrates green infrastructure and SuDS to manage surface water and enhance public spaces. Design and planning are being led by a sustainable development consultancy, with community input shaping the solutions (Cardiff News Room, 2023). This participatory approach has not only strengthened environmental stewardship but also fostered social cohesion, demonstrating that urban greening can be both an ecological and a social asset.
4. Copenhagen, Denmark has implemented green roof policies to improve building insulation, manage stormwater, and enhance urban biodiversity (City of Copenhagen., 2010; Ansel and Appl, 2025). While green roofs reduce buildings' energy consumption, ground-level greening offers greater social and environmental benefits in dense urban areas. The city now emphasizes street-level greenery like pocket parks, green walls, and community gardens. Copenhagen's Østerbro "Climate Quarter" exemplifies this approach, featuring urban greening projects that mitigate flood risks and improve public spaces, such as Sankt Kjeld's Square and Taasinge Plads (Larsen, 2018).
5. In Exeter, England, United Kingdom, the campus of the UK Met Office headquarters has earned a Biodiversity

Benchmark certificate (The Wildlife Trusts, 2025), highlighting its commitment to maintaining a rich ecological zone amid a developed area (Met Office, 2021, 2025a). The project includes pollinator-friendly meadows, water-efficient landscaping, and protected habitats for local wildlife. This initiative exemplifies how large organizations can incorporate biodiversity conservation into their urban footprint, serving as a model for businesses looking to integrate sustainability into their operations.

6. The Zero Carbon Initiative in Guildford, England, United Kingdom combines urban tree planting with SuDS to mitigate urban heat stress and manage stormwater (Zero Carbon Guildford, 2025) as well as installation of green screens and school living gate (Abhijith et al., 2025). The combination of these strategies exemplifies a holistic approach to urban greening, where multiple ecological and infrastructural benefits can be achieved simultaneously.
7. Nanjing, Jiangsu Province, China is home to the Vertical Forest project, an innovative urban afforestation initiative integrating nature-based solutions into architecture. This project features two towers in the Pukou District, standing at 200 and 108 m, with facades adorned by alternating balconies and greenery. Housing over 800 trees and 2,500 shrubs across 4,500 m^2, the design incorporates 27 native species. This approach enhances biodiversity, improves air quality, reduces surface temperatures, and regulates humidity, ultimately enhancing urban livability (Boeri Architetti, 2025).
8. New York City's MillionTreesNYC program, a landmark public–private partnership, aimed to plant one million trees across the city, focusing on underserved neighborhoods lacking green spaces (Fisher et al., 2014). Exceeding its goal ahead of schedule, the project demonstrated effective collaboration among government agencies, non-profits, and local communities. The success of MillionTreesNYC highlights the importance of long-term maintenance and community engagement in urban greening projects.
9. Singapore, a global leader in urban greening, has pioneered nature integration into its cityscape through the "City in Nature" plan. This initiative builds on decades of green infrastructure development, expanding tree-lined streets, rooftop gardens, and vertical greenery on high-rises (Government of Singapore, 2025). The city's green corridors connect parks and natural reserves, fostering wildlife and providing accessible green spaces for residents. Driven by strong government policies, technological innovation, and public–private collaboration, Singapore's approach exemplifies how urban greening can be seamlessly incorporated into high-density environments, balancing economic growth with ecological sustainability.
10. The Trees for Homes (TfH) programme in South Africa is an urban greening initiative promoting biodiversity enhancement and climate adaptation through household tree planting. The initiative was established around 1992 as a flagship urban greening initiative by the organization Food and Trees for Africa. In the Zandspruit neighborhood, the TfH project incorporated key sustainable livelihood principles, including being people-centered, responsive and participatory, multi-level, conducted in partnership, sustainable and dynamic. However, it was also noted that a key barrier to success in vulnerable communities was instability of tenure (Sachikonye et al., 2016). The TfH initiative has recently partnered with Johannesburg Inner City Partnership to train local environmental champions to act as Community Educators, who receive specialist training in climate change, tree care, environmental ethics, and community engagement.

3.2 Learning from successful urban greening projects

The 10 case studies of successful projects offer valuable lessons for enhancing urban living through NbS, highlighting the importance of active community involvement, integrated urban design, and long-term sustainability in creating climate-resilient urban spaces. Innovative approaches, including transforming existing buildings into green spaces, further demonstrate the potential for integrating nature into densely built environments. These successful strategies provide insights for cities worldwide to develop context-specific, innovative solutions for more climate-resilient urban spaces.

The longevity of well-designed green spaces is exemplified by the UK's urban parks, such as St. Thomas People's Park in Exeter, established in 1891 (Exeter Memories, 2009). Originally created as a safe space for children and vulnerable people, the park continues to serve the local community over a century later, underscoring the importance of sustainable design in GBGI projects.

Maximizing the public health benefits of GBGI projects requires effective cross-sector collaboration. In Surrey, efforts to integrate public health expertise into environmental initiatives have shown promise, with general practitioners engaging in discussions about air pollution and climate change (Zero Carbon Guildford, 2025). However, these efforts often remain isolated highlighting the need for greater coordination to address overarching priorities such as health inequalities and community empowerment.

To enhance GBGI project effectiveness, public health practitioners should be involved in co-creation and planning stages, while challenges persist, including limited financial and human resources. Research from Cardiff Council and Vale University Health Board area has revealed that air pollution disproportionately affects poorer areas (Cardiff Council, 2025), emphasizing the need to align GBGI projects with public health priorities. Overcoming these barriers requires a coordinated approach that integrates public health as a primary driver of greening initiatives, rather than a secondary co-benefit, ensuring that urban planning prioritizes both environmental and health outcomes.

4 Key challenges and strategic solutions for implementing urban greening

Urban greening initiatives, despite their well-documented benefits, face numerous challenges including financial and

regulatory constraints, space limitations, impacts on the vertical dispersion of air pollution, maintenance issues, year-round effectiveness limitations due to climate variations and plant species types, lack of interdisciplinary collaboration, and resistance to innovation within established structures and processes (Grunewald et al., 2021; Haaland and van den Bosch, 2015). Overcoming these obstacles require supportive policies, innovative financing, technological advancements, and inclusive community engagement. Some of the key challenges and strategic solutions (Figure 1) are presented in subsequent sub-sections.

4.1 Financial constraints, sustainable funding, and economic value

Urban greening strategies, such as GBGI and NbS offer long-term economic benefits such as reduced healthcare costs, energy savings, and improved property values. However, high upfront costs for land acquisition, infrastructure development, maintenance and misconceptions deter both public and private investment (Riedman et al., 2022; Rowley et al., 2022). Municipal budgets are often stretched, and many greening initiatives rely on single-source or short-term funding, limiting their long-term viability (C40 Cities Climate Leadership Group, 2022).

Government budgets and grants typically cover only capital costs for urban greening projects, excluding or providing limited-term funding for ongoing maintenance and operational expenses only. Consequently, many projects face neglect due to insufficient resources for routine upkeep, resulting in degraded green spaces that fail to achieve their intended outcomes (Albro, 2020).

Establishing sustainable funding mechanisms, such as public-private partnerships, green bonds, and community-led stewardship programmes, can help address the challenge of maintaining urban greening over time (Nash et al., 2021). Importantly, green and surface-based climate adaptations are likely cheaper to implement than large traditional gray infrastructure projects like sewer expansions (Xu et al., 2021). To encourage continuing public investment, it is crucial to demonstrate the financial rewards of NbS, particularly in terms of climate change resilience for issues like flooding.

The health benefits associated with NbS provide another compelling argument for urban greening initiatives, especially when addressing air quality and heat stress. While general wellbeing improvements may be subjective, quantifying health costs related to these issues and assessing potential cost reductions through environmental improvements can justify financial investments in urban greening. Techniques to quantify the economic value of ecosystem services provided by urban green infrastructure (Mell et al., 2016; Jones et al., 2019; Fletcher et al., 2021) can position these solutions as cost-effective approaches to building climate resilience, while offering a dual benefit of improved public health and enhanced urban sustainability (Elmqvist et al., 2015).

Innovative financing mechanisms can provide solutions to these challenges. For example, green bonds, corporate ESG funding, crowdfunding, and public–private partnerships (PPPs) offer possible solutions to ensure long-term investment. Green bonds provide a structured financial tool for municipalities and businesses to raise capital for environmental projects, while corporate ESG funding encourages businesses to invest in urban greening as part of their sustainability commitments. In particular, PPPs can help to distribute financial risk while leveraging private sector efficiency and public sector oversight (Naumann et al., 2011). Crowdfunding platforms raise small contributions for urban greening projects like tree planting and parks. Social impact funds and blended finance combine public, private, and philanthropic investments for social and environmental returns, as seen in the World Bank's Urban Resilience Fund (World Bank, 2015). Green Real Estate Investment Trusts (REITs) finance eco-friendly real estate with integrated green spaces (Hin Ho et al., 2013). Utility-backed financing reinvests savings from reduced water, energy, or stormwater costs into green infrastructure. Additionally, cities can introduce incentives such as tax breaks or subsidies for developers who incorporate greening elements into their projects. Establishing dedicated municipal budgets for urban greening and integrating it into broader urban resilience plans ensures that green infrastructure is not treated as an afterthought but as a critical component of sustainable city planning. Finally, training programs for local communities in green space management and ecological restoration can create job opportunities while fostering a culture of environmental stewardship.

Economic valuation is crucial for effective urban greening and management, requiring comprehensive decision-support tools that accurately assess both benefits and costs. Such tools can guide policymakers in determining optimal quantities, locations, and species of plants, while balancing urban greening with infrastructure development. Current frameworks fall short in capturing the full spectrum of benefits, including improved air quality, carbon storage, water management, energy efficiency, and public health (Tefera et al., 2023). Consequently, sustainable greening costs often seem to outweigh benefits, stemming from inadequate long-term benefit accounting, limited predictive modeling capabilities, and insufficient reliable data. Addressing these challenges is essential for making informed urban greening decisions that create more sustainable and liveable cities.

4.2 Regulatory and governance challenges

Fragmented governance structures, unclear policies, and bureaucratic delays hinder the planning and execution of urban greening projects. Multiple agencies, including local governments, urban planners, environmental departments, and private developers, often operate independently, leading to siloed decision-making and slow project approvals. For example, policies such as the UK's Greening Policy, which aims for 40% green cover in urban residential areas (Natural England, 2023), may struggle with enforcement due to the limited capacity of local authorities. Similarly, at the local scale, initiatives like the Guildford Borough Council's goal to double Biodiversity Net Gain (BNG) from the standard 10% to 20% encounter implementation challenges, as biodiversity policies are established at the county level but enforced by district and borough councils. Similar issues arise with other planning commitments, such as affordable housing and green infrastructure, which are often approved but not

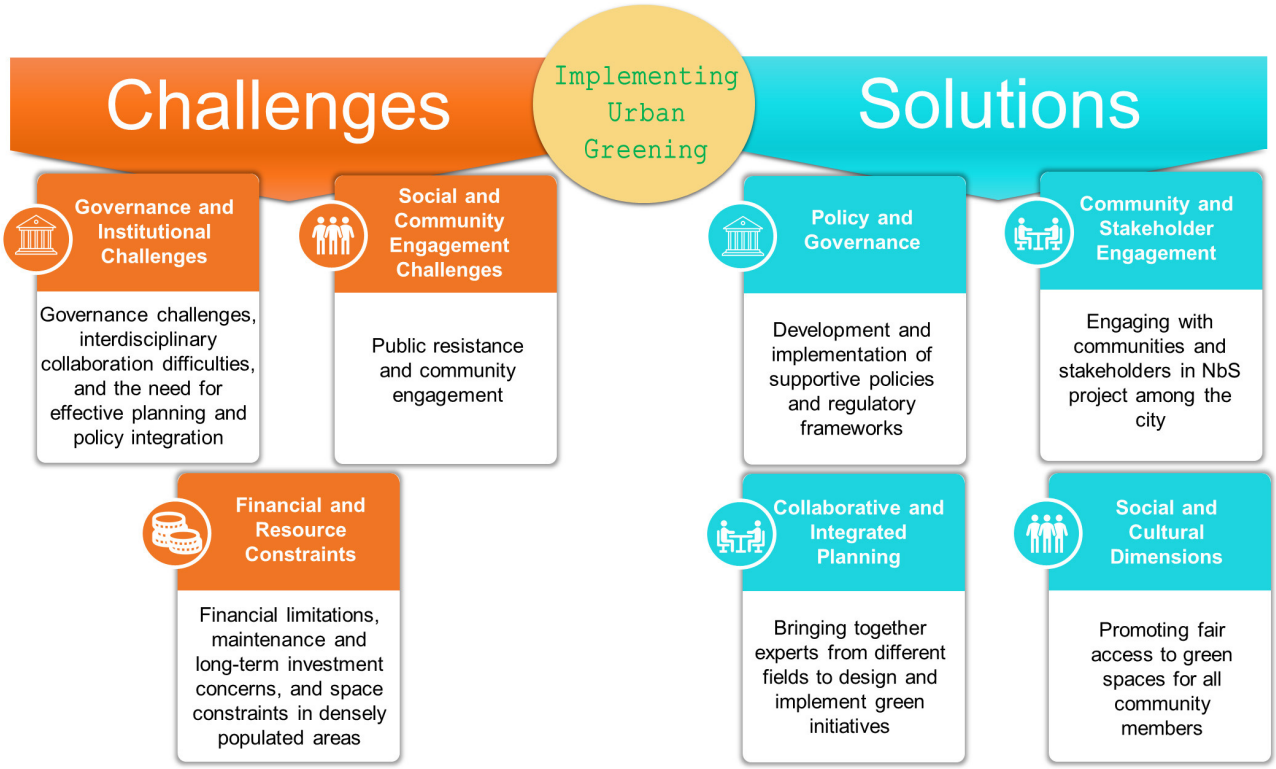


FIGURE 1
Challenges and solutions for implementing urban greening in cities. A summary of key barriers to urban greening, including limited land availability, high costs, regulatory constraints, and fragmented governance, alongside potential solutions such as multifunctional land use, streamlined policies, stakeholder collaboration, and innovative green infrastructure integration.

delivered due to cost and regulatory constraints. Addressing these challenges requires strategic planning, supportive policy frameworks, strengthening coordination among stakeholders, streamlining approval processes, and enhancing policy frameworks that actively promote urban greening.

Large city-scale urban greening should adopt agile management approaches that allow for flexibility and continuous improvement based on environmental and social feedback. Thus, urban greening should be embedded within broader climate adaptation plans to enhance resilience against environmental challenges such as heat waves, floods, and pollution. By aligning greening efforts with broader urban resilience strategies, cities can maximize the impact of adaptation options and ensure long-term sustainability. Governments should enforce green infrastructure requirements, such as BNG, green roofs, urban forests, and permeable surfaces in new developments. Strengthening urban greening legislation at national and local levels ensures that towns and cities can prioritize sustainable development over short-term economic gains.

Reducing bureaucratic delays and improving coordination between local governments, environmental agencies, and developers can accelerate project implementation. However, challenges in enforcing green infrastructure commitments can arise due to under-resourced planning departments. For instance, Surrey County Council, responsible for the local nature recovery strategy including a goal of planting 1.2 million trees, faced such difficulties in executing their green initiatives effectively (SCC, 2021).

4.3 Space constraints: balancing development and sustainability

Urban morphology plays a crucial role in determining the spatial extent and types of greening strategies that can be effectively implemented in cities (Owen et al., 2024; Kumar et al., 2024a,b). As urban areas expand, limited land availability forces green infrastructure to compete with development, making it crucial to plan ahead rather than retrofitting later. Existing urban structures, high costs, and regulatory barriers often constrain the implementation of green spaces, particularly when retrofitting developed areas. These challenges are compounded by the fact that stakeholders, such as shop owners and local communities, may have minimal influence over decisions regarding GBGI projects, especially in town centers.

To overcome these space constraints, urban planners must adopt innovative approaches and technologies. Integrating vertical gardens, rooftop gardens, green corridors, and pocket parks can maximize the use of limited space in developed areas. However, these solutions require strategic design to ensure they maximize ecological and social benefits without compromising urban functionality. Furthermore, urban planning strategies need to shift toward prioritizing multifunctional land use that balances development with environmental sustainability.

Effective implementation of GBGI solutions requires a holistic approach. This includes the ability to quantify the cost–benefit of green infrastructure compared to traditional gray solutions,

considering both immediate and long-term impacts. Moreover, it is essential to incorporate the "voice" and needs of local communities in GBGI interventions, whether in existing urban areas or new developments. By involving stakeholders in the decision-making process and adopting innovative planning strategies, cities can more effectively address space constraints and create sustainable, liveable urban environments that benefit both residents and the ecosystem.

Regarding green space distribution, community engagement is crucial (Riedman et al., 2022). It ensures that the focus extends beyond merely achieving percentage targets of green area coverage in a given region. Instead, it emphasizes placing appropriate solutions in strategic locations. The physical impacts of green spaces are spatially linked, diminishing with distance. Solely examining citywide percentages of green spaces may obscure the fact that benefits are not uniformly experienced across the entire urban landscape. Therefore, a targeted and communicative approach helps distribute benefits more effectively, particularly to disadvantaged communities.

4.4 Community engagement and public awareness campaigns

Community-driven initiatives, where residents are involved in the planning, maintenance, and stewardship of green spaces, can significantly enhance public support and project sustainability (Aslanoglu et al., 2025). Without proper engagement, communities may feel excluded from decision-making processes, leading to distrust and reluctance to support greening efforts. Additionally, in some cases residents and businesses may oppose greening projects due to misconceptions or factual concerns about land-use changes, maintenance responsibilities, and perceived inconveniences. These could be related with concerns about green gentrification, including increased property taxes, possibilities of fallen tree limbs, changes in neighborhood aesthetics and in traffic management; all these can potentially lead to opposition from local communities (Nobles and Moore, 2024; Meenar et al., 2022).

Effective community engagement involves transparent communication, co-creating projects with local residents, and addressing concerns about land use changes. Participatory decision-making processes, including town hall meetings, co-creation workshops, and citizen-led greening programs, empower communities to take ownership of green spaces. Public participation in tree planting, park maintenance, and urban gardening has been shown to foster a sense of ownership and encourages community stewardship of green spaces (e.g., Jennings et al., 2016; Bressane et al., 2024). This not only enhances public support for urban greening but also strengthens social cohesion and long-term stewardship of these spaces. The public can also actively participate in maintaining the green space, contributing to its design, and engaging in green space science outreach initiatives. Involving diverse stakeholders, such as local governments, non-profit organizations, and private developers, ensures that greening projects receive the necessary financial and logistical support while being aligned with broader urban development goals. The collaboration with communities in co-creating solutions is essential for reflecting the needs and preferences of local residents.

Educating citizens on the environmental and health benefits provided by urban greening features such as the green spaces further reinforce the importance of urban greening (Nguyen and Chidthaisong, 2024). Schools, workplaces, and community organizations can actively promote greening initiatives, emphasizing the role individuals can play in supporting these efforts (Zhang et al., 2024). Integration of urban greening into public health initiatives can also increase awareness of their benefits and encourage behavioral changes that support greener lifestyles and green social prescribing (Kruize et al., 2019). Governments and NGOs can leverage digital platforms, citizen science initiatives, social media, and interactive workshops to engage diverse audiences and inspire collective action through established behavior change levers that include identity-based social contagion, and the use of trusted messengers, normative appeals and "shared destiny narratives."

4.5 Equity and inclusivity in urban greening

Urban greening initiatives must prioritize equity and inclusivity to ensure fair distribution of benefits across all communities. Socially disadvantaged groups often face disproportionate environmental challenges, making equitable access to green infrastructure crucial (Hsu et al., 2022; Richards et al., 2023; Williams et al., 2025). To address this, there is a need for: (1) Ensuring that all communities, especially low-income and marginalized groups, benefit equally from improved air quality, cooling effects, and recreational opportunities; (2) Addressing disparities by recognizing that disadvantaged communities often suffer worse environmental consequences compared with more affluent areas, and tailor interventions accordingly; (3) Mitigating gentrification by implementing policies to prevent displacement and maintain affordable housing alongside greening projects; (4) Guiding urban greening efforts with social justice considerations to create healthier, more inclusive urban environments for all residents (Anguelovski and Corbera, 2023). By incorporating these principles, cities can develop urban greening strategies that not only improve environmental quality but also promote social equity and community wellbeing.

One of the major challenges of urban greening is the risk of green gentrification, where improved environmental quality leads to increased property values, ultimately displacing low-income residents (Jelks et al., 2021). While new parks, tree-lined streets, and urban forests enhance the overall liveability of a city, they can also attract wealthier residents and investors, making housing unaffordable for marginalized communities. To counter this, policymakers must integrate greening initiatives with affordable housing policies. Rent control measures, community land trusts, and mixed-income housing developments can help ensure that existing residents are not forced out due to rising costs. Additionally, designing green spaces that cater to diverse socio-economic groups, such as multipurpose parks, community gardens, and culturally relevant recreational areas, can promote inclusivity and prevent the exclusive transformation of neighborhoods. Implementing greening guidelines such as the 3–30–300 rule can help reduce inequity of access to greenspace, and thereby avoid

green gentrification, although there are substantial challenges to retrospectively achieving these targets in densely built up cities (Owen et al., 2024). Nevertheless, it remains a valuable aspiration for cities wanting to increase their greenspace, especially in residential areas.

Ensuring urban greening aligns with social and environmental justice requires prioritizing communities that historically lack access to green spaces. Low-income neighborhoods, often characterized by limited tree cover and fewer parks, experience higher temperatures, greater vulnerability to disasters, increased air pollution, and lower overall wellbeing. Greening efforts should be directed toward these areas to reduce health disparities and improve climate resilience (Rigolon et al., 2021). Public health benefits, such as reduced heat-related illnesses, improved air quality, and enhanced mental wellbeing, should be central to urban greening policies. Moreover, incorporating climate justice principles, where the needs of vulnerable populations are at the forefront of urban planning, ensures that greening efforts do not disproportionately benefit affluent areas while neglecting marginalized communities.

Equitable urban greening also requires addressing intersectional issues, such as accessibility, safety, and cultural inclusivity (Endalew Terefe and Hou, 2024). Green spaces must be designed to accommodate people of all ages, abilities, and backgrounds. Features such as wheelchair-accessible paths, shaded rest areas, and appropriately-lit public parks improve accessibility and usability for all residents. Safety concerns, particularly in marginalized neighborhoods, should be addressed through community-led monitoring programs and strategic placement of greenery to avoid creating isolated areas. Furthermore, integrating cultural elements into green spaces, such as indigenous planting practices, spaces for traditional celebrations, and art installations reflecting local heritage, enhances community identity and encourages greater use of urban greening. Where minority communities feel a sense of belonging in a local greenspace, their experiences of greenspace are often very positive (Palmer et al., 2025).

4.6 Interdisciplinary collaboration for effective implementation

Effective GBGI design and implementation requires expertise from various fields, including urban planning, environmental science, engineering, architecture, public health, economics and social sciences (Butt and Dimitrijević, 2022). However, a lack of interdisciplinary coordination often results in disjointed efforts that fail to maximize the full potential of urban greening. For example, planners may focus on aesthetics, or succumb to pressures for more housing, without considering ecological benefits or giving ecological benefits less priority, while engineers may prioritize stormwater management, the landscape architect will value biodiversity conservation, and so on.

The diverse knowledge base required to deliver a comprehensive plan, points to the need of a project manager or other champion that is able to coordinate all the various parts into one harmonious whole (Figure 2). While ambitious targets for tree cover and green space exist, there is criticism that public health practitioners are not leading environmental and greening initiatives to the standards imposed e.g., by air pollution reduction goals. Local authorities acknowledge efforts to engage with health professionals but work often happens in silos due to resource constraints and lack of effective communications. Despite good research and monitoring, effective collaboration remains limited.

Public health should not just be a co-benefit but a driving force in greening strategies, especially as pollution and health inequalities are often concentrated in disadvantaged areas. Encouraging cross-sector collaboration, knowledge-sharing, and integrated planning approaches can lead to more holistic and effective urban greening solutions. By proactively addressing these challenges, cities can harness the full potential of green infrastructure to enhance climate resilience, biodiversity, and overall liveability for future generations.

Urban greening calls for stronger collaboration between networks to achieve ambitious targets such as 40% green cover in urban areas (Natural England, 2023). Bringing together experts from urban planning, climate science, ecology, public health, and civil engineering can help foster new and holistic solutions that address multiple urban challenges simultaneously. Considering the needs of people in greenspace planning also involves understanding how their needs change as they go about their lives in the city, locally where they live, but also where they work and travel in between. Concepts such as "people-sheds" help formalize these requirements (Jones et al., 2022b). Cities should create cross-sector partnerships among local governments, businesses, academic institutions, and non-governmental organizations (NGOs) to pool resources, knowledge, and expertise. Such collaborations can accelerate project implementation, enhance innovation, and ensure urban greening aligns with broader sustainability and resilience goals.

4.7 Leveraging technological advancements

Data-driven planning plays an important role in optimizing urban greening initiatives. Technologies such as Geographic Information System, Machine learning, artificial intelligence, and availability of suitable, spatially coherent data (remote sensing, aerial imagery or satellite photography), enable spatial risk assessments at unprecedented resolution, thus enabling identification of areas most in need of green infrastructure (Fletcher et al., 2021). By harnessing these new tools, urban planners can maximize the ecological and social benefits of green spaces while ensuring efficient resource allocation. AI techniques have been applied to optimize green-blue infrastructure for climate resilience, analyzing complex environmental data to improve the design and management of green and blue spaces (Shaamala et al., 2024). Other works showed that AI-driven models coupled with publicly available satellite data can improve urban planning, resource allocation, and green blue space conservation (Hasan et al., In press). AI and multimodal datasets have the potential to help create holistic and adaptable systems that integrate green and blue infrastructure to address interconnected climate-related challenges

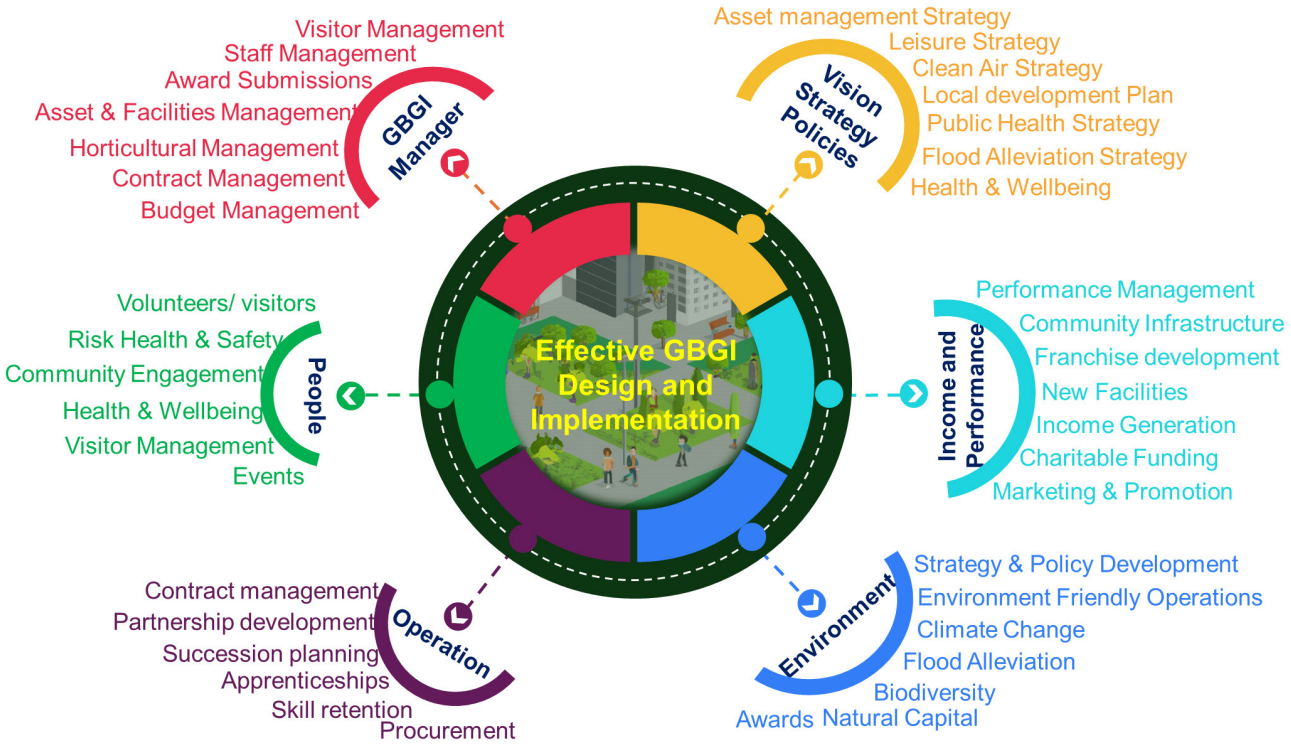


FIGURE 2
Dimensions for effective GBGI design and implementation. This diagram illustrates the interconnected dimensions for a successful GBGI project, including GBGI manager, people, operation, environment, income and performances and vision, strategy and policies. Adapted from APSE (2019).

such as flood control, biodiversity enhancement, and urban cooling (Islam M. R. et al., 2024).

The UK Met Office, for example, is committed to improving climate projections and ensuring their practical use for organizations and individuals (Met Office, 2025b). Additionally, organizations like UKRI are funding projects that generate large-scale environmental data, such as mapping the heat registry of every house in the UK to identify areas needing intervention. Ensuring that all funded research is open-source allows stakeholders, authorities and researchers to access valuable data and target interventions or research gaps. However, a major challenge lies in managing the volume of data and making it easily accessible. Addressing this requires investment in large-scale data infrastructure and smart interface solutions.

4.8 Long-term management strategies

Adaptive management strategies are essential for sustaining urban greening efforts in the long run, along with monitoring and evaluation. Regular environmental assessments using remote sensing technologies, satellite imagery, and GIS mapping can help monitor the health and effectiveness of green spaces and ecosystems. These tools enable urban planners to identify areas in need of intervention, track biodiversity trends, and assess the cooling impact of different greening solutions. Integrating urban greening with climate adaptation policies, such as flood mitigation through green infrastructure, drought-resistant landscaping, and urban heat island reduction strategies, enhances the resilience of cities against climate risks. Moreover, maintaining green spaces requires sustainable funding and governance structures. Many urban greening initiatives face long-term financial challenges, particularly when initial investments are not accompanied by maintenance plans (Pearsall et al., 2024). Innovative financing mechanisms, such as green bonds, corporate social responsibility (CSR) contributions, and public–private partnerships, can help secure ongoing funding.

5 Conclusion and future outlook

Urban areas have undergone a transformation from walkable, compact spaces to car-dominated sprawl. In recent years, there has been an increased focus on the development of sustainable and environmentally-friendly urban areas. This transition highlights the complex interplay of technological, social, and environmental factors that influence such developments. As urbanization accelerates, the demand for sustainable urban development has increased significantly. This urgency is further emphasized by global initiatives, such as the Sustainable Development Goals (SDGs), especially SDG11, which focuses on Sustainable Cities and Communities. Additionally, other SDGs, such as SDG 3, 7, 8, 9, 12, and 13 underline the importance of creating resilient and sustainable urban environments.

Urban greening is not just an aesthetic enhancement but a fundamental component of sustainable urban development.

Integrating GBGI into new and existing urban areas can help tackle environmental challenges while creating healthier, more livable spaces. With strategic planning, investment, and community engagement, urban greening can transform cities into climate-resilient, vibrant, and inclusive places for present and future generations, addressing the issues of extreme heat, air pollution, and loss of biodiversity. However, effective implementation of GBGI requires a multidisciplinary approach that aligns policy support, financial mechanisms, community engagement, and technological advancements. By addressing key challenges and leveraging collaborative governance, innovative funding, and data-driven decision-making, cities can successfully scale up GBGI, shifting from gray infrastructure dominance to climate-resilient inclusive urban landscapes.

Looking forward, urban greening will increasingly be shaped by technological advancements, innovative financing mechanisms, and a stronger emphasis on community participation and social equity. By embracing cutting-edge strategies and fostering multi-sectoral collaboration, future urban areas can strike a balance between development and sustainability, ensuring a greener, healthier, and more resilient world for future generations. Effective urban greening requires a comprehensive strategy that includes policy support, innovative financing, technological advancements, interdisciplinary collaboration, and community participation. Additionally, by addressing financial barriers, streamlining governance, utilizing data-driven decision-making, and fostering public engagement, cities can successfully implement urban greening initiatives. To ensure inclusive climate-resilient urban greening, urban areas must prevent green gentrification, prioritizing underprivileged communities, adopting adaptive management strategies, and securing long-term funding. As urban populations continue to grow and climate change intensifies, inclusive urban greening with careful and strategic planning is not just an option but a necessity for building just and sustainable cities.

Author contributions

PK: Project administration, Writing – original draft, Supervision, Investigation, Writing – review & editing, Conceptualization, Methodology, Funding acquisition. JS: Investigation, Methodology, Writing – review & editing, Data curation. KC: Data curation, Writing – review & editing, Methodology, Investigation, Visualization. AA: Methodology, Writing – review & editing, Resources, Validation. MFA: Validation, Writing – review & editing, Methodology. MA: Investigation, Validation, Writing – review & editing, Methodology. S-JC: Writing – review & editing, Validation, Investigation, Methodology. LC: Investigation, Writing – review & editing, Resources, Conceptualization. SDe: Validation, Writing – review & editing, Investigation, Methodology. SDi: Writing – review & editing, Validation, Conceptualization. CHH: Conceptualization, Validation, Writing – review & editing. FH: Conceptualization, Validation, Writing – review & editing. CH: Writing – review & editing, Conceptualization, Validation. LI: Writing – review & editing, Conceptualization, Validation. LJ: Funding acquisition, Validation, Conceptualization, Writing – review & editing. TR: Conceptualization, Validation, Funding acquisition, Writing – review & editing. BM: Conceptualization, Writing – review & editing, Validation. AM: Conceptualization, Writing – review & editing, Validation. RK: Conceptualization, Validation, Writing – review & editing. SK: Conceptualization, Validation, Writing – review & editing. ES: Writing – review & editing, Validation, Conceptualization. GO: Validation, Writing – review & editing, Conceptualization. KR: Conceptualization, Writing – review & editing, Validation. NR: Investigation, Writing – review & editing, Validation. VS: Conceptualization, Validation, Writing – review & editing. JW: Writing – original draft, Writing – review & editing, Validation. HW: Validation, Conceptualization, Writing – review & editing. RY: Conceptualization, Writing – review & editing, Validation.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work is conducted as part of the United Kingdom Research & Innovation (UKRI)-funded RECLAIM Network Plus (EP/W034034/1; EP/W033984/1). The network is jointly funded by three UKRI research councils: the Engineering and Physical Sciences Research Council (EPSRC), the Natural Environment Research Council (NERC), and the Arts and Humanities Research Council (AHRC). PK and co-leads acknowledge the support received through the UKRI funded GP4Streets (APP44894), DEFRAG (NE/W002892/1), GREENIN (APP55977), NERC-FAPESP funded GreenCities (NE/X002799/1; NE/X002772/1; FAPESP Grant No. 2022/02365-5), and the University Global Partnership Network (UGPN)-funded UGPN-NBS projects.

Acknowledgments

The authors gratefully acknowledge the contributions of all RECLAIM Conference panel members, attendees, co-leads, coordinating staff, event organizers, speakers, and collaborating partners. We also extend our appreciation to the University of Surrey's Global Centre for Clean Air Research (GCARE) team for their invaluable support in running both the conference and network operations. The names of co-authors are listed in alphabetical order following the core writing team.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abhijith, K. V., Kumar, P., Gallagher, J., McNabola, A., Baldauf, R. W., Pilla, F., et al. (2017). Air pollution abatement performances of green infrastructure in open road and built-up street canyon environments – a review. *Atmos. Environ.* 162, 71–86. doi: 10.1016/j.atmosenv.2017.05.014

Abhijith, K. V., Rawat, N., Emygdio, A. P. M., Le Den, C., Collins, K., Cartwright, P., et al. (2025). Demonstrating multi-benefits of green infrastructure to schools through collaborative approach. *Sci. Total Environ.* 958:177959. doi: 10.1016/j.scitotenv.2024.177959

Albro, S. L. (2020). "Chapter 5. Sustaining urban greening projects," in *Vacant to Vibrant: Creating Successful Green Infrastructure Networks* (Island Press), 105–123. Avaialble online at: https://link-springer-com.iclibezp1.cc.ic.ac.uk/chapter/10.5822/978-1-61091-901-2_6 (accessed March 10, 2025).

Alcántara-Ayala, I. (2025). Landslides in a changing world. *Landslides* 1–15. doi: 10.1007/s10346-024-02451-1

Anguelovski, I., and Corbera, E. (2023). Integrating justice in nature-based solutions to avoid nature-enabled dispossession. *Ambio* 52, 45–53. doi: 10.1007/s13280-022-01771-7

Ansel, W., and Appl, R. (2025). An international review of current practices and future trends: green roof policies. *CityGreen* 12. Avaialble online at: https://www.nparks.gov.sg/-/media/cuge/ebook/citygreen/cg2/cg2_02.pdf (accessed March 08, 2025).

APSE (2019). *Association for Public Service Excellence. Nurturing Skills for 21st Century Parks*. Avaialble online at: https://www.apse.org.uk/index.cfm/apse/members-area/special-interest-portals/parks-portal/research-papers/nurturing-skills-for-21st-century-parks/nurturing-skills-for-21st-century-parks/ (accessed March 13, 2025).

Aslanoglu, R., Kazak, J. K., Szewrański, S., Swiader, M., and Arciniegas, G., Chrobak, G., et al. (2025). Ten questions concerning the role of urban greenery in shaping the future of urban areas. *Build. Environ.* 267:112154. doi: 10.1016/j.buildenv.2024.112154

Astell-Burt, T., and Feng, X. (2019). Association of urban green space with mental health and general health among adults in Australia. *JAMA Netw Open* 2:e198209. doi: 10.1001/jamanetworkopen.2019.8209

Barrington-Leigh, C., and Millard-Ball, A. (2015). A century of sprawl in the United States. *Proc. Natl. Acad. Sci. U.S.A.* 112, 8244–8249. doi: 10.1073/pnas.1504033112

Boeri Architetti, S. (2025). *Nanjing Vertical Forest*. Avaialble online at: https://www.stefanoboeriarchitetti.net/en/project/nanjing-vertical-forest/ (accessed March 05, 2025).

Bowler, D. E., Callaghan, C. T., Felappi, J. F., Mason, B. M., Hutchinson, R., Kumar, P., et al. (2025). Evidence-base for urban green-blue infrastructure to support insect diversity. *Urban Ecosyst.* 28, 1–14. doi: 10.1007/s11252-024-01649-4

Bressane, A., Loureiro, A. I. S., and Almendra, R. (2024). Community engagement in the management of urban green spaces: prospects from a case study in an emerging economy. *Urban Sci.* 8:188. doi: 10.3390/urbansci8040188

Buccolieri, R., Gromke, C., Sabatino, D., i., and Ruck, S. B. (2009). Aerodynamic effects of trees on pollutant concentration in street canyons. *Sci. Total Environ.* 407, 5247–5256. doi: 10.1016/j.scitotenv.2009.06.016

Butt, A. N., and Dimitrijević, B (2022). Multidisciplinary and transdisciplinary collaboration in nature-based design of sustainable architecture and urbanism. *Sustainability* 14:10339. doi: 10.3390/su141610339

C40 Cities Climate Leadership Group (2022). *Policy Briefs. Greening City Budgets: Practical Approaches*. Avaialble online at: https://www.c40knowledgehub.org/s/article/Greening-city-budgets-Practical-approaches?language=en_US (accessed March 13, 2025).

Cardiff Council (2020). *One Planet Cardiff - Our Vision for a Carbon Neutral City by 2030*. Avaialble online at: https://www.oneplanetcardiff.co.uk/wp-content/uploads/OPC%20vision%20document.pdf (accessed March 03, 2025).

Cardiff Council (2025). *Clean Air Cardiff*. Avaialble online at: https://www.cardiff.gov.uk/ENG/resident/Parking-roads-and-travel/clean-air-cardiff/Pages/default.aspx (accessed March 10, 2025).

Cardiff News Room (2023). Avaialble online at: https://cardiffnewsroom.co.uk/releases/c25/31655.html (accessed March 12, 2025).

City of Adelaide. (2024). *Strategic Plan 2024–2028*. Avaialble online at: https://d31atr86jnqrq2.cloudfront.net/docs/strategic-plan-web.pdf (accessed March 03, 2025).

City of Copenhagen. (2010). *Green Roofs in Copenhagen*. City of Copenhagen. Technical and Environmental Administration. Avaialble online at: https://kk.sites.itera.dk/apps/kk_pub2/pdf/1017_sJ43Q6DDyY.pdf (accessed March 08, 2025).

City of Melbourne. (2021). *Urban Forest Strategy 2021-2032*. Avaialble online at: https://mvga-prod-files.s3.ap-southeast-4.amazonaws.com/public/2024-07/urban-forest-strategy.pdf (accessed March 03, 2025).

Cohen-Cline, H., Turkheimer, E., and Duncan, G. E. (2015). Access to green space, physical activity and mental health: a twin study. *J. Epidemiol. Commun. Health* 69, 523–529. doi: 10.1136/jech-2014-204667

Dainese, M., Martin, E. A., Aizen, M. A., Albrecht, M., Bartomeus, I., Bommarco, R., et al. (2019). A global synthesis reveals biodiversity-mediated benefits for crop production. *Sci. Adv.* 5:eaax0121. doi: 10.1126/sciadv.aax0121

Debele, S. E., Leo, L. S., Kumar, P., Sahani, J., Ommer, J., Bucchignani, E., et al. S. (2023). Nature-based solutions can help reduce the impact of natural hazards: a global analysis of NBS case studies. *Sci. Total Environ.* 902:165824. doi: 10.1016/j.scitotenv.2023.165824

Dembski, F., Wössner, U., Letzgus, M., Ruddat, M., and Yamu, C. (2020). Urban digital twins for smart cities and citizens: the case study of Herrenberg, Germany. *Sustainability* 12:2307. doi: 10.3390/su12062307

Department for Environment, Food and Rural Affairs (DEFRA) (2018). *A Green Future: Our 25 Year Plan to Improve the Environment*. HM Government. Avaialble online at: https://assets.publishing.service.gov.uk/media/5ab3a67840f0b65bb584297e/25-year-environment-plan.pdf (accessed March 10, 2025)

Devanand, V. B., Mubeen, A., Vojinovic, Z., Sanchez Torres, A., Paliaga, G., Abdullah, A. F., et al. (2023). Innovative methods for mapping the suitability of nature-based solutions for landslide risk reduction. *Land* 12:1357. doi: 10.3390/land12071357

Devon Gardens Trust (2025). *St Thomas Pleasure Gardens*. Avaialble online at: https://devongardenstrust.org.uk/gardens/st-thomas-pleasure-gardens (accessed March 10, 2025).

Di Marino, M., Tiitu, M., Saglie, M. I. L., and Lapintie, K. (2023). Conceptualizing 'green' in urban and regional planning – the cases of Oslo and Helsinki. *Euro. Plann. Stud.* 32, 1187–1209. doi: 10.1080/09654313.2023.2285811

Drottberger, A., Zhang, Y., Yong, J. W. H., and Dubois, M.-C. (2023). Urban farming with rooftop greenhouses: a systematic literature review. *Renew. Sustain. Energy Rev.* 188:113884. doi: 10.1016/j.rser.2023.113884

Elmqvist, T., Setälä, H., Handel, S., van der Ploeg, S., Aronson, J., Blignaut, J., et al. (2015). Benefits of restoring ecosystem services in urban areas. *Curr. Opin. Environ. Sustain.* 14:101. doi: 10.1016/j.cosust.2015.05.001

Endalew Terefe, A., and Hou, Y. (2024). Determinants influencing the accessibility and use of urban green spaces: a review of empirical evidence. *City Environ. Interact.* 24:100159. doi: 10.1016/j.cacint.2024.100159

EPSRC (2025). *Corporate Report. Tomorrow's Engineering Research Challenges*. Avaialble online at: https://www.ukri.org/publications/tomorrows-engineering-research-challenges/ (accessed March 03, 2025).

Exeter Memories (2009). *St Thomas Pleasure Ground*. Avaialble online at: https://www.exetermemories.co.uk/em/_parks/stthomas.php (accessed March 06, 2025).

Fisher, D. R., Campbell, L. K., and Svendsen, E. S. (2014). *MillionTreesNYC: The Integration of Research and Practice*. New York City Department of Parks and Recreation. Avaialble online at: https://research.fs.usda.gov/treesearch/45580 (accessed March 04, 2025).

Fletcher, D. H., Likongwe, P., Chiotha, S., Nduwayezu, G., Mallick, M., Uddin, M. N., et al. (2021). Using demand mapping to assess the benefits of urban green and blue space in cities from four continents. *Sci. Total Environ.* 785:147238. doi: 10.1016/j.scitotenv.2021.147238

França, A. L., and Ramos de Almeida, A. P. (2021). *Programa Cidades+Verdes*. Ministério do Meio Ambiente: Secretaria de Qualidade Ambiental. Avaialble online at: https://www.gov.br/mma/pt-br/acesso-a-informacao/acoes-e-programas/programa-projetos-acoes-obras-atividades/agendaambientalurbana/cidadesmaisverdes/ProgramaCidadesmaisVerdes.pdf (accessed March 11, 2025).

Galitskaya, P., Luukkonen, A., Roslund, M. I., Mänttäri, M., Yli-Viikari, A., Tyrväinen, L., et al. (2024). Green space quantity and exposure in relation to the risk of immune-mediated diseases: a scoping review. *BMC Public Health* 24:3358. doi: 10.1186/s12889-024-20655-x

Gavanas, N. (2024). "Intra-city clean, smart and sustainable mobility," in *Self-Sufficiency and Sustainable Cities and Regions*, ed. Eduardo Medeiros (Routledge), 200. doi: 10.4324/9781003498216-9

Government of India (2018). *Sustainable Development Goals and Urban Local Bodies*. United Nations Resident Coordinator's Office. Avaialble online at: https://www.local2030.org/library/540/Sustainable-Development-Goals-and-Urban-Local-Bodies.pdf (accessed March 11, 2025).

Government of Singapore (2025). *City in Nature: Key Strategies*. Avaialble online at: https://www.nparks.gov.sg/who-we-are/city-in-nature-key-strategies (accessed March 04, 2025).

Government of South Australia (2024). *Urban Greening Strategy: A Resilient and Liveable Adelaide for All*. Avaialble online at: https://yoursay.sa.gov.au/93122/widgets/434016/documents/284479 (accessed March 04, 2025).

Gregory, P. (2008). "City planning in ancient India," in *Encyclopaedia of the History of Science, Technology, and Medicine in Non-Western Cultures*, ed. H. Selin (Dordrecht: Springer), 576–577.

Grunewald, K., Bastian, O., Louda, J., Arcidiacono, A., Brzoska, P., Bue, M., et al. (2021). Lessons learned from implementing the ecosystem services concept in urban planning. *Ecosyst. Serv.* 49:101273. doi: 10.1016/j.ecoser.2021.101273

Haaland, C., and van den Bosch, C. K (2015). Challenges and strategies for urban green-space planning in cities undergoing densification: a review. *Urban Forestry Urban Green.* 14, 760–771. doi: 10.1016/j.ufug.2015.07.009

Hasan, M. M., Pramanik, M., Alam, I., Kumar, A., Avtar, R., and Zhran, M. (In press). Assessing the efficacy of artificial intelligence based city-scale blue green infrastructure mapping using Google Earth Engine in the Bangkok metropolitan region. *J. Urban Manage.* doi: 10.1016/j.jum.2024.11.009

Hin Ho, K., Rengarajan, S., and Han Lum, Y. (2013). "Green" buildings and Real Estate Investment Trust's (REIT) performance. *J. Property Investment Finance* 31, 545–574. doi: 10.1108/JPIF-03-2013-0019

Howard, E. (1902). *Garden Cities of Tomorrow*. London: S. Sonnenschein and Co., Ltd. Avaialble online at: https://www.gutenberg.org/files/46134/46134-h/46134-h.htm (accessed March 10, 2025).

Hsu, Y. Y., Hawken, S., Sepasgozar, S., and Lin, Z. H. (2022). Beyond the backyard: GIS analysis of public green space accessibility in Australian metropolitan areas. *Sustainability* 14:4694. doi: 10.3390/su14084694

ICLEI (2025). *Governos Locais pela Sustentabilidade*. Brazil. Avaialble online at: https://americadosul.iclei.org/how-brazilian-cities-are-implementing-nature-based-solutions/ (accessed March 11, 2024).

IPCC (2023). "Summary for policymakers," in *Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, eds. Core Writing Team, H. Lee and J. Romero (Geneva: IPCC), 1–34.

Islam, A., Pattnaik, N., Moula, Md. M., Rötzer, T., Pauleit, S., et al. (2024). Impact of urban green spaces on air quality: a study of PM_{10} reduction across diverse climates. *Sci. Total Environ.* 955:176770. doi: 10.1016/j.scitotenv.2024.176770

Islam, M. R., Talukdar, S., Rihan, M., and Rahman, A. (2024). Evaluating cooling effect of blue-green infrastructure on urban thermal environment in a metropolitan city: using geospatial and machine learning techniques. *Sustain. Cities Soc.* 113:105666. doi: 10.1016/j.scs.2024.105666

IUCN (2016). *Nature-Based Solutions to Address Global Societal Challenges*, eds. E. Cohen-Shacham, G. Walters, C. Janzen, and S. Maginnis (Gland: IUCN), xiii + 97. Avaialble online at: https://portals.iucn.org/library/sites/library/files/documents/2016-036.pdf (accessed March 11, 2025).

IUCN (2020). *IUCN Global Standard for Nature-Based Solutions: A User-Friendly Framework for the Verification, Design and Scaling up of NbS, 1st Edn*. Avaialble online at: https://portals.iucn.org/library/node/49070 (accessed March 09, 2025).

Jansen, M. (1985). Mohenjo-Daro, city of the Indus Valley. *Endeavour* 9, 161–169. doi: 10.1016/0160-9327(85)90072-9

Jelks, N. O., Jennings, V., and Rigolon, A. (2021). Green gentrification and health: a scoping review. *Int J Environ Res Public Health* 18:907. doi: 10.3390/ijerph18030907

Jennings, V., Larson, L., and Yun, J. (2016). Advancing sustainability through urban green space: cultural ecosystem services, equity, and social determinants of health. *Int. J. Environ. Res. Public Health* 13:196. doi: 10.3390/ijerph13020196

Johnson, D., and Geisendorf, S. (2022). Valuing ecosystem services of sustainable urban drainage systems: a discrete choice experiment to elicit preferences and willingness to pay. *J. Environ. Manage.* 307:114508. doi: 10.1016/j.jenvman.2022.114508

Jones, L., Anderson, S., Læssøe, J., Banzhaf, E., Jensen, A., Bird, D. N., et al. (2022a). A typology for urban green infrastructure to guide multifunctional planning of nature-based solutions. *Nat. Based Sol.* 2:100041. doi: 10.1016/j.nbsj.2022.100041

Jones, L., Fletcher, D., Fitch, A., Kuyer, J., and Dickie, I. (2024). Economic value of the hot-day cooling provided by urban green and blue space. *Urban Forestry Urban Green.* 93:128212. doi: 10.1016/j.ufug.2024.128212

Jones, L., Reis, S., Hutchins, M., Miller, J., He, B., Seifert-Dähnn, I., et al. (2022b). Airsheds, watersheds and more – the flows that drive intra-extra-urban connections, and their implications for nature-based solutions (NBS). *Nat. Based Sol.* 2:100040. doi: 10.1016/j.nbsj.2022.100040

Jones, L., Vieno, M., Carnell, E., Cryle, P., Holland, M., Nemitz, E., et al. (2019). Urban natural capital accounts: developing a novel approach to quantify air pollution removal by vegetation. *J. Environ. Econ. Policy* 8, 413–428. doi: 10.1080/21606544.2019.1597772

Kaur, S., Mishra, S. K., Goel, V., Kumar, M., Singh, R., Devi, M., et al. (2025). Impact of indoor plant-induced relative humidity on PM concentration in indoor urban environment. *Atmos. Pollut. Res.* 16:102468. doi: 10.1016/j.apr.2025.102468

Khamdamov, S.-J., and Usmanov, A. (2024). Sustainable cities and communities: urban planning and development strategies. *J. Artif. Intell. Digital Econ.* 1, 76–81. doi: 10.61796/jaide.v1i7.794

Konijnendijk, C. C. (2023). Evidence-based guidelines for greener, healthier, more resilient neighbourhoods: introducing the 3–30–300 rule. *J. Forestry Res.* 34, 821–830. doi: 10.1007/s11676-022-01523-z

Kruize, H., van der Vliet, N., Staatsen, B., Bell, R., Chiabai, A., Muiños, G., et al. (2019). Urban green space: creating a triple win for environmental sustainability, health, and health equity through behavior change. *Int. J. Environ. Res. Public Health* 16:4403. doi: 10.3390/ijerph16224403

Kumar, P. (2021). Climate change and cities: challenges ahead. *Front. Sustain. Cities* 3:645613. doi: 10.3389/frsc.2021.645613

Kumar, P., Abhijith, K. V., Barwise, Y. (2019a). *Implementing Green Infrastructure for Air Pollution Abatement: General Recommendations for Management and Plant Species Selection*, p. 9. doi: 10.6084/m9.figshare.8198261.v3

Kumar, P., Corada Perez, K., Debele, S. E., Emygdio, A. P. M., Abhijith, K. V., Hassan, H., et al. (2024a). Air pollution abatement from green-blue-grey infrastructure. *Innov. Geosci.* 2:100100. doi: 10.59717/j.xinn-geo.2024.100100

Kumar, P., Debele, S., Khalili, S., Halios, C. H., Sahani, J., Aghamohammadi, N., et al. (2024b). Urban heat mitigation by green and blue infrastructure: a review of drivers, effectiveness, and future needs. *Innovation* 5:100588. doi: 10.1016/j.xinn.2024.100588

Kumar, P., Druckman, A., Gallagher, J., Gatersleben, B., Allison, S., Eisenman, T. S., et al. L. (2019b). The Nexus between air pollution, green infrastructure and human health. *Environ. Int.* 133:105181. doi: 10.1016/j.envint.2019.105181

Langenheim, N., Sabri, S., Chen, Y., Kesmanis, A., Felson, A., Mueller, A., et al. (2022). Adapting a digital twin to enable real-time water sensitive urban design decision-making. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* 48, 95–100. doi: 10.5194/isprs-archives-XLVIII-4-W4-2022-95-2022

Larsen, R. S. (2018). *Copenhagen: World's First Climate-Resilient Neighborhood*. Global Opportunity Explorer. Avaialble online at: https://goexplorer.org/copenhagen-worlds-first-climate-resilient-neighborhood/ (accessed March 07, 2025).

McHarg, I. L. (1969). *Design with Nature: American Museum of Natural History, 1st edition* (Turtleback), 197.

Meenar, M., Heckert, M., and Adlakha, D. (2022). "Green enough ain't good enough": public perceptions and emotions related to green infrastructure in environmental justice communities. *Int. J. Environ. Res. Public Health* 19:1448. doi: 10.3390/ijerph19031448

Mell, I. C., Henneberry, J., Hehl-Lange, S., and Keskin, B. (2016). To green or not to green: establishing the economic value of green infrastructure investments in the Wicker, Sheffield. *Urban Forestry Urban Green.* 18, 257–267. doi: 10.1016/j.ufug.2016.06.015

Met Office (2021). *Marking Ten Year since Met Office Achievement of the Wildlife Trusts' Biodiversity Benchmark*. Avaialble online at: https://www.metoffice.gov.uk/binaries/content/assets/metofficegovuk/pdf/about-us/who-we-are/sustainability/10_year_anniversary_poster_a2-.pdf (accessed March 13, 2025).

Met Office (2025a). *Biodiversity at the Met Office*. Avaialble online at: https://www.metoffice.gov.uk/about-us/who-we-are/our-values/sustainability/biodiversity#:~:text=To%20date%20over%20500%20different,by%20our%20staff%20each%20year (accessed March 04, 2025).

Met Office (2025b). *UK Climate Projections (UKCP) Data*. Avaialble online at: https://www.metoffice.gov.uk/research/approach/collaboration/ukcp/data/index (accessed March 06, 2025).

Monachese, A., Gómez-Villarino, T., López-Santiago, J., Sanz, E., Almeida-Ñauñay, A., and Zubelzu, S. (2025). Challenges and innovations in urban drainage systems: sustainable drainage systems focus. *Water* 17:76. doi: 10.3390/w17010076

Mueller, W., Milner, J., Loh, M., Vardoulakis, S., and Wilkinson, P. (2022). Exposure to urban greenspace and pathways to respiratory health: an exploratory systematic review. *Sci. Total Environ.* 829:154447. doi: 10.1016/j.scitotenv.2022.154447

Mumford, L. (1938). *The Culture of Cities*. London: Secker and Warburg, 314.

Nash, C., Rumble, H., and Connop, S. (2021). "Stewardship innovation: the forgotten component in maximising the value of urban nature-based solutions," in

Urban Services to Ecosystems, Green Infrastructure Benefits from the Landscape to the Urban Scale, eds. C. Catalano, M. B. Andreucci, R. Guarino, F. Bretzel, M. Leone, and S. Pasta (London: Springer), 165–195. doi: 10.1007/978-3-030-75929-2_9

Natural England (2023). *Green Infrastructure Standards for England - Summary*. Green Infrastructure Framework - Principles and Standards for England. Avaialble online at: https://designatedsites.naturalengland.org.uk/GreenInfrastructure/downloads/Green%20Infrastructure%20Standards%20for%20England%20Summary%20v1.1.pdf (accessed March 03, 2025).

Nature-based Solutions Initiative (2025). *Home*. University of Oxford. Avaialble online at: https://www.naturebasedsolutionsinitiative.org/ (accessed March 11, 2025).

Naumann, S., McKenna, D., Kaphengst, T., Pieterse, M., and Rayment, M. (2011). *Design, Implementation and Cost Elements of Green Infrastructure Projects*. Final report to the European Commission, DG Environment, Contract no. 070307/2010/577182/ETU/F.1, Ecologic Institute and GHK Consulting. Avaialble online at: https://www.ecologic.eu/sites/default/files/project/2014/documents/design-implementation-cost-elements-of-green-infrastructure-projects-2011-naumann_0.pdf (Accessed: 12.03.2025).

Nemitz, E., Vieno, M., Carnell, E., Fitch, A., Steadman, C., Cryle, P., et al. (2020). Potential and limitation of air pollution mitigation by vegetation and uncertainties of deposition-based evaluations. *Philos. Trans. R. Soc. A* 378:20190320. doi: 10.1098/rsta.2019.0320

Nguyen, C. T., and Chidthaisong, A. (2024). Ecosystem services provided by urban green spaces in Bangkok Metropolis: public awareness and planning implications. *Urban Ecosyst.* 27, 855–868. doi: 10.1007/s11252-023-01482-1

Nieuwenhuijsen, M. J., Khreis, H., Triguero-Mas, M., Gascon, M., and Dadvand, P. (2017). Fifty shades of green pathway to healthy urban living. *Epidemiology* 28, 63–71. doi: 10.1097/EDE.0000000000000549

Nobles, E. C., and Moore, K. (2024). Barriers to equitable greening in urban environmental policies. *J. Environ. Policy Plann.* 26, 388–401. doi: 10.1080/1523908X.2024.2364624

Oke, T. R. (1982). The energetic basis of the urban heat Island. *Q. J. R. Meteorol. Soc.* 108, 1–24. doi: 10.1002/qj.49710845502

Owen, D., Fitch, A., Fletcher, D., Knopp, J., Levin, G., Farley, K., et al. (2024). Opportunities and constraints of implementing the 3-30-300 rule for urban greening. *Urban Forestry Urban Green.* 98:128393. doi: 10.1016/j.ufug.2024.128393

Page, J., Kåresdotter, E., Destouni, G., Pan, H., and Kalantari, Z. (2021). A more complete accounting of greenhouse gas emissions and sequestration in urban landscapes. *Anthropocene* 34:100296. doi: 10.1016/j.ancene.2021.100296

Palmer, A. K., Riley, M., Clement, S., Evans, K. L., Jones, L., Brockett, B. F., et al. (2025). In and out of place: diverse experiences and perceived exclusion in UK greenspace settings. *Environ. Plann. E*. doi: 10.1177/25148486251316124

Pearsall, H., Riedman, E., Roman, L. A., Grant, A., Davis, A.-L., Dentice, D., et al. (2024). Barriers to resident participation in tree-planting initiatives across a metropolitan area. *Urban Forestry Urban Green.* 95:128326. doi: 10.1016/j.ufug.2024.128326

Porter, R. (1995). *London, a Social History*. Harvard University Press. Avaialble online at: https://archive.org/search?query=external-identifier%3A%22urn%3Alcp%3Alondonsocialhist00port%3Aepub%3A32936d55-77b8-457d-9de9-b9e1ca956c9a%22 (accessed March 03, 2025).

Prescott, S. L., Millstein, R. A., Katzman, M. A., and Logan, A. C. (2016). Biodiversity, the human microbiome and mental health: moving toward a new clinical ecology for the 21st century? *Int. J. Biodivers.* 1:2718275. doi: 10.1155/2016/2718275

Pretty, J., and Barton, J. (2020). Nature-based interventions and mind–body interventions: saving public health costs whilst increasing life satisfaction and happiness. *Int. J. Environ. Res. Public Health* 17:7769. doi: 10.3390/ijerph17217769

Pretty, J., Griffin, M., Sellens, M., and Pretty, C. (2003). *Green Exercise: Complementary Roles of Nature, Exercise and Diet in Physical and Emotional Well-being and*. Essex: Centre for Environment and Society University of Essex.

Rahmati, Y. (2024). Artificial intelligence for sustainable urban biodiversity: a framework for monitoring and conservation. *arXiv* [Preprint]. doi: 10.48550/arXiv.2501 14766

Redfern, T. W., Macdonald, N., Kjeldsen, T. R., Miller, J. D., and Reynard, N. (2016). Current understanding of hydrological processes on common urban surfaces. *Progress Phys. Geogr.* 40, 699–713. doi: 10.1177/0309133316652819

Richards, D., Polyakov, M., Brandt, A. J., Cavanagh, J., Diprose, G., Milner, G., et al. (2023). Inequity in nature's contributions to people in Otautahi/Christchurch: a low-density post-earthquake city. *Urban Forestry Urban Green.* 86:128044. doi: 10.1016/j.ufug.2023.128044

Riedman, E., Roman, L. A., Pearsall, H., Maslin, M., Ifill, T., and Dentice, D. (2022). Why don't people plant trees? Uncovering barriers to participation in urban tree planting initiatives. *Urban Forestry Urban Green.* 73:127597. doi: 10.1016/j.ufug.2022.127597

Rigolon, A., Browning, M. H. E. M., McAnirlin, O., and Yoon, H. (2021). Green space and health equity: a systematic review on the potential of green space to reduce health disparities. *Int. J. Environ. Res. Public Health* 18:2563. doi: 10.3390/ijerph18052563

Roslund, M. I., Puhakka, R., Nurminen, N., Oikarinen, S., Siter, N., Grönroos, M., et al. (2021). Long-term biodiversity intervention shapes health-associated commensal microbiota among urban day-care children. *Environ. Int.* 157:106811. doi: 10.1016/j.envint.2021.106811

Rowley, S., Leishman, C., Olatunji, O., Zuo, J., and Crowe, A. (2022). *Understanding How Policy Settings Affect Developer Decisions*, AHURI Final Report No. 384. Melbourne: Australian Housing and Urban Research Institute Limited. Avaialble online at: https://www.ahuri.edu.au/research/final-reports/384 (accessed March 14, 2025).

Sachikonye, M. T., Dalu, T., and Gunter, A. (2016). Sustainable livelihood principles and urban greening in informal settlements in practice: a case of Zandspruit informal settlement, South Africa. *Dev. Southern Afr.* 33, 518–531. doi: 10.1080/0376835X.2016.1179101

Saqib, A., Ullua Khan, M. S., and Ahmad Rana, I. (2024). Bridging nature and urbanity through green roof resilience framework (GRF): a thematic review. *Nat. Based Sol.* 6:100182. doi: 10.1016/j.nbsj.2024.100182

SCC (2021). *1.2 Million New Trees for Surrey*. Surrey County Council. Avaialble online at: https://www.surreysays.co.uk/environment-and-infrastructure/1-2-million-new-trees-for-surrey/ (accessed March 13, 2025).

Schönbrunn (2025). *The History of the Park*. Schönbrunn Palace. Avaialble online at: https://www.schoenbrunn.at/en/about-schoenbrunn/gardens/history (accessed March 08, 2025).

Self, A. (2014). *The Birds of London*. London: Bloomsbury Publishing. ISBN 9781408194058. Avaialble online at: https://www.bloomsbury.com/uk/birds-of-london-9781408194058/ (accessed March 12, 2025).

Shaamala, A., Yigitcanlar, T., Nili, A., and Nyandega, D. (2024). Algorithmic green infrastructure optimisation: review of artificial intelligence driven approaches for tackling climate change. *Sustain. Cities Soc.* 101:105182. doi: 10.1016/j.scs.2024.105182

Sharma, A. (2006). *Water Harvesting Context in the Indian Subcontinent*. UNESCO G-WADI meeting on water harvesting Aleppo Syria 20-22, November 2006, 63–70.

Sondou, T., Dotsu, M. Y., Anoumou, K. R., Samon, S. P., Chenal, J., and Aholou, C. C. (2025). Urban planning through participatory democracy: analysis of citizen participation in urban planning in Ho (Ghana) and Kpalimé (Togo). *Sustainability* 17:1161. doi: 10.3390/su17031161

Stessens, P., Canters, F., and Khan, A. Z. (2021). Exploring options for public green space development: research by design and GIS-based scenario modelling. *Sustainability* 13:8213. doi: 10.3390/su13158213

Tefera, Y., Soebarto, V., Bishop, C., Kandulu, J., and Williams, C. (2023). A scoping review of urban planning decision support tools and processes that account for the health, environment, and economic benefits of trees and greenspace. *Int. J. Environ. Res. Public Health* 21:48. doi: 10.3390/ijerph21010048

The City of Copenhagen (2012). *The City of Copenhagen Cloudburst Management Plan 2012*. Avaialble online at: https://climate-adapt.eea.europa.eu/en/metadata/case-studies/the-economics-of-managing-heavy-rains-and-stormwater-in-copenhagen-2013-the-cloudburst-management-plan/cloudburst_management_plan_2012.pdf/%40%40download/file (accessed March 04, 2025).

The Wildlife Trusts (2025). *Biodiversity Benchmark*. The Wildlife Trusts. Avaialble online at: https://www.wildlifetrusts.org/partnerships/working-businesses/biodiversity-benchmark (accessed March 12, 2025).

UNESCO World Heritage Centre (2025). *The Persian Garden*. Avaialble online at: https://whc.unesco.org/en/list/1372/ (accessed March 11, 2025).

Urban Nature Atlas (2025). *Welcome to the Atlas*. Avaialble online at: https://una.city/ (accessed March 11, 2025).

Verma, G., Mondal, K., Islam, M., and Gupta, A. (2024). Recent advances in advanced micro and nanomanufacturing for wastewater purification. *ACS Appl. Eng. Mater.* 2, 262–285. doi: 10.1021/acsaenm.3c00711

Vesuviano, G., Fitch, A., Owen, D., Fletcher, D., and Jones, L. (2025). How well does the 3-30-300 rule mitigate urban flooding? *Urban Forestry Urban Green.* 104:128661 doi: 10.1016/j.ufug.2024.128661

Wang, F., Harindintwali, J. D., Wei, K., Shan, Y., Mi, Z., Costello, M. J., et al. (2023). Climate change: strategies for mitigation and adaptation. *Innov. Geosci.* 1:100015. doi: 10.59717/j.xinn-geo.2023.100015

Well, F., and Ludwig, F. (2022). Integrated planning and implementation of a blue-green architecture project by applying a design-build teaching approach. *Land* 11:762. doi: 10.3390/land11050762

Welsh Government (2018). *Statutory National Standards for Sustainable Drainage Systems: Designing, Constructing, Operating and Maintaining Surface Water Drainage Systems*, 63. Avaialble online at: https://www.gov.wales/sites/default/files/publications/2019-06/statutory-national-standards-for-sustainable-drainage-systems.pdf (accessed March 11, 2025).

Williams, C., Byrne, C., Evenden, S., Soebarto, V., Caddy-Retalic, S., Williams, C., et al. (2025). Urban green space provision: the case for policy-based solutions to support human health. *Med. J. Aust.* 222, 52569-1–52569-4. doi: 10.5694/mja2.52569

World Bank (2015). *Investing in Urban Resilience: Protecting and Promoting Development in a Changing World*, 120. Avaialble online at: https://documents1.worldbank.org/curated/en/739421477305141142/pdf/109431-WP-P158937-PUBLIC-ABSTRACT-SENT-INVESTINGINURBANRESILIENCEProtectingandPromotingDevelopmentinaChangingWorld.pdf (accessed March 12, 2025)

Xu, C., Liu, Z., Chen, Z., Zhu, Y., Yin, D., Leng, L., et al. (2021). Environmental and economic benefit comparison between coupled grey-green infrastructure system and traditional grey one through a life cycle perspective. *Resourc. Conserv. Recycl.* 174:105804. doi: 10.1016/j.resconrec.2021.105804

Xu, H., Omitaomu, F., Sabri, S., Zlatanova, S., Li, X., Song, Y., et al. (2024). Leveraging generative AI for urban digital twins: a scoping review on the autonomous generation of urban data, scenarios, designs, and 3D city models for smart city advancement. *Urban Inf.* 3:29. doi: 10.1007/s44212-024-00060-w

Yang, M., Ren, C., Wang, H., Wang, J., Feng, Z., Kumar, P., et al. (2024). Mitigating urban heat island through neighboring rural land cover. *Nat. Cities* 1, 533–532. doi: 10.1038/s44284-024-00091-z

Zero Carbon Guildford (2025). *Urban Greening*. Zero Carbon Guildford. Avaialble online at: https://www.zerocarbonguildford.org/urban-greening (accessed March 08, 2025).

Zhang, H., Zhan, Y., and Chen, K. (2024). Do education, urbanization, and green growth promote life expectancy? *Front. Public Health* 12:1517716. doi: 10.3389/fpubh.2024.1517716

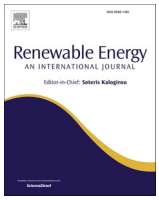
Renewable Energy 247 (2025) 123026



Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene



Vortex flow plasma reforming for hydrogen production from atomized water-methanol mixture and parameter optimization using RSM and ANN-GA

Neeraj Budhraja *, Amit Pal, R.S. Mishra

Department of Mechanical Engineering, Delhi Technological University, Delhi, 110042, India

ARTICLE INFO

Keywords:
Plasma reforming
Feed flow rate
Hydrogen production rate
Response surface methodology
Artificial neural network-genetic algorithm

ABSTRACT

Hydrogen energy is the cleanest energy. It has many applications in different sectors, such as power generation, industrial, and transportation. The conventional H_2 production methods are energy-consuming and generate a huge amount of pollution. Hence, a non-polluting method for H_2 production is necessary to utilize the benefit of cleaner fuel. The current work used a vortex flow plasma reforming process to generate H_2 from the methanol-water mixture. The process involved plasmolysis of methanol-water mixture to produce H_2. Additionally, the process parameters like feed flow rate (0.5–4.5 LPM), methanol concentration (5–35 %) and voltage (4–8 V) were optimized using RSM and ANN-GA models, and the influencing parameters were determined using the ANOVA test. The results showed an optimal hydrogen production rate of 229.2 mol/day (for 0.3 MHz), 265.0 mol/day (for 1.7 MHz) and 285.8 mol/day (for 2.4 MHz) for RSM at 4.5 LPM and 35 % methanol concentration while 267.0 mol/day for ANN-GA at 4.5 LPM, 33.8 % methanol concentration and 7.4 kV input voltage. The optimized parameters were used in the experimental setup, and an error of 4–7 % was observed for the RSM model and 2.5 % for the ANN-GA model. Therefore, the ANN-GA model showed better reliability than the RSM model. Further, the results may be used for commercial-scale H_2 production for various applications.

1. Introduction

Hydrogen is a clean fuel with many applications in different sectors, such as power generation, industrial, and transportation. However, hydrogen production requires enormous energy from conventional plants that pollute the environment [1]. Currently, many processes/methods are available for hydrogen production; however, the maximum hydrogen generation is through steam reforming (about 95 %), followed by electrolysis (about 4 %). The steam reforming is an energy-consuming and pollution-generating process [2,3]. And the raw materials are mainly natural gas and coal gas, which further degrades the environment [4,5]. Another hydrogen-producing method is plasma reforming, which can be a game changer in low energy consumption, eco-friendliness, and less polluting raw materials for hydrogen production [6–8]. The plasma generation requires electricity, which may be generated from renewable energy sources, further reducing dependence on conventional fuels during production [9] (see Table 10).

Plasma reforming is a process where a high-voltage plasma arc creates a very high-temperature zone (also called plasma arc zone), say about 20000 °C, with enough energy to break down hydrogen bonds from its compounds to produce hydrogen ions [10]. These hydrogen ions recombine to form hydrogen molecules (or gas) that can be separated and used for various applications [11,12]. *Pang* et al. [13] evaluated the functioning of the allo-thermal steam gasification and the non-thermal steam gasification of the wood particles on the hydrogen generation rate using the electrically heated drop tube reactor. Syngas composition produced using plasma-assisted gasification contained H_2 (38.5 %), CO (38.4 %), CO_2 (12.7 %), and CH_4 (10.3 %) at 760 °C, respectively. A 20 % increase was noticed in syngas production by non-thermal plasma systems compared to allo-thermal gasification. *Bulychev* et al. [14] analyzed the acoustic plasma method in liquid decomposition at low temperatures. The H_2 production rate attained was 2 L/min with a 100 mL capacity reaction chamber, and the feed rate of about 60–70 % of the energy efficiency factor was achieved depending upon the feed mixture composition.

Methanol and water are prominent sources of hydrogen production using different production methods. The hydrogen content in methanol and water is more than 10 %, while the low energy requirement for

* Corresponding author.
E-mail address: neeraj_budhraja@yahoo.com (N. Budhraja).

https://doi.org/10.1016/j.renene.2025.123026
Received 22 July 2024; Received in revised form 6 February 2025; Accepted 2 April 2025
Available online 3 April 2025
0960-1481/© 2025 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

hydrogen extraction gives them an advantage over other sources. Typically, the methanol has about 12.5 % of hydrogen content. The benefits of producing hydrogen from methanol are its less toxic nature, the fact that it is liquid at room temperature, and hydrogen generation at a comparatively lower temperature. These advantages developed an interest in methanol as a greener hydrogen production source. *Budhraja* et al. [15] simulated the hydrogen production process from methanol via a reforming process. The result showed a 0.57 hydrogen mole fraction with a more than 90 % methanol conversion percentage.

Plasma reforming has different operating parameters that influence the hydrogen production in a system [16,17]. Hence, understanding parameter performance in hydrogen production is crucial to enhance hydrogen production [18]. The influence of discharge frequency, input power, and liquid feed rate on the plasma-steam reforming of methanol and synergistic plasma-catalyst systems described a rise in the frequency accelerating the methanol conversion initially and then reduced. In contrast, 17.89 % and 21.86 % more methanol conversion and H_2 yield were obtained with activated catalyst than without catalyst [19]. *Huang* et al. [20] modified the plasma reforming system by adding a heater and varying the amplification ratio for hydrogen generation from the methanol-water mixture. Effects of increasing operating voltage and temperature showed approximately 50 % upsurge in hydrogen generation. In contrast, the influence of discharge power, type of carrier gas, and the moisture content of the raw material in the non-thermal arc plasma reactor showed that the gas yield enhanced with the rise in discharge power and higher moisture content; however, air as carrier gas showed the best gas yield and H_2/CO ratio against N_2 and Ar.

Once the influencing parameters were determined, a statistical relationship was developed between different parameters and response values. This relationship generates predicted values for each parameter input interaction [21–23]. Hence, the optimal output (or hydrogen production) was computed through the developed model. The researchers used different statistical model tools like response surface methodology (RSM) and artificial neural network (ANN) for optimization. *Shiraishi* et al. [24] studied the performance of the Ni electrode and Ni plate on the in-liquid plasma breakdown of methanol for hydrogen generation. The in-liquid plasma method achieved a hydrogen yield of 67 % for methanol conversion, and the hydrogen generation rate increased by 2.3–3.5 times with the Ni porous plate placed above the electrode. *Lian* et al. [25] examined the performance of O_2 to C ratios, steam/carbon (S/C) ratios, and specific energy input (SEI) on the oxidative methanol-reforming for onboard hydrogen generation using a gliding arc plasma reactor. Methanol conversion was accelerated with the rise of the O_2 to C ratio, although the lesser energy efficiency was recorded with the upsurge in either the S/C ratio or SEI. In contrast, the methanol conversion of 88 % and energy efficiency of 74 % was achieved by an O_2/C ratio of 0.30, an S/C ratio of 0.5, and an SEI of 24 kJ/mol, respectively.

The current work developed a vortex flow plasma reformer for hydrogen production from the water-methanol mixture. No study was observed for vortex flow plasma reforming in open literature. The plasma vortex created more spots for plasma and feed interaction, resulting in higher hydrogen production. The various input parameters were computed, and their influence on H_2 production was calculated. An RSM model was generated using parameters like feed flow rate, methanol concentration, and input voltage at different levels. The inputs were analyzed for optimal values, and their influences were recorded using Analysis of Variance (ANOVA). An ANN-GA model was also developed using the same parameters, and the data was analyzed for a better optimal solution. The results from RSM, ANN-GA, and experimental data were compared, and it was observed that ANN-GA has a better correlation between parameters and response than the RSM model. This way, the input parameters were optimized for better hydrogen production parameters and enhanced yield. Therefore, the output data may be used to develop a commercial-scale hydrogen production plant.

2. Material and method

The raw materials and the production process of hydrogen were proposed in this section. The mathematical modeling and parameter optimization were performed using response surface methodology (RSM) and artificial neural network-genetic algorithm (ANN-GA) models.

2.1. Raw materials

The commercial grade methanol (ONGC/LPG/09/2015), which was 99.9 % pure and distilled water (de-ionized/de-mineralized), was procured to prepare a water-methanol mixture at different methanol concentrations (5 %, 20 % and 35 %) to be fed into the reactor. All the high-grade chemicals did not require further processing to prepare samples for plasma reformer feed. The samples contained a water-methanol mixture in different compositions (i.e., Sample 1: 5 % methanol + 95 % water; Sample 2: 20 % methanol + 80 % water; and Sample 3: 35 % methanol + 65 % water, respectively).

2.2. Feed atomization

The water-methanol mixture cannot be fed directly to the plasma reformer/reactor; therefore, the feed atomization was carried out. Three different ultrasonic transducers (or atomizers) with frequencies of 0.3 MHz, 1.7 MHz, and 2.4 MHz were installed for feed atomization. The ultrasonic transducers and their circuit are shown in Fig. 1[26]. The working principle of the ultrasonic transducer depends on the sound frequency at the ultrasonic level generated by the metal diaphragm connected with the piezo-ceramic element. The ultrasonic sound waves vibrate the metal diaphragm to convert liquid feed into minute droplets with a diameter within a few nanometres (nm) range. The atomized feed was then fed into the plasma reformer for further processing. The specifications of the atomizers are depicted in Table 1.

2.3. Plasma reforming process

Water-methanol reforming was carried out in the plasma reformer designed and developed in the laboratory (Fig. 2). The plasma reformer consists of two electrodes (one cylindrical copper-based alloy and another neodymium-based ring-type magnet) placed in a cylindrical tube with an inner diameter of 20 mm [26]. A high-voltage generator was designed with a flyback transformer that generated a potential difference between 1.0 kV and 10 kV at the output terminals. The atomized feed was fed into the reformer and passed through the two electrodes. The plasma arc jumps from the Cu electrode to the magnet electrode when the distance between the two electrodes ranges from 7 mm to 13 mm. The plasma arc created a high-temperature zone that breaks down the hydrogen bond of the feed. Hence, water and methanol

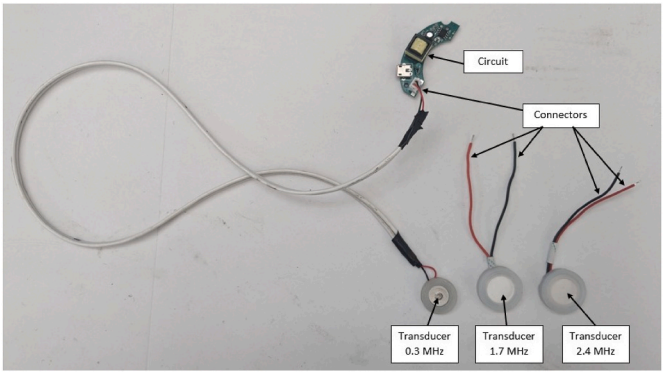


Fig. 1. Various ultrasonic transducers with circuit.

Table 1
Specifications of each atomizer.

Details	0.3 MHz	1.7 MHz	2.4 MHz
Model	SKU: 835845	SKU: 790464	SKU: 815786
Resonant Frequency (MHz)	0.3 (± 3 kHz)	1.7 (± 7 kHz)	2.4 (± 7 kHz)
Diameter (mm)	20	20	20
Power Rating (W)	2.5	2.5	2.5
Electrostatic capacity (pF)	3000	3000	2200

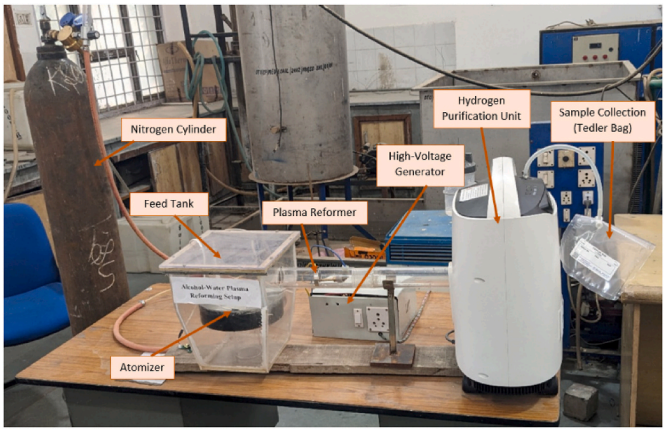


Fig. 2. Experimental setup for plasma reforming.

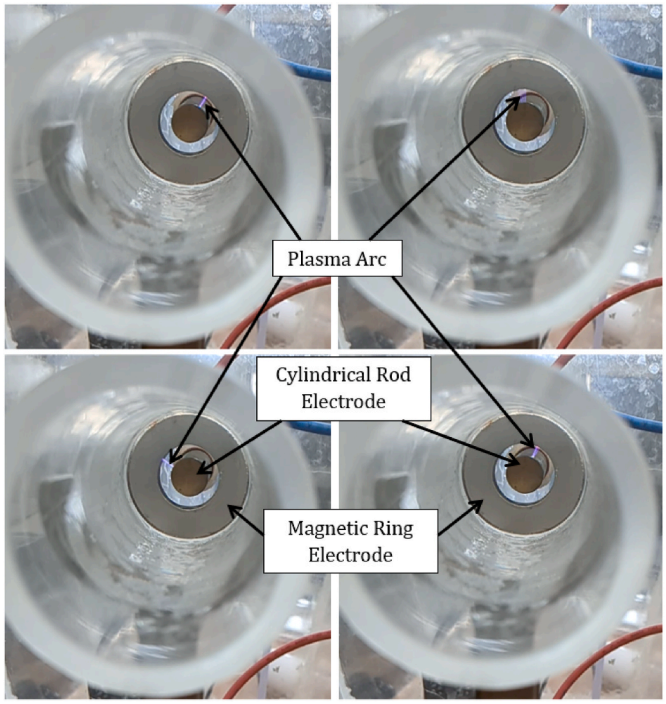


Fig. 3. Plasma arc vortex formation inside plasma reformer.

molecules split into species like H, CO, and O ions. These ions recombine to form compounds like H_2, CO_2, and CO.

The methanol-reforming process involves reactions (1), (2), and (3) as stated below [15]:

$$CH_3OH \rightarrow HCHO + H_2 \quad (1)$$

$$HCHO + H_2O \rightarrow HCOOH + H_2 \quad (2)$$

$$HCOOH \rightarrow CO_2 + H_2 \quad (3)$$

Whereas the other reactions involved in the methanol reforming process include reaction (4), (5), (6) where reaction (6) is water gas shift reaction [27].

$$CH_3OH + H_2O \leftrightarrow CO_2 + 3H_2 \quad (4)$$

$$CH_3OH \leftrightarrow CO + 2H_2 \quad (5)$$

$$CO + H_2O \leftrightarrow CO_2 + H_2 \quad (6)$$

The H_2 molecules were separated and stored in hydrogen storage bags, and tedlar bags of 1 L were used to collect samples to test the product's purity. The plasma arc vortex was formed within the plasma reformer is presented in Fig. 3[26].

2.4. Hydrogen purification

Hydrogen purification becomes very necessary to obtain pure hydrogen for applications such as fuel for fuel cells and other industrial applications. The gaseous mixture coming out of the plasma reformer consists of H_2, N_2, CO, CO_2, etc. The hydrogen purification was performed by passing a pressurized gaseous mixture through a zeolite sieve. The void diameter of the zeolite sieve was slightly more than the kinematic diameter of hydrogen molecules; thus, the hydrogen molecules passed through the sieve, whereas other gaseous molecules, being larger than the void diameter, got stuck. Once the first sieve got choked from larger gaseous molecules, the other sieve began working. After purification, a small amount of hydrogen gas passes through the first sieve in the reverse direction. The choked voids get unblocked and become ready for operation after the second sieve gets choked. Therefore, the zeolite sieve's automatic ON/OFF switching performs a continuous hydrogen purification. The specifications of the hydrogen purification unit used are mentioned in Table 2.

The purified hydrogen was taken for sample testing in tedlar bags for qualitative gas chromatography analysis. The specifications of the gas chromatography used are depicted in Table 3.

2.5. Hydrogen production rate

The amount of hydrogen gas produced during the plasma reforming process was measured in hydrogen production rate, R_{H_2} (mol/day) using Eq. (7),

$$R_{H_2} = \frac{W_{H_2}}{T \times 24} \quad (7)$$

Where W_{H_2} represents the amount of hydrogen produced in mol, and T represents the total time in hours.

2.6. Response surface methodology

Response Surface Methodology (RSM) provides a systematic method for conducting experiments and analyzing data. It allows for the creation of empirical models that establish the relationship between input parameters (such as feed flow rate, methanol concentration, and voltage) and the output response (hydrogen production rate). Response surface

Table 2
Specifications of hydrogen purification unit.

Make	Yuwell
Model	8F-5A/8F-5AW
Power	AC 220V @50Hz, 400 VA
Flow range	0.5–5.0 LPM
Sieve material	Zeolite
Output pressure	40–70 kPa
Net Weight	15.5 kg

Table 3
Specifications of the gas chromatography.

Make	Agilent
Model	8860 GC System
Detector	Thermal conductivity detector (TCD)
Column	HP-5
Carrier gas	Nitrogen
Oven temperature	50 °C (10 min) with 10 °C/min to 120 °C (5min.)
TCD temperature	180
Sample	0.1 mL

methodology (RSM) offers valuable insights into the impacts of individual factors and their interactions on the intended outcomes. Therefore, the experimental variables in the hydrogen production process were optimized using the RSM. As depicted in Table 5, the central composite design (CCD) generated a 20-run design of experiments (DoE) with three input variables at three levels. The RSM model generated empirical relationships between input and output variables to investigate the impacts of the various parameters at different levels. The various input variables involved in generating DoE for CCD were feed flow rate (FFR), methanol concentration (MeOH), and input voltage (Voltage).

A CCD design consisting of 6-central, 6-axial, and 8-factorial points was used to generate a quadratic Equation (8) that represented Y as the response variable (Hydrogen production rate); P, Q, and R as input variables (depicted in Table 2); β as intercept; β_1, β_2 and β_3 as first-order coefficients; β_4, β_5 and β_6 as interactive coefficients; β_7, β_8 and β_9 as a second-order coefficient and β_{10} as associated error, respectively.

$$Y = \beta_0 + \beta_1(P) + \beta_2(Q) + \beta_3(R) + \beta_4(PQ) + \beta_5(QR) + \beta_6(RQ) + \beta_7(P^2) + \beta_8(Q^2) + \beta_9(R^2) + \beta_{10} \quad (8)$$

In the earlier study, *Budhraja* et al. [27] used RSM-CCD to optimize hydrogen production from a water-methanol mixture via a reforming process. The optimized hydrogen selectivity of 84.81 % was achieved against the maximum simulated hydrogen selectivity of 82.76 %. Similarly, in another study, *Budhraja* et al. [28,29] optimized biohydrogen production using an anaerobic digester. The optimized hydrogen yield of 25.27 % was found against the maximum yield of 24.6 %. In another study, *Budhraja* et al. [30] optimized the biodiesel yield using the RSM model and observed a 92 % optimized yield against 89 % of the maximum experimental biodiesel yield.

2.7. Artificial neural network-genetic algorithm

Another tool used to optimize process parameters was the artificial neural network-genetic algorithm (ANN-GA) model. An ANN-GA model is a computational model that performs data interpretation and predictions based on the human brain and natural evolution. The correlation coefficient (R) for the hydrogen production rate was computed.

3. Results and discussion

This section describes the results obtained during experimentation and analyses the results through discussions. The effects of various process parameters, namely input voltage, methanol concentration, and feed flow rate, were studied. The three ultrasonic transducers with ultrasonic frequencies of 0.3 MHz, 1.7 MHz, and 2.4 MHz were used to atomize the water-methanol mixture, and the effects with different parameters were observed. The results obtained were discussed separately for each parameter, taking the other two parameters constant. The regression analysis and the interactive effects of the various parameters were also studied. The optimized results were obtained and compared with the experimental results obtained.

Table 4
Operating parameters for the RSM model.

Parameter	Unit	Levels		
		−1	0	1
P-Feed Flow Rate	LPM	0.5	2.5	4.5
Q-Methanol concentration	%	5	20	35
R-Voltage	kV	4	6	8

3.1. Empirical design & regression analysis

The three process parameters involved in the optimization process are feed flow rate (FFR), methanol concentration (MeOH), and input voltage (Voltage). Table 2 shows that a DoE was generated in CCD for three factors at three levels. A 20-run hydrogen production rate (HPR) design matrix was developed using three different ultrasonic frequency transducers (0.3 MHz, 1.7 MHz, and 2.4 MHz). The residual (or percentage error) was computed using actual and predicted values received through analysis and depicted in Table 3.

Experimental randomizations reduced the response variations. The predicted values were produced through a response model for each ultrasonic frequency transducer {Equation (9), Equation (10), and Equation (11)} with a 95 % confidence level. Here, the terms FFR, MeOH, and Voltage were linear; $(FFR)^2$, $(MeOH)^2$, and $(Voltage)^2$ were quadratic terms; and (FFR x MeOH), (MeOH x Voltage) and (Voltage x FFR) were interactive terms for hydrogen production rate at relative frequencies.

$$HPR\ (0.3\ MHz) = -10.661 + 61.049\ FFR - 0.006\ MeOH - 1.950\ Voltage + 0.595\ FFR{*}MeOH + 0.753\ FFR{*}Voltage + 0.070\ MeOH{*}Voltage - 7.549\ FFR^2 - 0.007\ MeOH^2 + 0.051\ Voltage2 \quad (9)$$

$$HPR\ (1.7\ MHz) = -11.837 + 68.784\ FFR - 0.937\ MeOH + 2.013\ Voltage + 0.492\ FFR{*}MeOH + 0.697\ FFR{*}Voltage + 0.032\ MeOH{*}Voltage - 6.694\ FFR^2 + 0.021\ MeOH^2 - 0.269\ Voltage2 \quad (10)$$

$$HPR\ (2.4\ MHz) = -21.500 + 81.859\ FFR - 0.658\ MeOH + 4.952\ Voltage + 0.345\ FFR{*}MeOH + 0.781\ FFR{*}Voltage + 0.011\ MeOH{*}Voltage - 7.711\ FFR^2 + 0.021\ MeOH^2 - 0.499\ Voltage2 \quad (11)$$

The F-values and p-values of the parameters were calculated from analysis of variance (ANOVA) tests. The proportion between the mean square errors was defined as the F-value. A higher F-value determined the significance of a parameter. Similarly, the variation between the group means described the p-value. The relevance of a parameter was described by the p-value <0.001, respectively. Here, the ANOVA tables were developed for each ultrasonic frequency (Table 4, Tables 5, and Table 6), and each parameter's significance was determined for the hydrogen production rate. Table 4, Tables 5, and Table 6 concluded that the feed flow rate and the methanol concentration were the most significant parameters through F-values and p-values calculations for each ultrasonic frequency transducer.

The coefficient of regression (R^2) values of each ultrasonic frequency transducer were also calculated for the hydrogen production rate. It determined the model's adequacy and showed the competence of the outcomes with the statistically predicted values. The values closer to unity showed better predictions for the model. Through Table 7, it was concluded that the R^2 values of the outputs from each ultrasonic frequency transducer were closer to 1 (unity). It showed a good correlation developed among the experimental and predicted values for hydrogen production rate.

Table 5
Matrix design of 20-runs (Experimental).

Run	Parameters			Hydrogen Production Rate (mol/day)								
				0.3 MHz			1.7 MHz			2.4 MHz		
	FFR (LPM)	MeOH (%)	Voltage (kV)	Actual Value	Predicted Value	Residual	Actual Value	Predicted Value	Residual	Actual Value	Predicted Value	Residual
1	0.5	35	4	26.7	23.6	3.10	34.2	32.6	1.61	39.1	41.0	−1.93
2	2.5	35	6	147.2	153.9	−6.67	166.8	174.4	−7.58	196.3	193.5	2.83
3	2.5	20	6	132.3	131.4	0.94	148.1	149.4	−1.31	169.8	172.2	−2.38
4	2.5	20	6	132.3	131.4	0.94	148.1	149.4	−1.31	169.8	172.2	−2.38
5	4.5	35	8	231.1	230.2	0.92	267.0	266.0	0.96	282.2	286.5	−4.28
6	0.5	5	4	15.0	15.2	−0.18	21.8	23.7	−1.94	31.7	29.2	2.50
7	2.5	5	6	109.3	105.6	3.75	145.6	134.1	11.52	164.6	160.3	4.31
8	2.5	20	6	132.3	131.4	0.94	148.1	149.4	−1.31	169.8	172.2	−2.38
9	2.5	20	6	132.3	131.4	0.94	148.1	149.4	−1.31	169.8	172.2	−2.38
10	0.5	35	8	28.7	29.6	−0.92	35.1	33.6	1.50	40.2	40.0	0.23
11	2.5	20	8	140.7	135.5	5.23	155.3	150.7	4.64	181.3	172.5	8.85
12	0.5	20	6	22.6	21.7	0.87	27.3	24.0	3.34	35.7	31.5	4.15
13	2.5	20	4	119.5	127.6	−8.15	145.3	146.0	−0.70	166.2	167.9	−1.71
14	2.5	20	6	132.3	131.4	0.94	148.1	149.4	−1.31	169.8	172.2	−2.38
15	4.5	5	8	139.6	142.0	−2.36	191.7	194.3	−2.59	232.1	231.9	0.15
16	0.5	5	8	9.9	12.7	−2.85	16.4	20.9	−4.50	21.9	26.8	−4.94
17	2.5	20	6	132.5	131.4	1.14	148.1	149.4	−1.31	169.8	172.2	−2.38
18	4.5	20	6	176.8	180.6	−3.79	221.9	221.3	0.60	254.1	251.1	2.99
19	4.5	35	4	215.7	212.1	3.59	257.4	253.9	3.52	278.2	275.0	3.16
20	4.5	5	4	134.0	132.4	1.66	183.5	186.0	−2.48	219.8	221.8	−2.01

Table 6
ANOVA table of 0.3 MHz ultrasonic frequency for HPR.

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	76566.51	9	8507.39	397.76	<0.0001	significant
A-Feed Flow Rate	63091.25	1	63091.25	2949.84	<0.0001	
B-Methanol concentration	5837.06	1	5837.06	272.91	<0.0001	
C-Voltage	152.88	1	152.88	7.15	0.0233	
AB	2545.41	1	2545.41	119.01	<0.0001	
AC	72.60	1	72.60	3.39	0.0952	
BC	35.70	1	35.70	1.67	0.2254	
A^2	2507.36	1	2507.36	117.23	<0.0001	
B^2	7.45	1	7.45	0.3481	0.5683	
C^2	0.1151	1	0.1151	0.0054	0.9430	
Residual	213.88	10	21.39			
Lack of Fit	213.85	5	42.77	6415.40	<0.0001	significant
Pure Error	0.0333	5	0.0067			
Cor Total	76780.39	19				

Table 7
ANOVA table of 1.7 MHz ultrasonic frequency for HPR.

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	1.064E+05	9	11817.72	408.97	<0.0001	significant
A-Feed Flow Rate	97357.69	1	97357.69	3369.18	<0.0001	
B-Methanol concentration	4060.23	1	4060.23	140.51	<0.0001	
C-Voltage	54.29	1	54.29	1.88	0.2005	
AB	1743.45	1	1743.45	60.33	<0.0001	
AC	62.16	1	62.16	2.15	0.1732	
BC	7.41	1	7.41	0.2565	0.6235	
A^2	1971.81	1	1971.81	68.24	<0.0001	
B^2	63.96	1	63.96	2.21	0.1677	
C^2	3.19	1	3.19	0.1104	0.7465	
Residual	288.97	10	28.90			
Lack of Fit	288.97	5	57.79			
Pure Error	0.0000	5	0.0000			
Cor Total	1.066E+05	19				

3.2. Interactive effects of parameter

The analysis of the interactive effects of the parameters generated a relationship between two parameters taken together to the response value. The relationship was developed using model equations for each response value. A 3D surface contour graph was developed with two parameters, and a response value described the interactive effects of the parameters considered during the analysis.

The interactive effects of different parameters with a 0.3 MHz ultrasonic frequency transducer for HPR are shown in Fig. 4. In Fig. 4 (a), the FFR and MeOH were taken together. In contrast, in Fig. 4 (b), the MeOH and Voltage and in Fig. 4 (c), the Voltage and FFR were taken

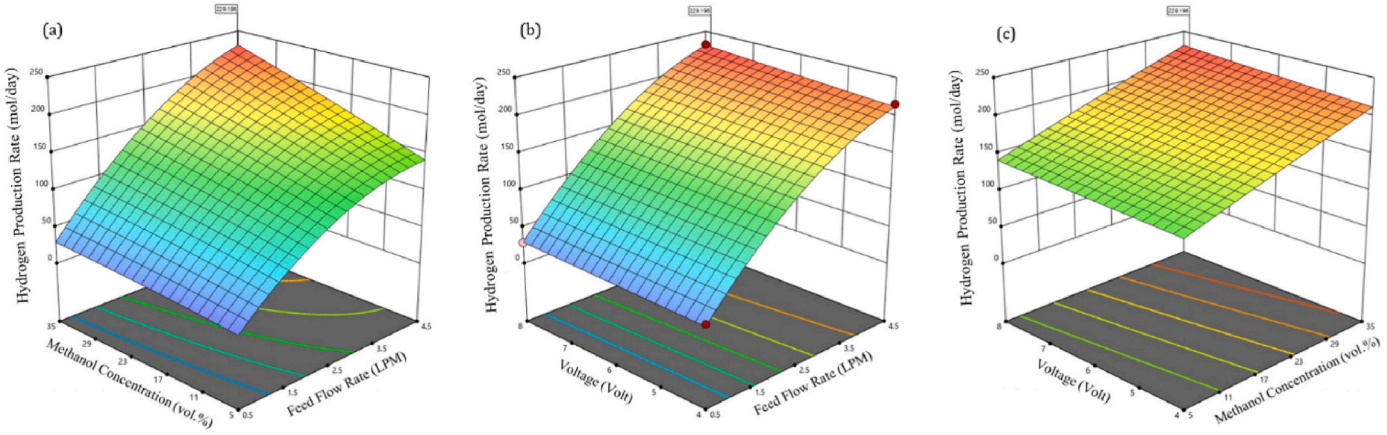


Fig. 4. 3D Surface contour curve of hydrogen production rate for 0.3 MHz transducer frequency (a) MeOH vs. FFR; (b) Voltage vs FFR; and (c) Voltage vs MeOH.

together. The results showed that the FFR of 4.5 LPM and MeOH of 35 % vol. generated the highest HPR. Fig. 4 (b) showed the highest HPR at 4.5 LPM and all input voltages. Similarly, Fig. 4 (c) showed peak HPR at higher MeOH at all voltage ranges. Here, the peak HPR was 229.2 mol/day achieved with optimized parameters.

The interactive effects of different parameters on HPR using a 1.7 MHz ultrasonic frequency transducer are shown in Fig. 5. Fig. 5 (a) generated interaction between FFR and MeOH; Fig. 5 (b) generated interaction between MeOH and Voltage; and Fig. 5 (c) generated interaction between Voltage and FFR. The improvement in the HPR from the 0.3 MHz ultrasonic frequency transducer was seen, and the peak was attained at 265.0 mol/day. Here, Fig. 5 (a) showed a peak value of HPR at FFR of 4.5 LPM and MeOH of 35 % vol., whereas Fig. 5 (b) showed the highest value at 4.5 LPM and all voltage ranges. Similarly, the peak value in Fig. 5 (c) was at higher methanol concentrations and all voltage ranges.

The interactive effects of different parameters with a 2.4 MHz ultrasonic frequency transducer for HPR are presented in Fig. 6. In Fig. 6 (a), the FFR and MeOH were taken together, while in Fig. 6 (b), the MeOH and Voltage were brought together, whereas in Fig. 6 (c), the Voltage and FFR were taken together. The interactions of parameters showed a maximum HPR of 285.8 mol/day. Fig. 6 (a) showed a peak value at 4.5 LPM and 35 % vol. methanol concentration. Fig. 6 (b) showed maximum HPR at 4.5 LPM and all voltage ranges. Similarly, Fig. 6 (c) showed the peak value of HPR for all voltage ranges and at 35 % vol. methanol concentration. The discussion showed that the higher ultrasonic frequency contributed to a higher hydrogen production rate. Also, the higher methanol concentrations and feed flow rates contributed to enhanced hydrogen production rate, which showed that hydrogen generation was mainly due to methanol conversion.

Table 8 shows the percentage error for hydrogen production rate at optimal parameters of different transducer frequencies against experimental results. It was observed that a maximum error of 6.7 % and a minimum error of 4.1 % were computed when optimal parameter values from the RSM model were used in the experimentation setup. Therefore, an error of the 4–7 % order is within the tolerable limits (see Table 9).

3.3. Artificial neural network (ANN)

The correlation coefficient (R) value for the hydrogen production rate reached unity, showing a reliable regression model (see Fig. 7). The high R values observed across all networks indicate that the data was effectively trained using the DOE, reinforcing the reliability and robustness of the regression model. Here, the ultrasonic transducer frequencies (0.3 MHz, 1.7 MHz and 2.4 MHz) were also considered as the optimizing parameter along with feed flow rate (0.5–4.5 LPM), methanol concentration (5–35 %) and voltage (4–8 V), respectively. It will help find the optimized ultrasonic frequency for atomization that will enhance the hydrogen production rate. The optimized value of HPR achieved was 267.0 mol/day at 4.5 LPM feed flow rate, 33.8 % vol. methanol concentration, and 7.4 kV input voltage for 1.7 MHz transducer frequency. Therefore, the 1.7 MHz ultrasonic transducer frequency would be selected for optimized hydrogen production rate. In contrast, the experimental results showed 260.4 mol/day HPR at the above parameters. Hence, an error of about 2.5 % was observed for the ANN-GA model against experimental results.

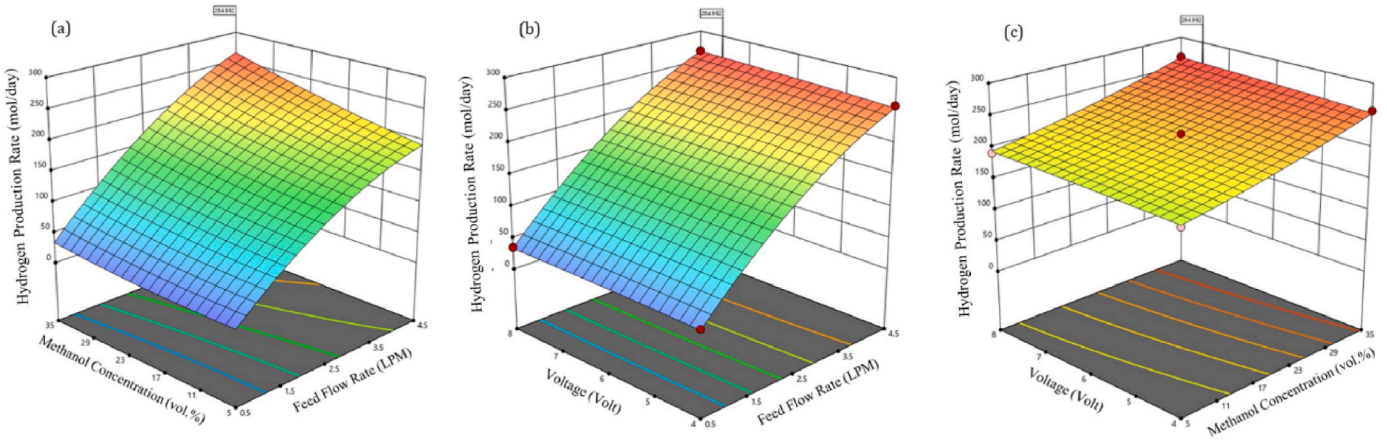


Fig. 5. 3D Surface contour curve of hydrogen production rate for 1.7 MHz transducer frequency (a) MeOH vs. FFR; (b) Voltage vs FFR; and (c) Voltage vs MeOH.

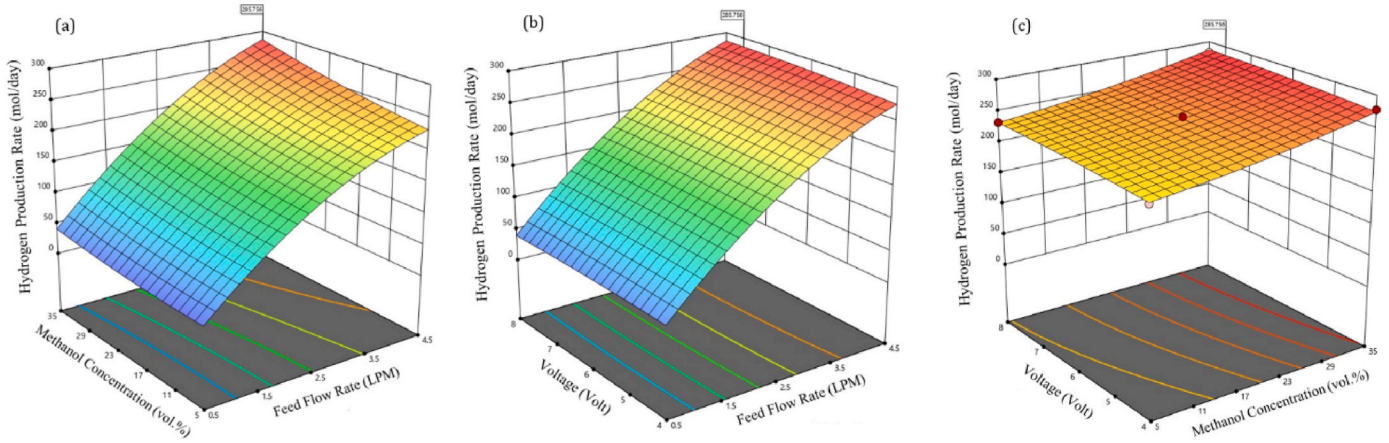


Fig. 6. 3D Surface contour curve of hydrogen production rate for 2.4 MHz transducer frequency (a) MeOH vs. FFR; (b) Voltage vs FFR; and (c) Voltage vs MeOH.

Table 8
ANOVA table of 2.4 MHz ultrasonic frequency for HPR.

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	1.286E+05	9	14287.37	608.92	<0.0001	significant
A-Feed Flow Rate	1.205E+05	1	1.205E+05	5136.32	<0.0001	
B-Methanol concentration	2752.28	1	2752.28	117.30	<0.0001	
C-Voltage	51.53	1	51.53	2.20	0.1692	
AB	856.98	1	856.98	36.52	0.0001	
AC	78.13	1	78.13	3.33	0.0980	
BC	0.8450	1	0.8450	0.0360	0.8533	
A^2	2616.47	1	2616.47	111.51	<0.0001	
B^2	60.87	1	60.87	2.59	0.1383	
C^2	10.95	1	10.95	0.4667	0.5100	
Residual	234.64	10	23.46			
Lack of Fit	234.64	5	46.93			
Pure Error	0.0000	5	0.0000			
Cor Total	1.288E+05	19				

Table 9
Regression table for hydrogen production rate.

Transducer Frequency	0.3 MHz	1.7 MHz	2.4 MHz
R^2	0.9972	0.9973	0.9982
Adjusted R^2	0.9947	0.9949	0.9965
Predicted R^2	0.9791	0.9797	0.9835
Adeq Precision	66.4889	64.4920	75.8035
Std. Dev.	4.62	5.38	4.84
Mean	115.54	137.89	158.11
C.V. %	4.00	3.90	3.06

3.4. Optimal values (comparative results)

This section discussed the comparative results obtained from RSM, ANN-GA, and experimental data. The optimal values of the parameters obtained from the RSM and ANN-GA model were used to perform experiments, and the results from the experimental data were computed for percentage error. The percentage errors for RSM values were between 4 % (for 1.7 MHz ultrasonic transducer frequency) and 7 % (for 0.3 MHz ultrasonic transducer frequency), whereas the percentage error for the ANN-GA model was about 2.5 % (for 1.7 MHz ultrasonic transducer frequency). Here, the ultrasonic transducer frequencies (0.3 MHz, 1.7 MHz and 2.4 MHz) were also considered as the optimizing parameter along with other process parameters; therefore, no separate results were obtained for different ultrasonic transducer frequencies. The comparative result showed that the ANN-GA model is superior in optimization to the RSM model, with a lower error percentage during analysis.

Table 10
Error percentage for the optimal parameters from the RSM model.

Transducer Frequency (MHz)	Parameters			Hydrogen Production rate (mol/day)		%age error
	Feed Flow Rate (LPM)	Methanol concentration (%)	Voltage (kV)	Predicted	Experimental	
0.3	4.5	35	7.86	229.2	213.8	6.7
1.7	4.5	35	7.50	264.9	254.0	4.1
2.4	4.5	35	7.54	285.8	270.5	5.4

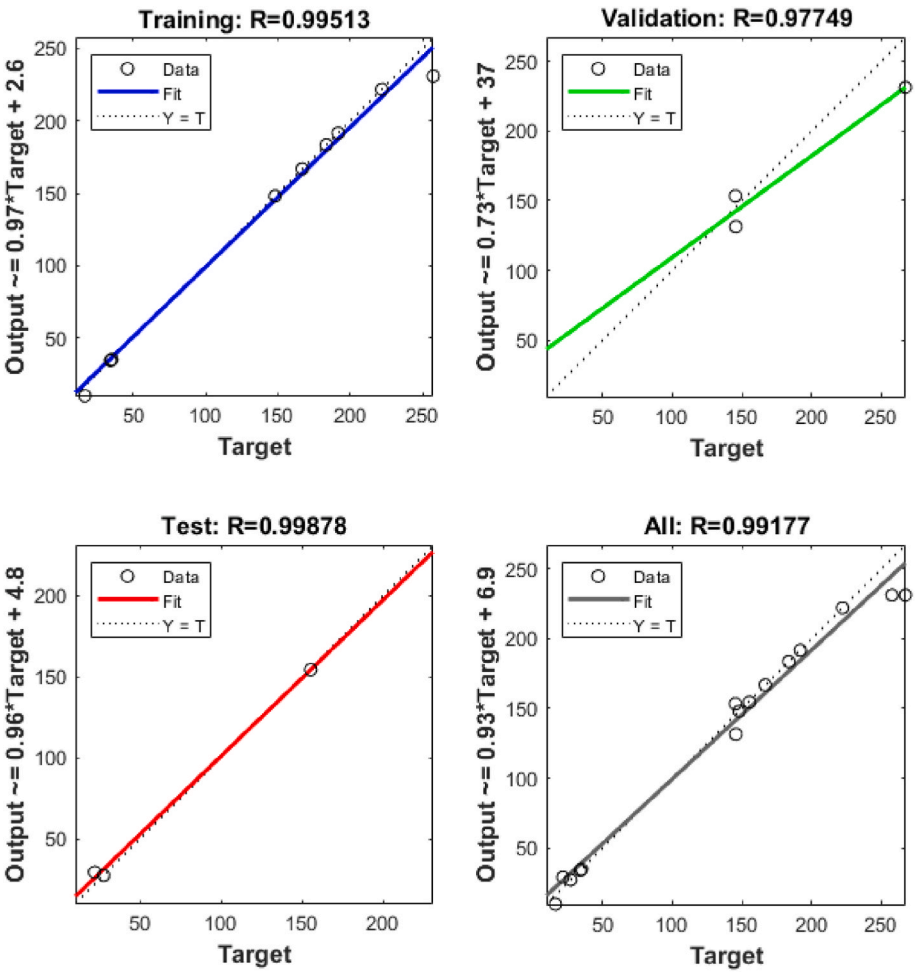


Fig. 7. Coefficient of correlation (R) for hydrogen production rate.

3.5. Hydrogen purity analysis

The purity of a compound is an essential criterion; hence, the gaseous mixture was collected in a 1 L tedlar bag and tested for the gas composition using Agilent make gas chromatography (GC) having thermal conductivity detector (TCD). The composition of a gaseous mixture was found to have the following gases: hydrogen, carbon dioxide, nitrogen, and oxygen, as observed in Fig. 8 (a). It was observed that about 32.2 % of hydrogen content was present in the gaseous mixture.

The gas mixture was then purified using an oxygen concentrator-type hydrogen purification unit, and the samples were collected after the gas purification. The same procedure was followed to determine the purified gas composition. It was observed that the hydrogen content in the sample rose to about 99 %, and a small fraction of about 1 % of oxygen was present after the purification, which could be the oxygen released during methanol-reforming.

4. Conclusions

The experimental analysis of hydrogen production using plasma reforming of water-methanol mixture was performed. The study was conducted on the influence of three operating parameters: input voltage, methanol concentration, and feed flow rate at different ultrasonic transducer frequencies, i.e., 0.3 MHz, 1.7 MHz, and 2.4 MHz, respectively. The operating parameters varied to understand their interactive influence on the hydrogen production rate. The data was optimized using response surface methodology and an artificial neural network-genetic algorithm. The following outcomes were observed.

- Higher ultrasonic frequency has positive impacts on hydrogen production rate.
- The F-values and p-values of the methanol concentration and feed flow rate showed they were significant parameters in the plasma reforming process.
- The R^2 value from RSM showed good competence of the outcome with the statistically predicted values.
- The optimal hydrogen production rate of 229.2 mol/day (for 0.3 MHz), 265.0 mol/day (for 1.7 MHz), and 285.8 mol/day (for 2.4 MHz) for RSM were observed at 4.5 LPM and 35 % methanol concentration while 267.0 mol/day for ANN-GA at 4.5 LPM, 33.8 % methanol concentration, and 7.4 kV input voltage.
- The higher methanol concentration enhanced the hydrogen production rate.
- The ANN-GA model also showed correlation coefficient values close to unity; thus, the model was reliable.
- The experimental tests on the optimized values of RSM and ANN-GA models showed 4–7 % and about 2.5 % percentage errors, demonstrating the ANN-GA model's superiority over RSM.

Therefore, the experimental analysis of plasma reforming of a water-methanol mixture showed the potential of generating a gaseous mixture rich in hydrogen gas. The hydrogen production was successfully optimized for an enhanced production rate. The results obtained were

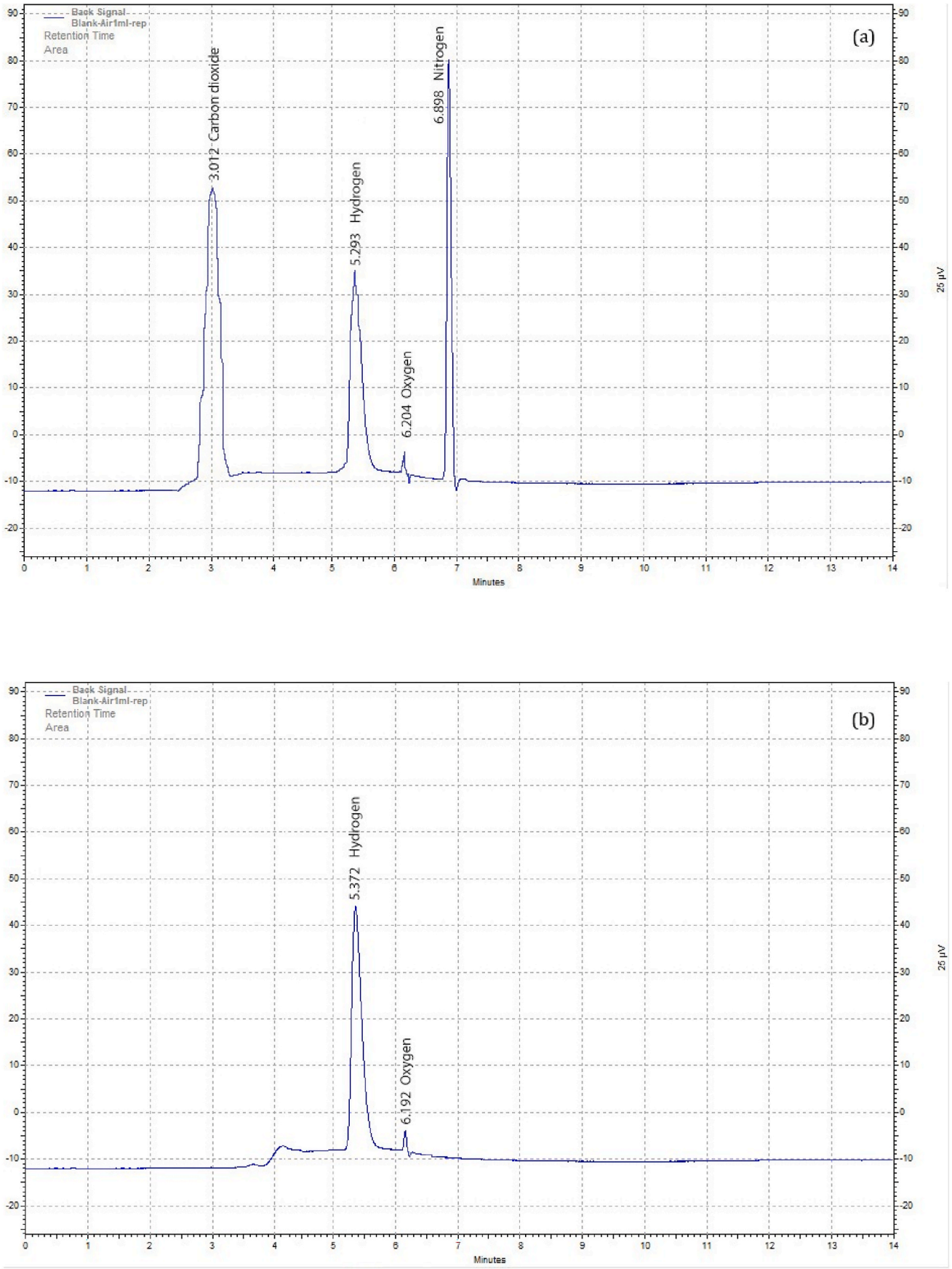


Fig. 8. Chromatogram for (a) gaseous mixture and (b) purified gas.

reliable. Further work may be performed to up-scale the plasma reactor for commercialization, and other process parameters may also be considered for optimization.

CRediT authorship contribution statement

Neeraj Budhraja: Writing – original draft, Visualization, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Amit Pal:** Writing – review & editing, Validation, Supervision, Software, Resources, Methodology, Data curation. **R.S. Mishra:** Writing – review & editing, Validation, Supervision.

Data availability

All the data used in the manuscript have been appropriately cited with the corresponding reference.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors are thankful to the I. C. Engine Lab, Department of Mechanical Engineering, Delhi Technological University, for developing a plasma reformer to conduct experiments within the lab.

References

[1] F. Qureshi, M. Yusuf, H. Kamyab, S. Zaidi, M. Junaid Khalil, M. Arham Khan, M. Azad Alam, F. Masood, L. Bazli, S. Chelliapan, B. Abdullah, Current trends in hydrogen production, storage and applications in India: a review, Sustain. Energy Technol. Assessments 53 (2022) 102677, https://doi.org/10.1016/j.seta.2022.102677.

[2] N. Budhraja, A. Pal, R.S. Mishra, Plasma reforming for hydrogen production: pathways, reactors and storage, Int. J. Hydrogen Energy 48 (2023) 2467–2482, https://doi.org/10.1016/j.ijhydene.2022.10.143.

[3] Q. Jin, X. Meng, P. Wu, Y. Li, M. Xu, R. Zhou, M. Yang, H. Xu, Methanol steam reforming for hydrogen production over $NiTiO_3$ nanocatalyst with hierarchical porous structure, RSC Adv. 13 (2023) 16342–16351, https://doi.org/10.1039/D3RA02891G.

[4] M. Shamsi, S. Moghaddas, E. Naeiji, S. Farokhi, Techno-economic, energy, exergy, and environmental comparison of hydrogen production from natural gas, biogas, and their combination as feedstock, Arabian J. Sci. Eng. 48 (2023) 8971–8987, https://doi.org/10.1007/s13369-022-07581-z.

[5] Q. Wu, Y. Wang, H. Zhou, X. Qiu, D. Mei, Measurement of reaction temperature distribution inside of methanol steam reforming microreactor using infrared thermography, Appl. Therm. Eng. 238 (2024) 121909, https://doi.org/10.1016/j.applthermaleng.2023.121909.

[6] M.K. Awasthi, R.K. Rai, S. Behrens, S.K. Singh, Low-temperature hydrogen production from methanol over a ruthenium catalyst in water, Catal. Sci. Technol. 11 (2021) 136–142, https://doi.org/10.1039/D0CY01470B.

[7] Y. Xin, B. Sun, X. Zhu, Z. Yan, X. Sun, Hydrogen-rich syngas production by liquid phase pulsed electrodeless discharge, Energy 214 (2021) 118902, https://doi.org/10.1016/j.energy.2020.118902.

[8] Y.-C. Pu, S.-R. Li, S. Yan, X. Huang, D. Wang, Y.-Y. Ye, Y.-Q. Liu, An improved Cu/ZnO catalyst promoted by Sc2O3 for hydrogen production from methanol reforming, Fuel 241 (2019) 607–615, https://doi.org/10.1016/j.fuel.2018.12.067.

[9] M. Kheirollahivash, F. Rashidi, M.M. Moshrefi, Experimental study and kinetic modeling of methane decomposition in a rotating arc plasma reactor with different cross-sectional areas, Int. J. Hydrogen Energy 44 (2019) 17460–17469, https://doi.org/10.1016/j.ijhydene.2019.05.011.

[10] Y. Xin, B. Sun, X. Zhu, Z. Yan, X. Zhao, X. Sun, T. Ohshima, Characteristics and pathways of hydrogen produced by pulsed discharge in ethanol-water mixtures, Int. J. Hydrogen Energy 45 (2020) 1588–1596, https://doi.org/10.1016/j.ijhydene.2019.11.101.

[11] N. Budhraja, A. Pal, A. Kumar, in: A. Kumar, A. Pal, S.S. Kachhwaha, P.K. Jain (Eds.), Investigations on Ethanol as the Raw Material for Hydrogen Production, Storage, and Applications, Springer Singapore, Singapore, 2021, pp. 317–326, https://doi.org/10.1007/978-981-15-9678-0_28.

[12] Y. Hayakawa, S. Kambara, T. Miura, Hydrogen production from ammonia by the plasma membrane reactor, Int. J. Hydrogen Energy 45 (2020) 32082–32088, https://doi.org/10.1016/j.ijhydene.2020.08.178.

[13] Y. Pang, T. Hammer, D. Müller, J. Karl, Investigation on the influence of non-thermal plasma on reaction degree of wood gasification in a drop tube reactor, Fuel 253 (2019) 95–105, https://doi.org/10.1016/j.fuel.2019.04.165.

[14] N.A. Bulychev, M.A. Kazaryan, A.S. Averyushkin, A.A. Chernov, A.L. Gusev, Hydrogen production by low-temperature plasma decomposition of liquids, Int. J. Hydrogen Energy 42 (2017) 20934–20938, https://doi.org/10.1016/j.ijhydene.2016.09.226.

[15] N. Budhraja, A. Pal, R.S. Mishra, Modeling and analysis for enhanced hydrogen production in process simulation of methanol reforming, Energy Sources, Part A Recovery, Util. Environ. Eff. 45 (2023) 11553–11565, https://doi.org/10.1080/15567036.2023.2262414.

[16] O. Akande, B. Lee, Plasma steam methane reforming (PSMR) using a microwave torch for commercial-scale distributed hydrogen production, Int. J. Hydrogen Energy 47 (2022) 2874–2884, https://doi.org/10.1016/j.ijhydene.2021.10.258.

[17] S.M.A. Mousavi, W. Piavis, S. Turn, Reforming of biogas using a non-thermal, gliding-arc, plasma in reverse vortex flow and fate of hydrogen sulfide contaminants, Fuel Process. Technol. 193 (2019) 378–391, https://doi.org/10.1016/j.fuproc.2019.05.031.

[18] B. Wang, S. Liu, Y. Peng, C. Wang, J. Zou, Heptane dry reforming and coupling with partial oxidation in gliding arc discharge plasma for H2 production, Fuel Process. Technol. 221 (2021) 106943, https://doi.org/10.1016/j.fuproc.2021.106943.

[19] Y. Xin, B. Sun, X. Zhu, Z. Yan, X. Zhao, X. Sun, Hydrogen production from ethanol solution by pulsed discharge with TiO2 catalysts, Int. J. Hydrogen Energy 43 (2018) 9503–9513, https://doi.org/10.1016/j.ijhydene.2018.04.062.

[20] D.-Y. Huang, J.-H. Jang, W.-R. Tsai, W.-Y. Wu, Improvement in hydrogen production with plasma reformer system, Energy Proc. 88 (2016) 505–509, https://doi.org/10.1016/j.egypro.2016.06.070.

[21] L. Chen, P. Li, S. Xia, R. Kong, Y. Ge, Multi-objective optimization for membrane reactor for steam methane reforming heated by molten salt, Sci. China Technol. Sci. 65 (2022) 1396–1414, https://doi.org/10.1007/s11431-021-2003-0.

[22] P. Li, L. Chen, S. Xia, R. Kong, Y. Ge, Multi-objective optimal configurations of a membrane reactor for steam methane reforming, Energy Rep. 8 (2022) 527–538, https://doi.org/10.1016/j.egyr.2021.11.288.

[23] H.-Y. Lian, J.-L. Liu, X.-S. Li, X. Zhu, A.Z. Weber, A.-M. Zhu, Plasma chain catalytic reforming of methanol for on-board hydrogen production, Chem. Eng. J. 369 (2019) 245–252, https://doi.org/10.1016/j.cej.2019.03.069.

[24] R. Shiraishi, S. Nomura, S. Mukasa, R. Nakano, R. Kamatoko, Effect of catalytic electrode and plate for methanol decomposition by in-liquid plasma, Int. J. Hydrogen Energy 43 (2018) 4305–4310, https://doi.org/10.1016/j.ijhydene.2018.01.060.

[25] H.-Y. Lian, X.-S. Li, J.-L. Liu, X. Zhu, A.-M. Zhu, Oxidative pyrolysis reforming of methanol in warm plasma for an on-board hydrogen production, Int. J. Hydrogen Energy 42 (2017) 13617–13624, https://doi.org/10.1016/j.ijhydene.2016.10.166.

[26] N. Budhraja, A. Pal, R.S. Mishra, Experimental investigation of enhanced hydrogen production from atomised water-methanol mixture in a modified plasma reformer, Thermal Science and Engineering Progress 61 (2025) 103535, https://doi.org/10.1016/j.tsep.2025.103535.

[27] N. Budhraja, A. Pal, R.S. Mishra, Optimizing methanol reforming parameters for enhanced hydrogen selectivity in an aspen hysys simulator using response surface methodology, Energy Technol. 11 (2023), https://doi.org/10.1002/ente.202300203.

[28] N. Budhraja, A. Pal, R.S. Mishra, Simulation and optimization of biohydrogen production from biomass feed via anaerobic digestion, Chem. Eng. Technol. 47 (2024) 706–715, https://doi.org/10.1002/ceat.202300395.

[29] N. Budhraja, Simulation and optimization for biohydrogen production potential of various organic waste via anaerobic digestion, Fuel 360 (2024) 130563, https://doi.org/10.1016/j.fuel.2023.130563.

[30] N. Budhraja, A. Pal, R.S. Mishra, Parameter optimization for enhanced biodiesel yield from Linum usitatissimum oil through solar energy assistance, Biomass Convers Biorefin (2022), https://doi.org/10.1007/s13399-022-03649-w.

Check for updates

ChemistrySelect

Research Article
doi.org/10.1002/slct.202404835

Chemistry Europe
European Chemical Societies Publishing

www.chemistryselect.org

Zn/Cu Layered Double Hydroxide as a Catalyst for the Synthesis of Highly Functionalized Pyrazole Derivatives

Charu Gupta,[a] Bhamini Pandey,[a] Poonam Singh,[a] and Ram Singh*[a]

A 2D layered double hydroxides (LDHs) of Zn/Cu that belongs to a diverse category of anion-exchangeable layered structures was prepared. The synthetic inorganic LDHs material was characterized by employing various techniques such as Fourier transform infrared (FTIR), Scanning electron microscope (SEM), Powder X-ray diffraction (XRD), and thermogravimetric analysis (TGA). The prepared heterogeneous catalyst mediated the synthesis of highly functionalized pyrazole derivatives via a one-pot reaction between arylaldehyde, malononitrile, and phenylhydrazine in ethanol at room temperature with excellent yields of 76%–94%. Notably, the catalyst can be simply retrieved from the reaction mixture using filtration, and it can be reused multiple times without experiencing a substantial decrease in its activity.

1. Introduction

The study of two-dimensional materials has gained significant attention in recent years, particularly in the context of developing adsorbents, heterogeneous solid catalysts, catalyst supports, metal-sulfur batteries, electrodes, and so on[1–3] Layered double hydroxides (LDHs) are 2D materials characterized by their general formula $[M^{2+}_{(1-x)}M_x^{3+}(OH)_2](A^{n-})_{x/n}.zH_2O]$ belong to a diverse category of anion-exchangeable layered structures. The commonly used metal cations in divalent and trivalent forms include Mg^{2+}, Zn^{2+}, Mn^{2+}, Co^{2+}, Ca^{2+}, Fe^{2+}, and Co^{3+}, Fe^{3+}, Mn^{3+}, Al^{3+}, and Cr^{3+}, respectively and the anions are usually carbonate, chloride, nitrate, or acetate.[4–6] The versatile behavior of LDHs is derived from their facile synthesis and ability to modify their positively charged layers and negatively charged interlamellar region.[7] This versatility results in diverse applications of LDHs in various fields. Kamali et al. showed the removal of As(III) ion from aqueous solutions by using a new nanospherical graphene that was modified with LDHs[8] and also contributed to various organic syntheses.[9–11] Furthermore, recent years have witnessed a surge in research focusing on utilizing LDHs-based catalysts in multicomponent reactions (MCRs). These reactions involve the simultaneous participation of multiple reactants where LDHs play a pivotal role in enhancing the selectivity and efficiency of the reaction.[12–15]

LDHs act as heterogenous catalysts and offer distinct advantages including easy and efficient separation, contributing to cost-effective catalyst recovery.[16,17] MCRs comprise an important class of chemical transformations that have been applied to almost all fields of organic and medicinal chemistry.[18] These convergent procedures incorporate three or more reactants into a final product in one-pot, thus combining high levels of complexity and diversity generation with low synthetic cost and high atom economy.[19] Most of the reactions performed are multicomponent one-pot reactions. MCR has been a successful method for the synthesis of heterocyclic compounds including pyrazoles.[20,21] Pyrazole scaffold leads to diversified applications in different areas such as technology, medicine, and agriculture.[22] They are also described as inhibitors of protein glycation, antibacterial, antifungal, anticancer, antidepressant, anti-inflammatory, antituberculosis, antioxidant as well as antiviral agents.[23–28]

Generally, the synthesis of 5-amino-pyrazole-4-carbonitrile derivatives is a three-component reaction of aromatic aldehyde, malononitrile, and phenylhydrazine. For the synthesis of pyrazoles and their derivatives many catalysts and reagents have been studied such as Fe_3O_4@Alg,[29] sulfonic acid-functionalized polyvinyl alcohol (SPVA),[30] Fe_3O_4@SiO_2@Tannic acid,[31] NaCl,[28] 2-amino-6-oxo-3-(piperidinylamidino)-4-(4methoxyphenyl)-6,7-dihydro-pyrano[2,3-d]-5,7-thiazol derivative (MATYPd Complex),[32] graphene oxide–TiO_2,[33] oxone,[34] taurine,[35] and molecular iodine.[32] LDHs materials have found many applications in the heterogeneous catalysis field including photocatalysts, electrocatalysts, and thermocatalysts.[37] The composition, particle size, morphology, and surface defect structure, as well as the electronic properties of LDHs can be modulated by various synthetic strategies and their porous nature and larger surface area made them a suitable catalyst for various reactions including Michael addition affording the coordinatively unsaturated active sites (e.g., low-coordinated steps, edges, and/or corner atoms) which facilitate heterogeneous catalysis.[38] In this manuscript, we utilized LDHs materials as a catalyst for a multicomponent one-pot approach for the synthesis of substituted 5-amino-pyrazole-4-carbonitrile framework.

[a] C. Gupta, B. Pandey, P. Singh, R. Singh
Department of Applied Chemistry, Delhi Technological University, Delhi 110042, India
E-mail: ramsingh@dtu.ac.in

Supporting information for this article is available on the WWW under https://doi.org/10.1002/slct.202404835

© 2025 Wiley-VCH GmbH

Research Article
doi.org/10.1002/slct.202404835

Chemistry Europe
European Chemical Societies Publishing

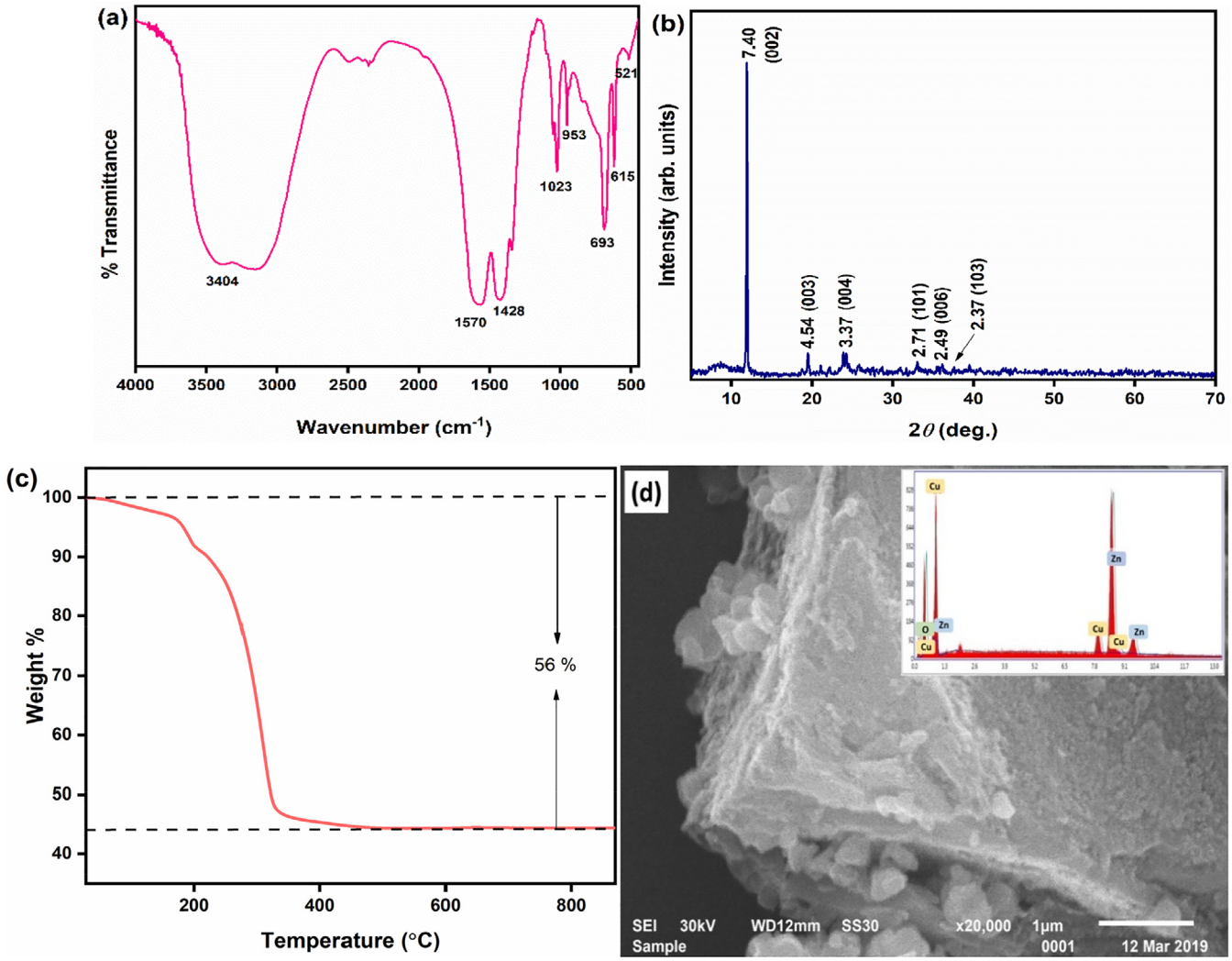


Figure 1. (a) FTIR spectrum, (b) PXRD pattern, (c) TGA, and (d) SEM micrograph (EDX spectrum in inset).

Table 1. The effect of temperature with catalyst (40 mg) on the yield of the product **4a**.

Entry	Temp. (°C)	Time (min)	Yield (%)
1	80	60	70
2	70	60	70
3	60	60	75
4	50	60	77
5	40	60	80
6	RT (~30 °C)	60	83
7	RT (~30 °C)	30	90
8	RT (~30 °C)	12	90

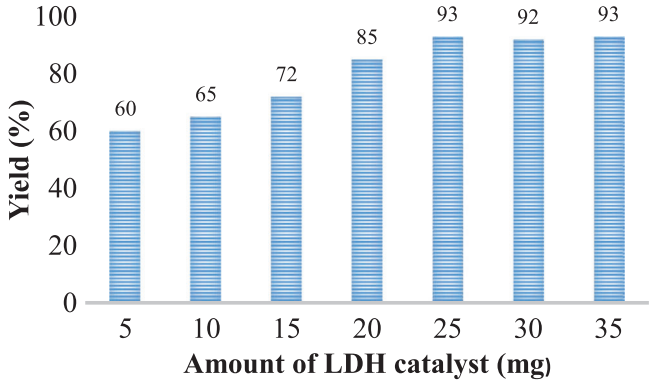


Figure 2. Catalyst optimization with yield for the synthesis of pyrazole **4a**.

2. Experimental Section

2.1. General Procedure for Synthesis of Catalyst Zn/Cu LDHs

Zn/Cu LDHs was prepared in accordance with the method specified in[39] which is briefly described as follows: Salts of zinc acetate dihydrate (4.49 mmol) and copper acetate monohydrate (0.49 mmol) were dissolved in 20 mL of deionized (DI) water. Then, 35 mL of H_2O_2 solution (5 mL H_2O_2 + 30 mL DI water) was added to the metal salts solution with continuous stirring, followed by ultrasonication for 30 min. Afterward, the resulting solution was stirred for 24 h and later dried at 60 °C. The obtained blue color powder was characterized with FTIR spectra, Powder X-ray diffraction (PXRD),

© 2025 Wiley-VCH GmbH

23656549, 2025, 13, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/slct.202404835 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

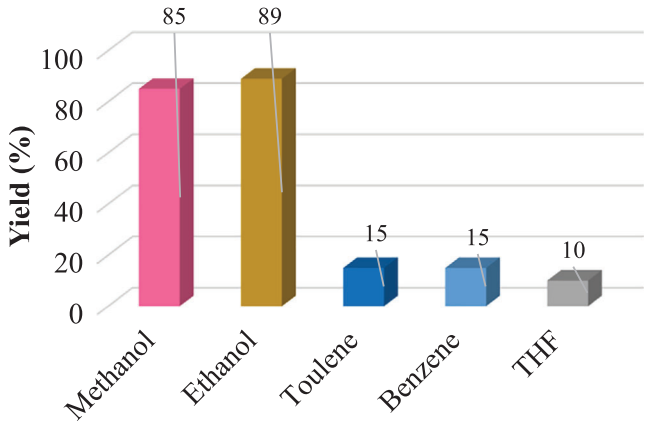


Figure 3. Solvent optimization with yield for the synthesis of pyrazole **4a**.

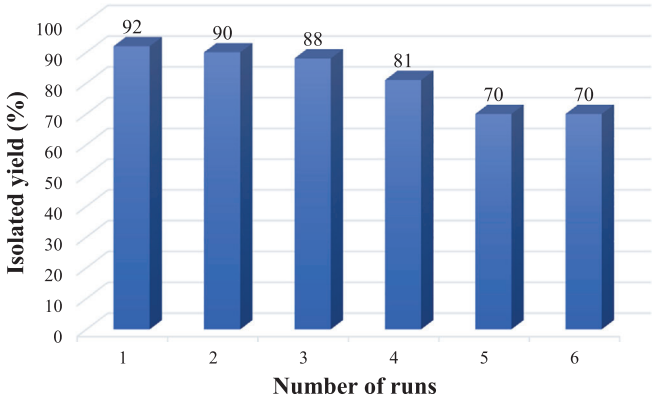


Figure 4. Reusability of the catalyst for the product **4a**.

thermogravimetric analysis (TGA), and Scanning electron microscope (SEM). This was used for the reaction.

2.2. General Procedure for the Preparation of Pyrazole Derivatives (4)

The synthesized catalyst (25 mg) was introduced into a reaction mixture containing benzaldehyde (1.0 mmol) and malononitrile (1.0 mmol) in ethanol solvent (3 mL). The mixture was stirred at room temperature until the formation of a white precipitate was observed. Subsequently, phenylhydrazine (1.0 mmol) was added to the resulting reaction mixture at room temperature under stirring for a duration of 5–10 min. The progress of the reaction was monitored using a thin-layer chromatography (TLC) plate in hexane and ethyl acetate (4:1 v/v). After completion of the reaction, water (30 mL) was added to the reaction mixture and filtered. The catalyst was separated, dried, and reused for the reaction. The filtrate was mixed with ethyl acetate and separated. The organic layer was subjected to a rotary evaporator to obtain the crude product that was further re-crystallized from ethanol (The spectra given in Supporting Information).

5-Amino-3-(2-nitrophenyl)-1-phenyl-1H-pyrazole-4-carbonitrile **(4a)**: IR (KBr, cm^{-1}) 3421, 3295, 3057, 2218, 1601, 1533, 1491, 1333, 1253, 1162, 1070, 990; ^1H NMR ($CDCl_3$): δ 8.31 (s, 2H, NH_2), 8.27 (d, 1H, J = 8.0 Hz, ArH), 7.99 (d, 1H, J = 8.6 Hz, ArH), 7.60 (t, 1H, ArH), 7.40 – 7.36 (m, 2H, ArH), 7.32 – 7.25 (m, 2H, ArH), 7.13 (d, 1H, J = 8.8 Hz, ArH), 6.94 (t, 1H, ArH); HRMS (ESI-MS) (m/z): calcd for $C_{16}H_{11}N_5O_2$, 305.29; found, 306.09

5-Amino-1,3-diphenyl-1H-pyrazole-4-carbonitrile **(4b)**: IR (KBr, cm^{-1}) 3312, 3027, 2204, 1617, 1597, 1442, 1358, 1257, 1132, 1065, 928; ^1H NMR ($CDCl_3$): δ 7.65 – 7.62 (m, 3H, ArH, NH_2), 7.38 – 7.23 (m, 6H, ArH), 7.10 (d, 2H, J = 7.6 Hz, ArH), 6.8 (t, 1H, ArH); HRMS (ESI-MS) (m/z): calcd for $C_{16}H_{12}N_4$, 260.30; found, 262.12

5-Amino-3-(4-methylphenyl)-1-phenyl-1H-pyrazole-4-carbonitrile **(4c)** IR (KBr, cm^{-1}): 3306, 3026, 2914, 2117, 1595, 1483, 1252, 812; ^1H NMR ($CDCl_3$): δ 7.67 (s, 2H, NH_2), 7.55 (d, 2H, J = 8.4 Hz, ArH), 7.36 – 7.23 (m, 4H, ArH), 7.17 (d, 1H, J = 8.0 Hz, ArH), 7.10 (d, 1H, J = 8 Hz, ArH), 6.86 (t, 1H, ArH), 2.36 (s, 3H, ArH); HRMS (ESI-MS) (m/z): calcd for $C_{17}H_{14}N_4$, 274.33; found, 275.28.

5-Amino-3-(4-nitrophenyl)-1-phenyl-1H-pyrazole-4-carbonitrile **(4d)** IR (KBr, cm^{-1}): 3294, 3026, 2422, 2120, 1595, 1494, 1315, 1151, 849; ^1H NMR ($CDCl_3$): δ 8.21 (d, 2H, J = 8.8 Hz, ArH), 7.78 – 7.70 (m, 3H, ArH, NH_2), 7.36 – 7.26 (m, 3H, ArH), 7.15 (d, 2H, J = 7.6 Hz, ArH), 6.95 (t, 1H, ArH); HRMS (ESI-MS) (m/z): calcd for $C_{16}H_{11}N_5O_2$, 305.29; found, 305.07.

5-Amino-3-(cinnamaldehyde)-1-phenyl-1H-pyrazole-4-carbonitrile **(4e)** IR (KBr, cm^{-1}): 3312, 3057, 3031, 2973, 2264, 1596, 1559, 1450, 1292, 1133, 1067; ^1H NMR ($CDCl_3$): δ 7.52 (d, 1H, J = 9.2 Hz, ArH), 7.44 (d, 2H, J = 7.4 Hz, ArH), 7.34 (t, 2H, ArH), 7.28 – 7.24 (m, 4H, ArH, NH_2), 7.06 – 7.02 (m, 2H, ArH), 7.01 (d, 1H, J = 9.2 Hz), 6.86 (t, 1H, ArH), 6.68 (d, 1H, J = 16 Hz); HRMS (ESI-MS) (m/z): calcd for $C_{18}H_{14}N_4$, 286.33; found, 287.13.

5-Amino-3-(4-chlorophenyl)-1-phenyl-1H-pyrazole-4-carbonitrile **(4f)** IR (KBr, cm^{-1}): 3681, 3311, 2972, 2844, 2203, 1597, 1560, 1485, 1265, 1131, 1033; ^1H NMR ($CDCl_3$): δ 7.64 (s, 1H, NH_2), 7.58 (d, 2H, J = 8.4 Hz, ArH), 7.36 – 7.26 (m, 5H, ArH), 7.11 (d, 2H, J = 8.4 Hz, ArH), 6.95 (t, 1H, ArH); HRMS (ESI-MS) (m/z): calcd for $C_{16}H_{11}ClN_4$, 294.74; found, 296.07.

5-Amino-3-(4-methoxyphenyl)-1-phenyl-1H-pyrazole-4-carbonitrile **(4g)** IR (KBr, cm^{-1}): 3312, 3023, 2834, 2554, 2223, 1595, 1498, 1294, 1103, 1026, 881; ^1H NMR ($CDCl_3$): δ 7.67 (s, 1H, ArH), 7.60 (d, 2H, J = 8.8 Hz, ArH), 7.33 (m, 3H, ArH, NH_2), 7.18 (d, 2H, J = 7.6 Hz, ArH), 6.91 (d, 2H, J = 8.8 Hz, ArH), 6.85 (t, 1H, ArH), 3.83 (s, 3H, OCH_3).

5-Amino-3-(3-methoxy-4-hydroxyphenyl)-1-phenyl-1H-pyrazole-4-carbonitrile **(4h)** IR (KBr, cm^{-1}): 3488, 3317, 2847, 2117, 1669, 1591, 1423, 1114; ^1H NMR ($CDCl_3$): δ 7.62 (s, 1H, ArH), 7.30 (d, 1H, J = 2 Hz, ArH), 7.23 – 7.18 (m, 3H, ArH, NH_2), 7.03 – 6.25 (m, 3H, ArH), 6.84 – 6.77 (m, 2H, ArH), 3.91 (s, 3H, OCH_3); HRMS (ESI-MS) (m/z): calcd for $C_{17}H_{14}N_4O_2$, 306.32; found, 306.30.

5-Amino-3-(3-chlorophenyl)-1-phenyl-1H-pyrazole-4-carbonitrile **(4i)**: IR (KBr, cm^{-1}): 3325, 3057, 2960, 2554, 2277, 1694, 1581, 1479, 1254, 1141, 1075, 885; ^1H NMR ($CDCl_3$): δ 7.58 (s, 2H, NH_2), 7.47 (s, 1H, ArH), 7.38 (d, 1H, J = 7.2 Hz, ArH), 7.22 – 7.15 (m, 4H, ArH), 7.02 (d, 2H, J = 8 Hz, ArH), 6.81 (t, 1H, ArH).

5-Amino-3-(4-fluorophenyl)-1-phenyl-1H-pyrazole-4-carbonitrile **(4j)** IR (KBr, cm^{-1}): 3310, 3035, 2979, 2200, 1598, 1502, 1445, 1231, 1131, 834; ^1H NMR ($CDCl_3$): δ 7.66 – 7.60 (m, 3H, ArH), 7.29 – 7.22 (m, 3H, ArH, NH_2), 7.11 – 7.03 (m, 4H, ArH), 6.86 (t, 1H, ArH); HRMS (ESI-MS) (m/z): calcd for $C_{16}H_{11}FN_4$, 278.29; found, 275.27.

3. Results and Discussion

3.1. Preparation and Characterization of Zn/Cu LDHs Catalyst

The preparation of the catalyst Zn/Cu LDHs was done according to the literature method.[39] Salts of zinc acetate dihydrate and copper acetate monohydrate were dissolved in DI water. To this solution, aqueous H_2O_2 solution was added with continuous stirring, followed by ultrasonication and stirring. This was filtered and the solid was dried at 60 °C. The obtained blue color powder

© 2025 Wiley-VCH GmbH

23656549, 2025, 13, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/slct.202404835 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

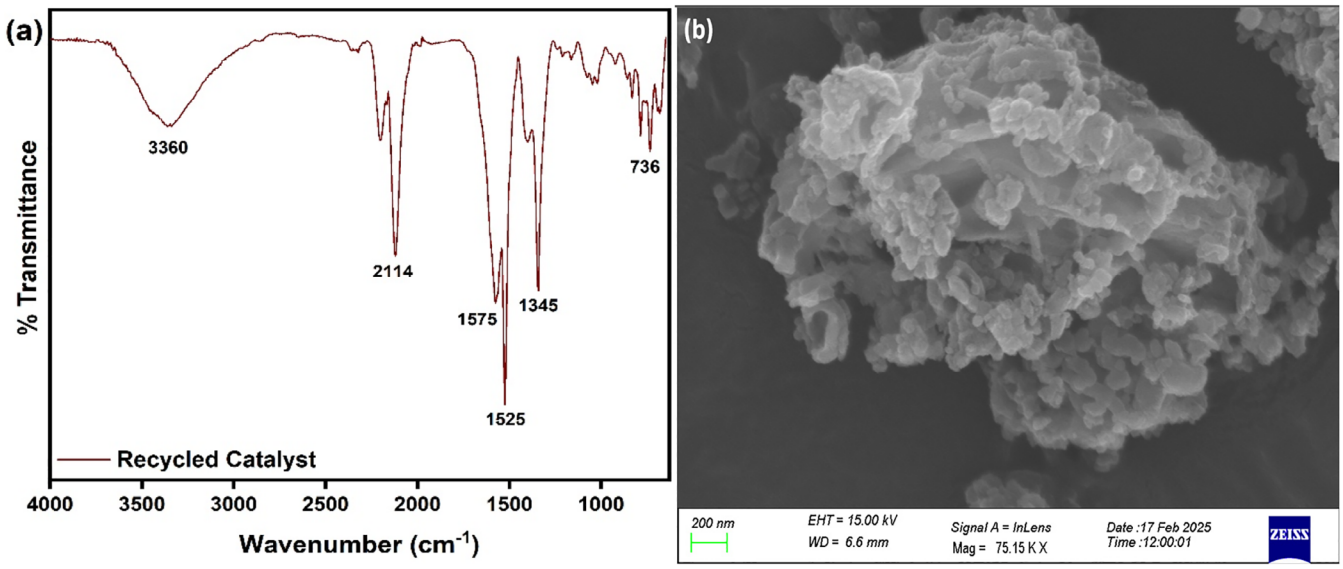


Figure 5. (a) FTIR spectrum and (b) FESEM of recycled catalyst.

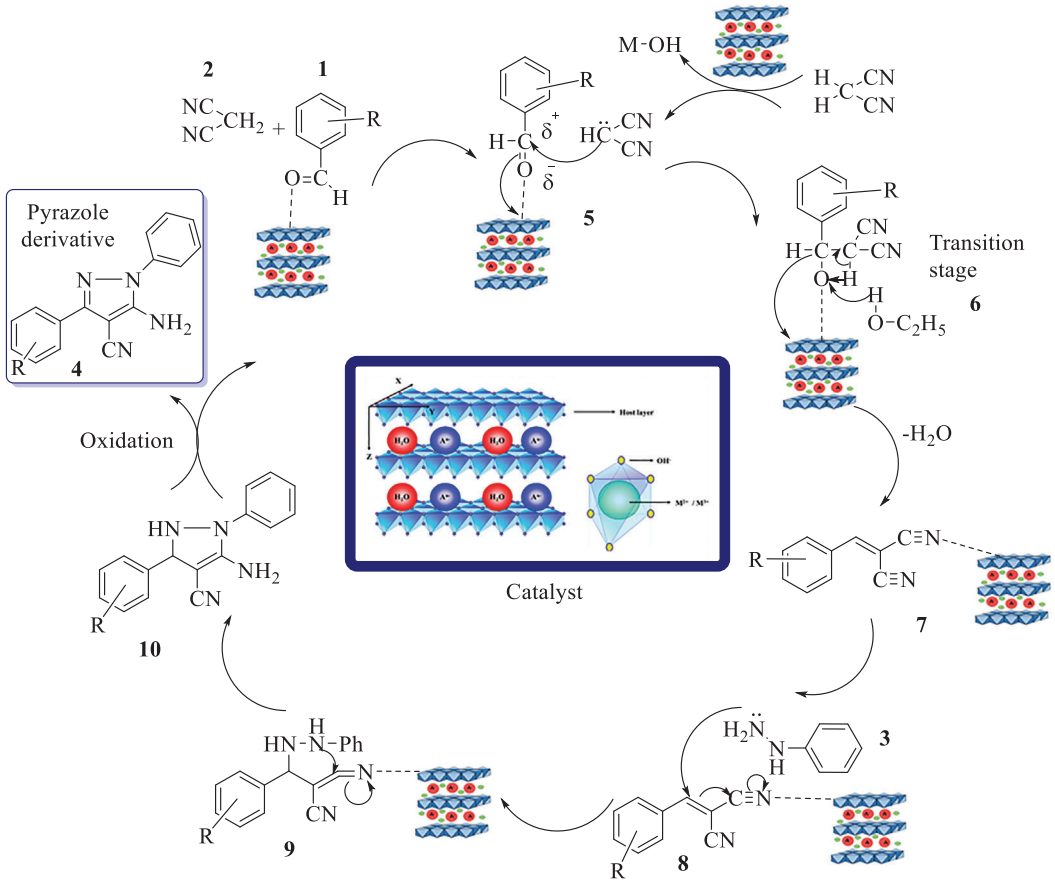


Figure 6. A proposed reaction mechanism for the synthesis of pyrazole derivatives.

was characterized with FTIR spectra, PXRD, TGA, SEM, and used for the reaction.

FTIR spectroscopy was used to analyze the presence of characteristic bands of hydrotalcite-like materials in the infrared spectrum of LDHs. In Figure 1a, the signal within the range of 3000–3500 cm^{-1} corresponds to the stretching vibrations associated with the hydrogen bonds formed by the —OH groups and water molecules present in the internal gallery region. The interlayer acetate anions are characterized by the presence of antisymmetric and symmetric stretching vibrations at 1570 and

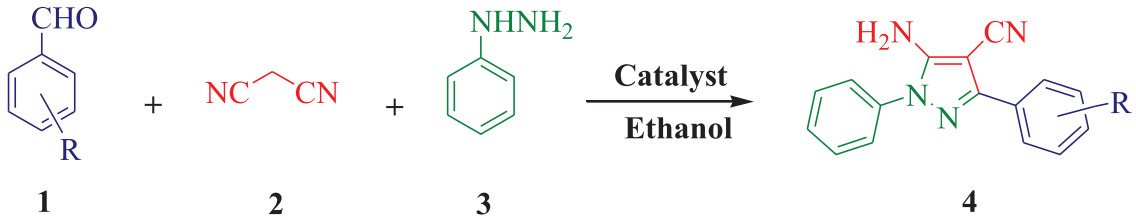
© 2025 Wiley-VCH GmbH

23656549, 2025, 13, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/slct.202404835 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

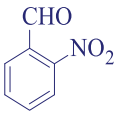
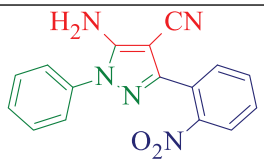
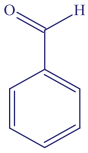
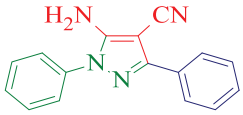
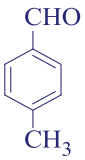
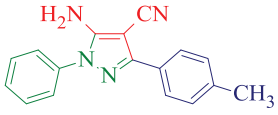
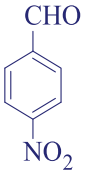
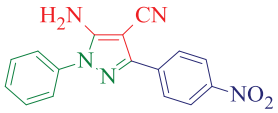
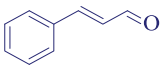
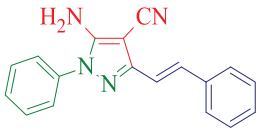
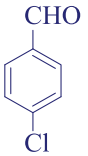
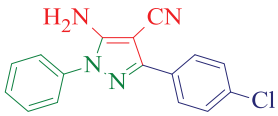
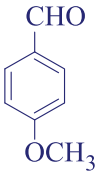
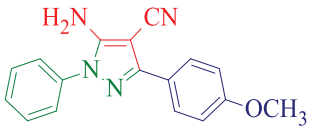
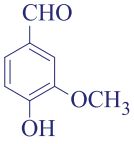
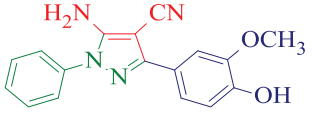
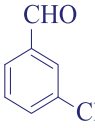
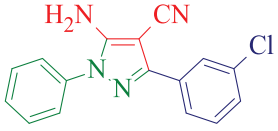
Research Article
doi.org/10.1002/slct.202404835

Chemistry Europe
European Chemical Societies Publishing

Table 2. Synthesis of 5-amino-3-aryl-1-phenyl-1H-pyrazole-4-carbonitrile derivatives (**4**).



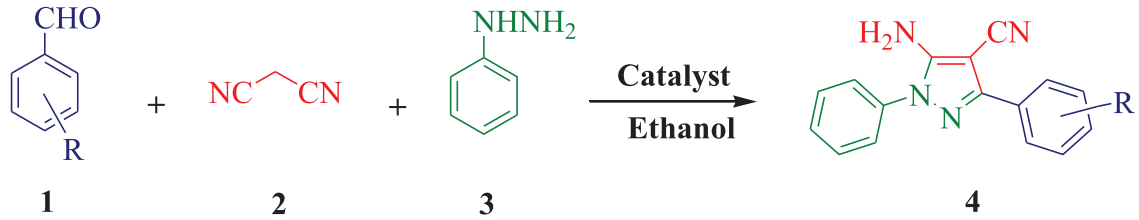
1 **2** **3** **4**

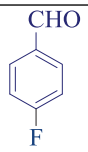
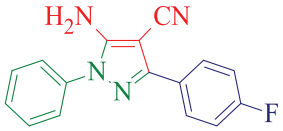
Entry	Aldehyde (1)	Product (4)	Time (in min)	Yield (%)	Melting Point (°C) Found	Melting Point (°C) Literature
4a			12	92	158	160–162[44]
4b			8	87	159	158–159[44]
4c			14	94	119	120[44]
4d			9	93	163	164–165[45]
4e			12	84	156	NA
4f			10	92	124	128–129[45]
4g			15	88	108	106–108[46]
4h			18	76	118	120–121[47]
4i			12	90	135–136	139–140[48]

© 2025 Wiley-VCH GmbH

23656549, 2025, 13, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/slct.202404835 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 2. (Continued)



Entry	Aldehyde (1)	Product (4)	Time (in min)	Yield (%)	Melting Point (°C) Found	Literature
4j			7	93	140	144–146[49]

1428 cm^{-1} respectively.[40] The bands at 1023, 693, and 615 cm^{-1} correspond to C—H rocking vibration of the CH_3 group, O—C—O bending, and CO_2^- rocking vibrations, respectively. These observations confirm the existence of interlayered water molecules and acetate anions. Additionally, a small band can also be seen around 521 cm^{-1} which can be attributed to the vibrational mode of M—O bonds.[41]

The PXRD pattern of the catalyst synthesized using the acid hydrolysis route is presented in Figure (1b). The obtained d-spacing values are also given in the PXRD pattern. The pattern shows a sharp and intense peak at 12.13° which is followed by the weak intensity peaks at 19.54, 24.22, 33.01, 35.53, and 37.86° corresponding to (002), (003), (004), (101), (006), and (103) hkl planes, respectively, confirming the formation of the lamellar structure of LDHs.[42] Furthermore, thermal analysis technique was employed to study the stability of catalyst (Figure 1c). As the temperature rises, three-level mass reduction became apparent with a total weight loss of 56%. The initial decline in mass, occurring at temperatures below 210 °C, can be attributed to the release of water from the sample, primarily situated on the surface and within the layers. The second and third degradation steps observed below 380 °C, might be associated with the dihydroxylation of brucite layers and decomposition of acetate anions, respectively.[43]

The morphology of synthesized catalyst was examined using SEM/EDX analysis technique (Figure 1d). The SEM micrograph showed that LDHs particles are present in plate-like arrangement. The elements present in the synthesized catalyst were confirmed by EDX analysis which also indicated that there exists a uniform distribution of zinc, copper, carbon, and oxygen elements in the lattice. The porosity and surface area of synthesized Zn/Cu LDHs was found to be 3.656 nm and 61.842 m^2g^{-1}, respectively as determined using adsorption-desorption isotherm (BET method). The result was in accordance with the literature value.[39] These results demonstrated that the catalyst exhibited a moderately high surface area offering ample number of active sites for the reaction to take place.

3.2. Synthesis of Pyrazole Derivatives Using LDHs Zn/Cu Catalyst

We have started the reaction with electron-withdrawing aryl aldehyde, 2-nitrobenzaldehyde (**1a**) with the view that if a reaction proceeds with this substrate, the process will be useful to another substrate also. Knoevenagel condensation is performed at room temperature using **1a** and malononitrile (**2**) reactants to yield 2-arylidenemalononitrile (**7**), which goes through a transition stage in the solvent ethanol. Further, Michael addition takes place between Knoevenagel product and phenylhydrazine (**3**) which undergoes rearrangements to give the final product, 5-amino-3-(2-nitrophenyl)-1-phenyl-1H-pyrazole-4-carbonitrile (**4a**) on oxidation.

To optimize the amount of catalyst and reaction temperature, a series of reactions were performed using reagents (**1a**), (**2**), and (**3**) over a varied range of temperature and catalyst amounts. First, we optimized the temperature by performing the reaction of **1a** with **2** and **3** using a random 40 mg of the catalyst (Table 1). We observed an increase in isolated yield in lowering the temperature from 80 °C to RT (~30 °C). Under optimized conditions, **4a** was obtained in 90% yield at RT in 12 min (Table 1, entry 8). Also, refluxing/heating to longer periods gave less yields with multiple unidentified spots on TLC. After fixing the temperature to RT, we optimized the amount of catalyst for the same reaction as 25 mg where **4a** was obtained in 92% yield (Figure 2).

The effect of solvent on the product yield was also studied which suggested that with nonpolar solvents such as toluene and benzene the product was obtained with negligible yield (Figure 3). However, polar solvents like ethanol and methanol resulted in significant yields within a short period of time as polar solvents released proton that was abstracted by oxygen of arylaldehyde (**6**) in the transition stage which undergoes rearrangement to form benzylidene malononitrile (**7**) by removing water molecule. The synthesized solid heterogeneous catalyst was recovered at the end of pyrazole synthesis, washed with ethanol, dried in the oven for 6–7 h, and reused for the conse-

© 2025 Wiley-VCH GmbH

23656549, 2025, 13, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/slct.202404835 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

quent run under vacuum at 70 °C. The dried catalyst was further reused for the next cycle of the same reaction and it showed decreasing in catalytic activity from cycle 4 (Figure 4).

The stability of the recycled catalyst was studied using FTIR and FESEM (Figure 5). The FTIR spectrum of recycled catalyst confirmed the presence of characteristic peaks of LDHs with a slight shift and weakening of the intensity of the absorption bands suggesting minor structural changes. The appearance of additional peaks can be attributed to the possible interaction between the reaction intermediates and the catalyst. Additionally, in the FESEM micrograph there appear small aggregated particles on the surface of the catalyst which may be due to the deposition of reaction residues due to chemical or thermal stress. The plate-like morphology of the LDHs is still visible, however some layers appear fragmented with rough edges.

A mechanism is proposed for the formation of highly functionalized pyrazoles using LDHs catalyst (Figure 6). In the catalyst, the metal centers (Zn/Cu) impart acidic sites and the acetate ions/ hydroxyl ions act as a weak base. The Knoevenagel condensation takes place between the arylaldehyde derivatives (**1**) and malononitrile (**2**) and the reaction is initiated by the attack of the polarized carbonyl oxygen from the arylaldehyde derivative on the metal center of the catalyst (**5**). The hydroxyl group present in the catalyst detached the acidic proton from malononitrile to produce the carbonium anion which reacts with the carbonyl group of benzaldehyde to give an intermediate or transition stage (**6**) which undergoes rearrangement to give the benzylidene malononitrile (**7**). The Michael addition takes place between benzylidene malononitrile (**7**) and phenylhydrazine (**3**) to give the Michael addition product (**10**). After that, the Michael addition product undergoes intermolecular cyclization which further undergoes rearrangement to give final product **4** on oxidation.

To observe the prosperity of the reaction, a wide variety of structurally different aromatic aldehydes, including electron releasing and electron-withdrawing substituents, were used, and gave the corresponding products in excellent yields without the formation of any major byproducts (Table 2).

4. Conclusion

We have successfully prepared the heterogenous solid catalyst, liquid double hydroxides (LDHs) of Zn/Cu, and characterized it using FTIR spectra, PXRD, TGA, and SEM. Furthermore, the prepared catalyst served as an efficient heterogenous catalyst in the preparation of pyrazole derivatives. The one-pot multicomponent cycloaddition of arylaldehyde, malononitrile, and phenylhydrazine in the presence of prepared LDHs afforded the pyrazole derivatives in 76%–94% yield. The developed procedure showed easy work up with catalyst reusability of four cycles with almost consistent catalytic efficiency.

Acknowledgements

The author Charu Gupta is thankful to Delhi Technological University (DTU) and Bhamini Pandey is thankful to CSIR-UGC for fellowships. All the authors are thankful to DTU for infrastructural support.

Conflict of Interests

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: Catalysis · Layered double hydroxides · One-pot synthesis · Pyrazole Derivatives · Zn/Cu

[1] K. Khan, A. K. Tareen, M. Aslam, R. Wang, Y. Zhang, A. Mahmood, Z. Ouyang, H. Zhang, Z. Guo, *J. Mater. Chem. C* **2020**, *8*, 387–440.
[2] , E. E. Abdel-Hady, H. F. Mohamed, S. H. Hafez, A. M. Fahmy, A. Magdy, A. S. Mohamed, E. O. Ali, H. R. Abdelhamed, O. M. Mahmoud, *Sci. Rep.* **2023**, *13*, 6435.
[3] P. Dh, S. Ramasamy, S. Angaiah, *ES Energy Environ* **2023**, *20*, 902.
[4] M. Xu, M. Wei, *Adv. Funct. Mater.* **2018**, *28*, 1802943.
[5] D. Yu, F. Wu, J. He, L. Bai, Y. Zheng, Z. Wang, Zhang, *Appl. Catal. B: Environ.* **2023**, *320*, 121880.
[6] M. Duan, S. Liu, Q. Jiang, X. Guo, J. Zhang, S. Xiong, *Chin. Chem. Lett.* **2022**, *33*, 4428–4436.
[7] Z. P. Xu, J. Zhang, M. O. Adebajo, H. Zhang, C. Zhou, *Appl. Clay Sci.* **2011**, *53*, 139–150.
[8] N. Kamali, J. B. Ghasemi, G. M. Ziarani, S. Moradian, A. Badiei, *Chin. J. Chem. Eng.* **2023**, *53*, 374–380.
[9] J. Babamoradi, R. Ghorbani-Vaghei, S. Alavinia, *Int. J. Biol. Macromol.* **2022**, *209*, 1542–1552.
[10] S. Rohani, G. M. Ziarani, A. Ziarati, A. Badiei, *Appl. Surf. Sci.* **2019**, *496*, 143599.
[11] R. Ghiai, S. Alavinia, R. Ghorbani-Vaghei, A. Khazaei, R. Karimi-Nami, I. Karakaya, *J. Mater. Chem. A* **2024**, *12*, 5474–5492.
[12] E. Nope, A. G. Sathicq, J. J. Martínez, H. Rojas, G. Romanelli, *Catal. Today* **2021**, *372*, 126–135.
[13] M. Dib, M. Kacem, S. Talbi, H. Ouchetto, K. Ouchetto, A. Essoumhi, A. Hafid, M. Khouili, *Curr. Chem. Biol.* **2023**, *17*, 160–169.
[14] K. E. Farouki, M. Kacem, M. Dib, H. Ouchetto, A. Hafid, M. Khouili, *Curr. Organocat.* **2024**, *11*, 154–174.
[15] P. Padmaja, B. S. Reddy, V. G. Ugale, P. N. Reddy, *Curr. Org. Synth.* **2022**, *19*, 591–615.
[16] S. He, Z. An, M. Wei, D. G. Evans, X. Duan, *Chem. Commun.* **2013**, *49*, 5912–5920.
[17] J. Pathak, P. Singh, *Appl. Organomet. Chem.* **2024**, *38*, e7507.
[18] L. Reguera, D. G. Rivera, *Chem. Rev.* **2019**, *119*, 9836–9860.
[19] A. Dómling, W. Wang, K. Wang, *Chem. Rev.* **2012**, *112*, 3083.
[20] S. K. Biswas, D. Das, *Mini-Rev. Org. Chem.* **2022**, *19*, 552–568.
[21] R. L. Mohlala, E. M. Coyanis, *Phys. Sci. Rev.* **2024**, *9*, 3529–3545.
[22] C. P. Pandhurnekar, A. J. Mungole, H. C. Pandhurnekar, B. G. Yadao, *J. Adv. Sci. Res.* **2021**, *2021(ICITNAS)*, 37–43.
[23] S. Mortada, K. Karrouchi, E. H. Hamza, A. Oulmidi, M. A. Bhat, H. Mamad, Y. Aalilou, S. Radi, M. H. Ansar, A. Masrar, M. E. A. Faouzi, *Sci. Rep.* **2024**, *14*, 1312.
[24] A. Ansari, A. Ali, M. Asif, *New J. Chem.* **2017**, *41*, 16–41.
[25] R. A. Mohamed-Ezzat, M. A. Omar, A. Temirak, A. S. Abdelsamie, M. M. Abdel-Aziz, S. A. Galal, G. H. Elgemeie, H. I. El Diwani, K. J. Flanagan, M. O. Senge, *J. Mol. Struct.* **2024**, *1311*, 138415.
[26] H. Kumar, K. K. Bansal, A. Goyal, *Anti-Infect. Agents* **2020**, *18*, 207–223.
[27] V. N. Badavath, V. Jayaprakash, *Front. Drug Des. Dis.* **2021**, *10*, 47–91.
[28] A. Khandebharad, S. Sarda, M. Soni, C. Gill, P. Kulkarni, B. Agrawal, *Curr. Chem. Lett.* **2018**, *7*, 57–64.

© 2025 Wiley-VCH GmbH

23656549, 2025, 13, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/slct.202404835 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

[29] S. Amirnejat, A. Nosrati, S. Javanshir, *Appl. Organomet. Chem.* **2020**, *34*, e5888.
[30] A. S. Patki, K. N. Patil, S. Kusuma, D. B. Muley, A. H. Jadhav, *Res. Chem. Intermed.* **2021**, *47*, 2751–2773.
[31] F. S. Pashaki, M. Nikpassand, *Front. Chem.* **2021**, *9*, 724745.
[32] M. A. E. A. A. Ali El-Remaily, T. El-Dabea, M. Alsawat, M. H. Mahmoud, A. A. Alfi, N. El-Metwaly, A. M. Abu-Dief, *ACS Omega* **2021**, *6*, 21071–21086.
[33] S. Kumari, A. Shekhar, D. D. Pathak, *New J. Chem.* **2016**, *40*, 5053–5060.
[34] R. G. Jacob, D. H. D. Oliveira, T. J. Peglow, J. E. Nascimento, R. H. Bartz, *J. Braz. Chem. Soc.* **2019**, *30*, 2144–2152.
[35] G. Mali, B. A. Shaikh, S. Garg, A. Kumar, S. Bhattacharyya, R. D. Erande, A. V. Chate, *ACS Omega* **2021**, *6*, 30734–30742.
[36] J. Chen, D. Chen, J. Kuang, Y. Ma, *J. Org. Chem.* **2021**, *86*, 9289–9298.
[37] G. Fan, F. Li, D. G. Evans, X. Duan, *Chem. Soc. Rev.* **2014**, *43*, 7040–7066.
[38] K. Madasamy, S. Kumaraguru, V. Sankar, S. Mannathan, M. Kathiresan, *New J. Chem.* **2019**, *43*, 3793–3800.
[39] B. Pandey, P. Singh, *Appl. Organomet. Chem.* **2023**, *37*, e7072.
[40] B. Pandey, P. Singh, V. Kumar, *Bull. Mater. Sci.* **2022**, *45*, 169.
[41] P. Meenakshi, U. Sitharaman, N. Rajamani, *J. Rare Earths.* **2017**, *35*, 474–479.
[42] G. Rathee, S. Kohli, S. Panchal, N. Singh, A. Awasthi, S. Singh, S. Hooda, R. Chandra, *ACS Omega* **2020**, *5*, 23967–23974.
[43] J. Pathak, P. Singh, *Polym. Compos.* **2022**, *43*, 7836–7844.
[44] D. Mishra, R. Singh, C. Rout, *J. Chem. Pharm. Res.* **2017**, *9*, 16–19.
[45] M. Srivastava, P. Rai, J. Singh, J. Singh, *New J. Chem.* **2014**, *38*, 302–307.
[46] B. K. Patel, S. D. Patel, *J. Chem. Pharm. Res.* **2014**, *6*, 364–368.
[47] H. Kiyani, M. Bamdad, *Res. Chem. Intermed.* **2018**, *44*, 2761–2778.
[48] N. Yadav, R. Singh, Y. K. Tyagi, *Org. Prep. Proced. Int.* **2024**, *56*, 342–349.
[49] P. Arora, J. K. Rajput, *J. Mater. Sci.* **2017**, *52*, 11413–11427.

Manuscript received: October 17, 2024

© 2025 Wiley-VCH GmbH

23656549, 2025, 13, Downloaded from https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/slct.202404835 by Delhi Technological University, Wiley Online Library on [06/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License