# SCHOLARLY PUBLICATIONS

*A CURRENT AWARENESS BULLETIN*
*OF RESEARCH OUTPUT*

## *@DTU*

(145th Edition)

***JANUARY 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 Fifth** 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 **January, 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 Copper Sulfide Doped 2D-MXene-Based Ultrasensitive Label-Free Electrochemical Immunosensor for EpCAM Antigen Detection ,***6.Sweety, 6.1.Saroj Paneru , 3.Devendra Kumar,*** Applied chemistry, DTU

2. A Review on Effects of Waste Tire Rubber Polymer on Fresh, Mechanical, and Impact Resistance Properties of Rubberized Concrete , Rahul Kumar*, **3.Nirendra Dev***, Civil Engineering, DTU

3. A review on Kerr optical frequency combs ***,8.1. Parth Dipendra , Ashish Kumar Satya, 3. A. S. Rao,*** Applied physics ,DTU

4. A two-stage ensemble approach for the analysis of Parkinson's Disease using speech signals, ***6. Kavita Bhatt, 3.N Jayanthi and 3. Manjeet Kumar***, E&C, DTU

5. A Novel Hybrid Model for Efficient Prediction of Diabetes Using Machine Learning, ***8. Akshit Jain, 6.Shagun Jain***, Production and Industrial & Software Engg., DTU

6. Advancements in and multi-objective transportation problems: techniques and comparative analysis, ***6.Rajpal Rajbhar, 3. Laxminarayan Das***, Mathematics, DTU

7. Aerodynamic Influences on Structures with Varied Plan Configurations, Rahul Kumar Meena, ***6. Abhishek Prakash Paswan, 6. Prateek Roshan*** and Manvendra Verma, Civil Engg. DTU

8. Air Quality Analysis in Desert Region in the Northern State of India: GIS Based Approach, ***6. Vignesh Mohan, 3. Rajeev Kumar Mishra*** and Vijay Kumar Soni, Environmental Engg. DTU

9. Analysing the performance of heptagonal photonic crystal fibre in terahertz regime, **8.1. Gurmeet Singh, 6.Shubham Sharma, 3. Ajeet Kumar,** Applied Physics, DTU

10. ANFIS-Based Resonant Controller for Mitigating Torque Ripples and Addressing Parametric Variation in PMSM-Driven Electric Vehicle, ***6.Shilpa Ranjan 3. Madhusudan Singh and 3.Mini Sreejeth***, Electrical Engg. DTU

11. Biodegradable Polymeric Film for Food Packaging, 3. Raminder Kaur, 3.Neelam Yadav, Applied Chemistry, DTU

12. CNN Based System in Vehicle for different Road Surface and Lane Markings, ***8.1 Kukku Yaduvanshi, 3.Narendra Kumar, 8.1 Parikshit Dahiya***, ***3. S Indu***, and ***8.1 Sangamithr Chembil Sudheer***, Electrical, ECE, DTU

13. Congestion Control Scheme for Link Quality Improvement Using Bio-Inspired Fuzzy Inference System–Based CoAP in IoT, ***6.Raveena Yadav, 3.Vinod Kumar***, CSE, DTU

14. Design and analysis of SRAM cell using swing restoration inverter for low power applications, **3. Ram Murti Rawat, 3. Vinod Kumar**, Computer Science and engineering, DTU

15. Design and Analysis of Twin-Core Side-Polished SPR PCF Sensor for Preliminary Malaria Detection, ***6.Devi Kangan B S, 6.Akash Khamaru 3. Ajeet Kumar***, Applied Physics, DTU

16. Design and Implementation of an IoT-Based Patient Monitoring System for Comprehensive Vital Tracking and Secure Data Access, Vihaan Ravishankar, Mohammad Ibrahim, Anjali Jain and ***6.1Neelam Verma,*** Electrical Engg., DTU

17. Effi-CNN: real-time vision-based system for interpretation of sign language using CNN and transfer learning, Pranav, ***3.Rahul Katarya, DTU***

18. Empirical Analysis of Nature-Inspired Algorithms forAutism Spectrum Disorder Detection Using 3D Video Dataset, Aneesh Panchal, ***6.Kainat Khan*** and ***3.Rahul Katarya,*** CSE, DTU

19. Enhanced Red Luminescence in Dy3+/Eu3+ Co-doped Calcium Alumino Borosilicate Glasses for w-LED Applications, ***6.Videsh Kumar, 6.Sheetal Kumari, 6.Shristy Malik***, Ravita ,Tripti Sharma, Aman Prasad, and ***3.A. S. Rao***, Applied Physics, DTU

20. Experimental investigation of Homogeneous Charge Compression Ignition (HCCI) engine with vaporized fuel injection, ***3.Naveen Kumar*** Pawan Kishore Jha, and ***3.R. C. Singh***, Mechanical Engg., DTU

21. Highly coherent on-chip mid-infrared supercontinuum generation from 1.5 to 25 μm in CdTe-based Rib waveguide ,***6.Deepak Garg, 3. Ajeet Kumar***,Applied Physics DTU

22. Highly Sensitive Black Phosphorus–Layered SPR-PCF Refractometric Sensor, ***6.Plakshi Gupta, 6.Akash Khamaru*** and ***3.Ajeet Kumar***, Applied Physics, DTU

23. Inclusive and supportive aging: mental health and social care strategies for India's elderly **6.Neha Narula** ,USME, DTU

24. Influence of Magnesium Content on Tensile Properties and Deformation Behaviour of Al–Mg–Si Alloy Sheets, 3***.Rakesh Kumar, 3.Vijay Gautam***,Mechanical Engg. DTU

25. Insights into Temperature sensitivity analysis of polarity controlled charge plasma based Tunable Arsenide/Antimonide Tunneling Interfaced Junctionless TFET, S. Sharma, J. Madan, and ***3. R. Chaujar***, Applied Physics, DTU

26. Modeling the enablers for sustainable adoption of water ionizers: an ISM-MICMAC approach, ***6.Saurabh Pandey, 3.Deepti Aggrawal***, USME, DTU

27. Network Partitioning Problem and UAVs' Integration for Efficient Connectivity Restoration: A Systematic Review, Aditi Zear, ***3.Virender Ranga*** , Kriti Bhushan, IT,DTU

28. Nonlinear effects associated with Alfvén ion cyclotron wave,and turbulence generation in presence of magnetic islands in solar wind plasmas***, 6. Ritu, 6. Jyoti, 3.Suresh C. Sharma***, R. P. Sharma, applied Physics, DTU

29. Object Shape Classification Using 4x4 Ultrasonic Sensor Array, ***6. Sambhav Jain***, Kuldeep Singh and ***3. Qasim Murtaza,*** Mechanical Engg., DTU

30. PATANLS: Performance Analysis of Trust-Aware Node Localization in WSNs***, 6.Yogendra Mohan, 3. Rajesh Kumar Yadav***, Manisha Manjul, CSE, DTU

31. Performance Appraisal of a Domesticated Potato Solar Dryer: An Experimental Study, ***6. Shimpy, 6. Mahesh Kumar, 3. Anil Kumar***, Mechanical, DTU

32. Performance of Aircraft with varying cant angles and sweep
    Angles, ***8.1 Aman Jha***, ***8.1 Aman Ayush***, ***6. Arun Kumar*** and ***3. M. Zunaid,*** Mechanical, DTU

33. Performance optimization of lead-free KGeCl3 based perovskite solar cells using SCAPS-1D, ***6. Rahul Kundara, 3. Sarita Baghel***, Applied Physics DTU

34. Polyvinyl Alcohol: Recent Advances and Applications in Sustainable Material, Teena Saini, ***6.1 Jagram Meena***, Vratika Verma, Sakshi Saini & Rajdeep Malik, Applied Chemistry, DTU

35. Real-Time Classification of Leukocytes Using Deep Learning in Microscopic Imaging, Sagar Bharat Shah, Siddharth Agarwal, Rohan Bagulwar, ***8.1Yash Sant,*** Aditya Dilip, Ayush Singh, Pujitha Bobbili and Dhanshri Ahir, Software Engg. DTU

36. Recent developments in solar energy–based systems for essential oil extraction from aromatic and medicinal plants: an overview, ***6. Gourav Dhingra*** and ***3. Anil Kumar***, Mechanical Engg., DTU

37. Sea sonal Vari a tion of Aero sol Dis tri bu tion and Optical Depth over the Indo-Gangetic Belt, B. Kumara, B. P. Singh, ***3. P. Chandra***, Yadavb, V. Singhb, and J. Masih, Biotechnology, DTU

38. Signal Noise Reduction Algorithm for Object Localization and Tracking using a 2-Dimensional Ultrasonic Sensor Array, ***6.Sambhav Jain***, Kuldeep Singh, Mechanical, DTU

39. Small kernelled deep capsule networks for handwritten Ge'ez numerals recognition, ***3.Ruchika Malhotra, 6. Maru Tesfaye Addis***, SE & CSE, DTU

40. Software defect prediction based on multi-filter wrapper feature selection and deep neural network with attention mechanis, ***3. Ruchika Malhotra, 6.Sonali Chawla*** and Anjali Sharma, Software Engg., DTU

41. Stimulus Towards ESG in India and GHG Protocol with Measurable Metrological Tools: Progresses and Challenges, ***6. S. Chourasia*** and S. M. Pandey, Mechanical Engg., DTU

42. Strategically engineered defects in 2D-SnS for water splitting: Green energy solution for powering LEDs with hydroelectric cells, ***6.Hemant K. Arora***, Jyoti Shah , R.K. Kotnala, ***3.A.S. Rao and 3. Nitin K. Puri,*** Applied Physics, DTU

43. Supposititious Sarcasm Detection and Sentiment Analysis Coping Hindi Language in Social Networks Harnessing Zipf- Mandelbrot Probabilistic Optimisation and Perplexity Entropy Learning, ***6.HIMANI POKHRIYAL, 3.GOONJAN JAIN,*** Applied Mathematics, DTU

44. Sustainable material as a column filler in soft clay bed reinforced with encased column: numerical analysis, ***6.Srijan*** & ***3.A. K. Gupta,*** Civil Engg. DTU

45. T&D Circuit Segmentation for Improved Resilience—A Texas Utility Case Study, Sarat Chandra Vegunta, Nelson Bacalao, Giovani Sasso Scarpati, Azadeh Shadman and ***8.1 Ram Nath***, Mechanical Engg. DTU

46. THYMES: A Framework for Detecting Suicidal Ideation from Social Media Posts Using Hyperbolic Learning, Surendrabikram Thapa, Mohammad Salman, Siddhant Bikram Shah, ***8.1 Shuvam Shiwakoti***, Qi Zhang, Liang Hu, Imran Razzak, , Usman Nasee, SE, DTU

47. TRANSIT-EEG — A Framework for Cross-Subject Classification with Subject Specific Adaptation***, 6.Chirag Ahuja and 3. Divyashikha Sethia,*** CSE, DTU

48. Ultrasonic-Assisted Adsorptive Uptake of Malachite Green Dye Onto Zn/Cu-Layered Double Oxide: Spectrophotometric and Smartphone-Based Detection, Bhamini Pandey***, 3. Poonam Singh*** and Ravinder Kumar, Applied Chemistry, DTU

49. Unified dimensionality reduction techniques in chronic liver disease detection, ***8.Anand Karna, 8.Naina Khan, 8.Rahul Rauniyar*** and ***3. Prashant Giridhar Shambharka***r, Computer Engg. DTU

50. Unveiling the Impact of Interfacial Trap Charges on Strained VS-FeFinFETs for Improved Reliability:Device to Circuit level Assessment, ***6.Kajal Verma, 3.Rishu Chaujar***, Aplied Physics, DTU

51. Validation of a Parameter Design methodology for Air-core two-coil WPT system with multi-resonant LC compensation, ***6.Akshay, 6. Gaurav Yadav, 3. Mukhtiar Singh***, Electrical Engg. DTU

52. VSG-DC-Based Grid Forming Inverter Control for Standalone Microgrid, ***6.Deepak Kumar, 6.Jitendra Nath Rai*** and ***3.ST Nagarajan***, Electrical Engg. DTU

53. What Helps to Detect What? Explainable AI and Multi-Sensor Fusion for Semantic Segmentation of Simultaneous Crop and Land Cover Delineation Land Use, Saman Ebrahim, ***8.1 Saurav Kumar***, Envirnmental Engg., 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*

Check for updates

*Applied Organometallic Chemistry*

WILEY

Applied Organometallic Chemistry

RESEARCH ARTICLE

# A Copper Sulfide Doped 2D-MXene-Based Ultrasensitive Label-Free Electrochemical Immunosensor for EpCAM Antigen Detection

Sweety | Saroj Paneru | Devendra Kumar

Department of Applied Chemistry, Delhi Technological University, Delhi, India

**Correspondence:** Devendra Kumar (dkumar@dce.ac.in)

**Received:** 12 July 2024 | **Revised:** 28 October 2024 | **Accepted:** 24 December 2024

**Funding:** The authors received no specific funding for this work.

**Keywords:** CuS | electrochemical immunosensor | EpCAM antigen | MXene

## ABSTRACT

Herein, a novel CuS-anchored 2D-$Ti_3C_2T_x$ composite material was synthesized by a one-pot hydrothermal method. This CuS/$Ti_3C_2T_x$ hybrid was used to prepare a novel label-free electrochemical immunosensor for EpCAM antigen detection. The immunosensor has been fabricated by electrophoretic deposition of CuS/$Ti_3C_2T_x$ hybrid onto the ITO electrode followed by the immobilization of anti-EpCAM. The electrochemical biosensing studies of this fabricated immunosensor, BSA/anti-EpCAM/CuS/$Ti_3C_2T_x$@ITO demonstrated a wide linear range (0.01 fg/mL to 100 ng/mL) and low LOD (0.0057 fg/mL), together with high sensitivity (461.85 $\mu A\,fg^{-1}\,mL\,cm^{-2}$) for EpCAM antigen detection. The clinical validation of the proposed immunosensor was conducted with serum samples. Moreover, the immunosensor was found to be reproducible, specific, selective, and stable over a long period of 42 days.

## 1 | Introduction

MXenes and their composites are very advantageous materials owing to their excellent thermal and electrical conductivity, biocompatibility, low toxicity, high hydrophilicity, large surface area, and interlayer spacing in their sheets [1]. Additionally, these materials possess good ion and electron transport behavior because of tailorable hydroxyl, fluoride, and oxygen containing functional groups on their surfaces [2]. MXene is a 2D inorganic material, generally expressed as carbide/nitride/carbonitride of early transition metals in the periodic table. The name, MXene, is derived from its precursor, MAX phase ($M_{n+1}AX_n$), where M = Sc, Ti, Zr, Hf, V, Nb, Ta, Cr, Mo, W; A = group IIIA or IVA element; X = C and/or N; and $n = 1, 2, 3$ [3]. It is remarkably different from other 2D materials such as transition metal dichalcogenide, graphene, and graphitic carbon nitride; in respect of the bonding between different layers. These 2D materials have weak van der Waals force of attraction between their layers, whereas MXenes are bounded together by strong metallic bonds via 'A' layers in the MAX phase [4]. MXenes having the general formula, $M_{n+1}X_nT_x$, are obtained by the selective etching of "A" layers from the MAX phase, where $T_x$ represents the surface functional groups like −O, −OH, −F [5]. These hydrophilic moieties present on the surface provide superior water dispersibility to the MXenes [6]. These functional groups can be easily tailored by different surface functionalization methods like hydrothermal method [7], in situ polymerization [8], and heat and milling treatment [9]. MXenes are now considered to have unique properties and potential for their applications in different fields such as catalysis [10], energy storage [11], biomedicine [12], and sensing [13]. Furthermore, the MXenes have found applications in the environmental remediation because it is not only made up of non-toxic elements like Ti, C, and/or N, but their degradation byproducts ($CO_2$ and $N_2$) are also environment-friendly [14]. Nowadays, 2D-layered MXene-based electrodes have gained popularity as research hotspots in different fields. MXenes have

© 2025 John Wiley & Sons Ltd.

shown a great potential for the selective and sensitive detection of the cancer biomarkers in medical science [15].

Cancer biomarkers are either found in tumor cells or secreted by the tumor cells. The presence of these biomarkers above baseline indicates the existence of a tumor in the body [16]. Epithelial cell adhesion molecule (EpCAM) is one of the most common biomarkers, which confirms the tumor arising from the different parts of the body including head, lung, neck, and prostate [17]. EpCAM is a cell-to-cell adhesion membrane, and it works as a promoter in the cell proliferation. Therefore, EpCAM may be a good target owing to its overexpression on most of the cancer cells as well as on the tumor-initiating cells [18]. Thus, so far, several reports have been developed for the cancer cell analysis based on the immunological reaction between the EpCAM antigen and monoclonal antibody (anti-EpCAM). Microchip technology [19], fluorescence microscopy [20] and flow cytometry [21] are typical examples of the techniques used for the detection with varying degrees of effectiveness. However, the requirements for the labeled chemicals and expensive equipment limit the use of fluorescence microscopy and flow cytometry techniques. The limitations of microchip technology stem from the need for proficient personnel and intricate chip fabrication procedures. The quantitative detection of biomarkers is necessary to measure the risk appetite, early diagnosis and appropriate treatment, thereby improving the survival rate [22]. The use of electrochemical techniques has received a lot of attention now because of their inherent benefits, which include low cost, quick reaction, ease of downsizing, and compatibility with the microfabrication technology [23–26].

Electrochemical biosensor could be a reliable technique for the non-invasive detection of the biomarkers. These biosensors possess a tendency to measure the changes in the electrical parameters (i.e., potential, current, and impedance) due to the immunochemical interactions among the antibodies and their respective antigens [27]. Electrochemical biosensors based on MXene show a remarkable detection characteristics owing to their large surface area and drug-loading capacity [28]. The literature review reveals that there are very few reports for EpCAM detection based on MXene [29]. Recently, Zhangmin et al. developed a highly sensitive electrochemical biosensor using MXene and AuNPs for multiple exosome protein analysis [30]. In our previous work, we have also developed a MXene@ITO-based electrochemical immunosensor for the cancer biomarker detection [31]. Further, we have successfully modified the 2D-MXene with 2D-$TiO_2$ to improve the electrochemical parameters of the fabricated immunosensor, namely, BSA/anti-EpCAM/$TiO_2$/$Ti_3C_2T_x$@ITO for the detection of EpCAM antigen [32].

The most common drawback with MXene is that their nanosheets are not only easily oxidizable in the solutions but also tend to agglomerate, thereby affecting the electrochemical properties of these biosensors [33]. Therefore, it becomes very necessary to design a high-performance electrode material based on MXene to improve the different characteristics such as linear dynamic range, sensitivity, specificity, and adhesion of the analyte. MXene and its composite-based electrochemical biosensors are found to permit the effective immobilization of biomolecules in the proper environment to retain their bioactivity due to their unique structures [34].

Copper sulfide (CuS), a p-type semiconductor transition metal chalcogenide, is well-known for its high metal-like electrical conductivity, good chemical stability, and inexpensive [35]. Therefore, MXene-$Ti_3C_2$/CuS nanocomposite has been used for the colorimetric determination of cholesterol because of its high catalytic activity and synergetic effect [36]. By considering the high electrical conductivity and ion intercalation behavior of $Ti_3C_2T_x$, and high stability and catalytic power of the CuS, poses a synergistic effect in the hybrid of CuS and $Ti_3C_2T_x$ [37]. Nevertheless, to the best of our knowledge, there is no report based on $Ti_3C_2T_x$ with CuS filler as a biosensing platform used for the cancer biomarker detection. Therefore, it is interesting to study the electrochemical behavior of CuS/$Ti_3C_2T_x$ hybrid, as an electrode material for EpCAM detection.

This study focuses on the fabrication of an electrochemical biosensor based on CuS/$Ti_3C_2T_x$ hybrid for EpCAM antigen detection. Herein, CuS has been incorporated into the $Ti_3C_2T_x$ sheets by one-pot solvothermal method. The incorporation of CuS enhances the interlayer spacing in the sheets of $Ti_3C_2T_x$, thereby increasing the rate of electron transfer through the redox probe. The good anchoring of CuS onto the sheets further improves the specific surface area providing more space for the effective immobilization of antibodies. This hybrid is further used to fabricate an immunosensor (BSA/anti-EpCAM/CuS/$Ti_3C_2T_x$@ITO) for the detection of EpCAM antigen. This platform shows the high sensitivity and a broad linear range for EpCAM detection. Furthermore, the validation of the proposed immunosensor with spiked serum samples is paving the way to consider it a reliable alternative for measuring the EpCAM antigen in people.

## 2 | Materials and Methods

### 2.1 | Chemicals

Atlas antibodies supplied us the epithelial cell adhesion molecule antibody (anti-EpCAM). The following materials were obtained from the Sigma-Aldrich:Artificial human serum, EpCAM antigen, N-hydroxy succinimide (NHS), hydrochloric acid (HCl), titanium aluminum carbide ($Ti_3AlC_2$), sodium tetrafluoroborate ($NaBF_4$), sodium thiosulfate ($Na_2S_2O_3.5H_2O$), copper sulfate ($CuSO_4.5H_2O$), N-(3-dimethyl aminopropyl)-N′-ethyl carbodiimide hydrochloride (EDC), and bovine serum albumin (BSA). Thermo Fischer Scientific has provided all the analytical-grade buffer salts. All the solutions were prepared by using Milli-Q water prior to the experimental studies.

### 2.2 | Synthesis of $Ti_3C_2T_x$ and CuS/$Ti_3C_2T_x$ Hybrid

$Ti_3C_2T_x$ was prepared by little modification in the hydrothermal method as reported earlier [38]. On the other hand, the CuS/$Ti_3C_2T_x$ hybrid was synthesized for the first time by using the one-pot hydrothermal method. Approximately 150 mg of each $Ti_3C_2T_x$, $Na_2S_2O_3.5H_2O$ and $CuSO_4.5H_2O$ was dispersed in 10 mL of DIW separately. Subsequently, mix all the solutions under constant stirring and transfer them to a Teflon-lined autoclave for 12 h, maintained at 180°C. The mixture so obtained was then washed with DIW multiple times under centrifugation and then dried in a vacuum oven at 60°C.



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7975 by Delhi Technological University, Wiley Online Library on [20/01/2025]. 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 comparison purpose, the CuS was also synthesized by earlier reported method with a little modification [39]. About 0.8 g of sodium thiosulfate (0.125 M) and copper sulfate (0.5 M) was dissolved in 40 and 20 mL of DIW, respectively. After 30 min, both the solutions were mixed under stirring and transferred to an autoclave for 12 h at 180°C. The resulting dark-colored suspension was washed with DIW and ethanol under centrifugation, respectively. Subsequently, the precipitate was vacuum-dried at 70°C to obtain the powdered CuS.

## 2.3 | Electrophoretic Deposition of CuS/$Ti_3C_2T_x$ Hybrid on ITO Electrode

CuS/$Ti_3C_2T_x$ hybrid was deposited onto the glass substrate coated with indium-tin-oxide (ITO) by electrophoretic deposition technique consisting of two electrodes system. Before deposition, a colloidal solution was prepared with CuS/$Ti_3C_2T_x$ hybrid (0.1 mg/mL) in DIW. The electrophoretic deposition was then carried out for 18 s at an ideal DC potential of 10 V.

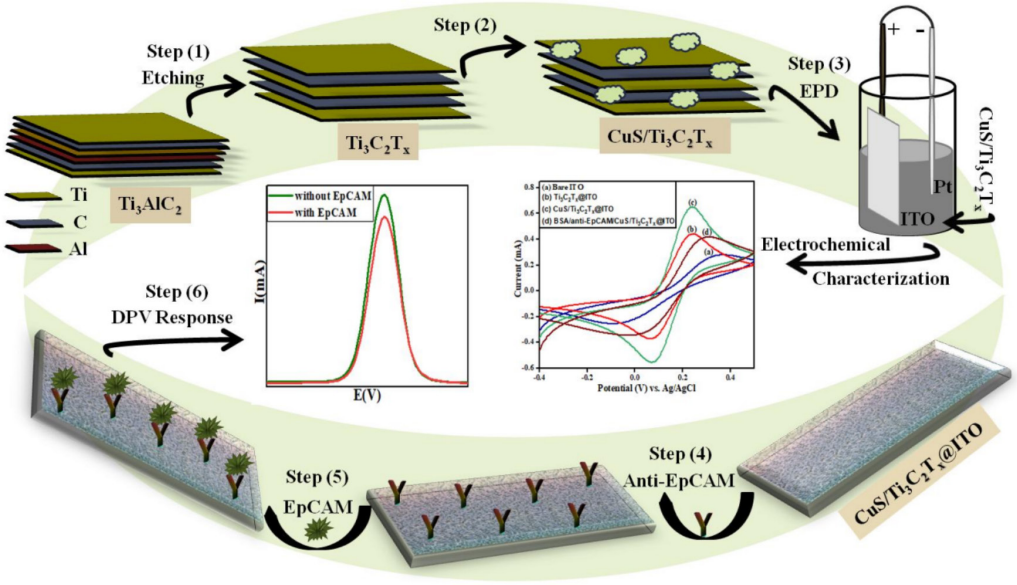



**FIGURE 1** | Schematic diagram for the synthesis of CuS/$Ti_3C_2T_x$ and fabrication of immunosensing platform for the detection of EpCAM antigen.

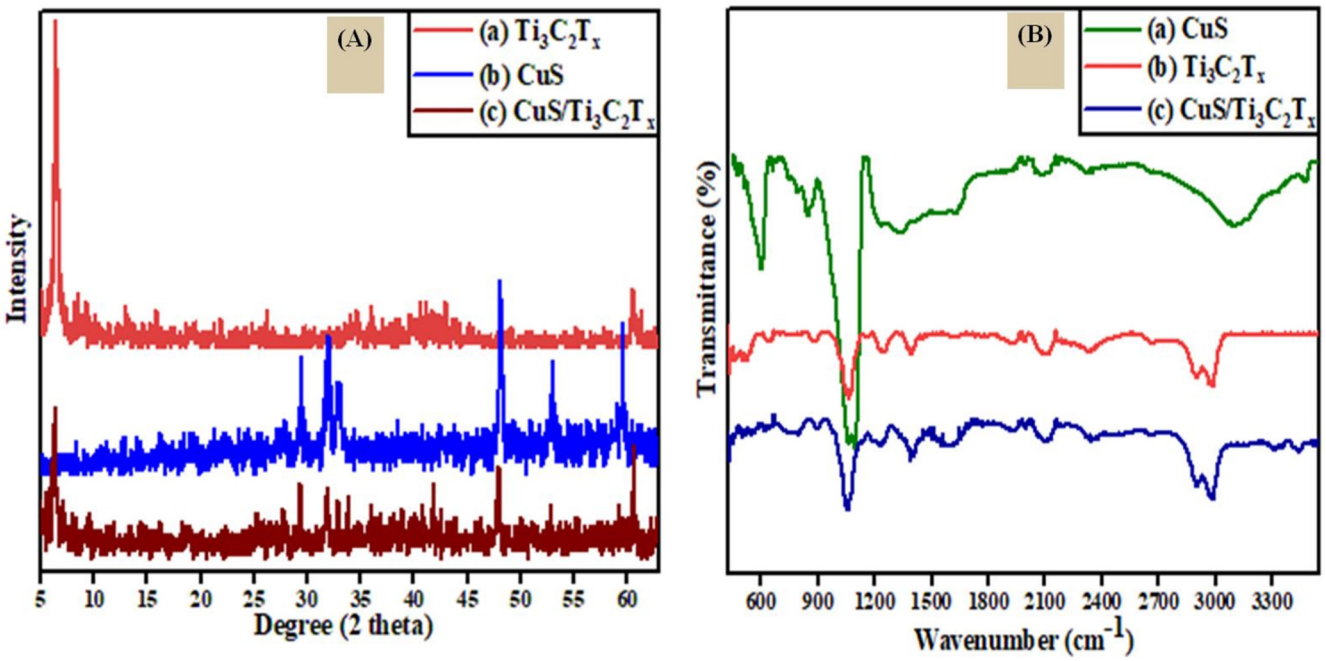



**FIGURE 2** | (A) XRD spectra of $Ti_3C_2T_x$ (a), CuS (b), and CuS/$Ti_3C_2T_x$ (c) and (B) FTIR spectra of CuS (a), $Ti_3C_2T_x$ (b), and CuS/$Ti_3C_2T_x$ (c).



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7975 by Delhi Technological University, Wiley Online Library on [20/01/2025]. 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

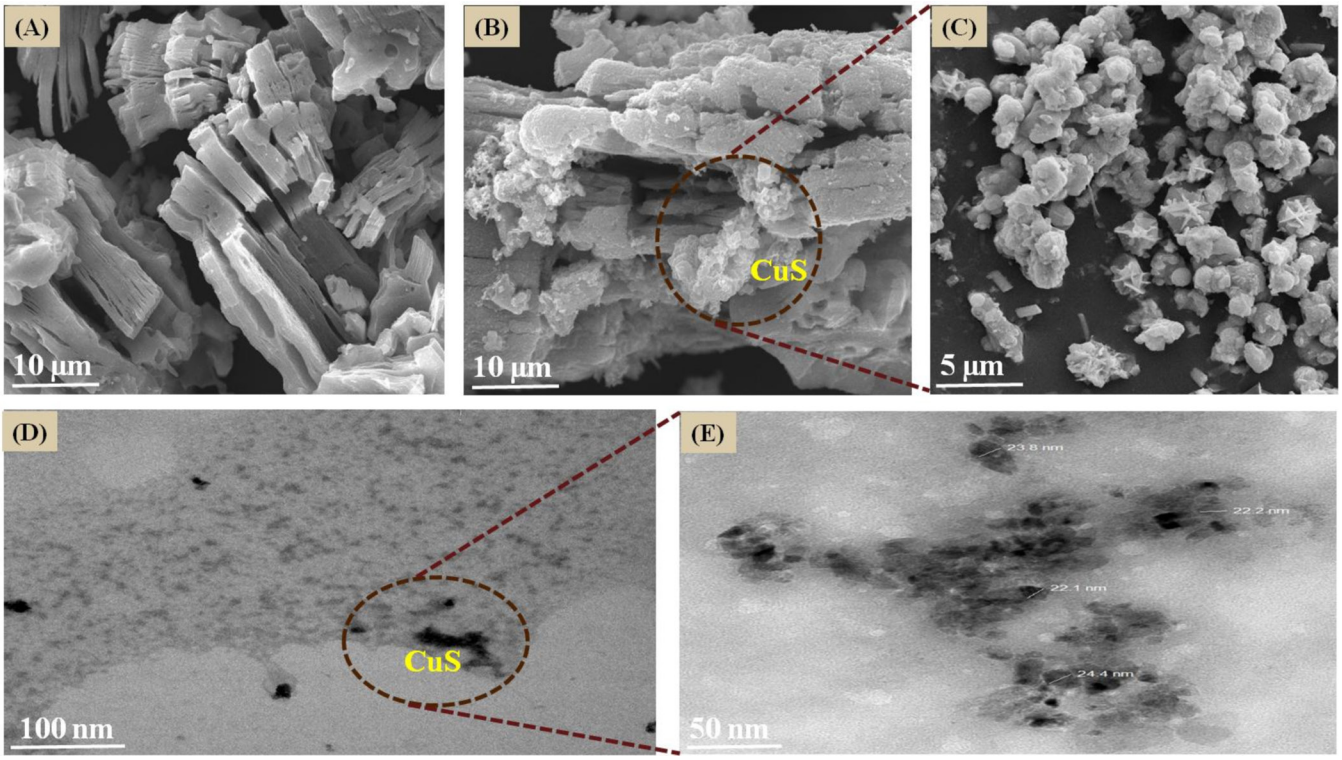



**FIGURE 3** | (A) SEM image of $Ti_3C_2T_x$, (B) and (C) SEM image of $CuS/Ti_3C_2T_x$, (D) and (E) TEM image of $CuS/Ti_3C_2T_x$.

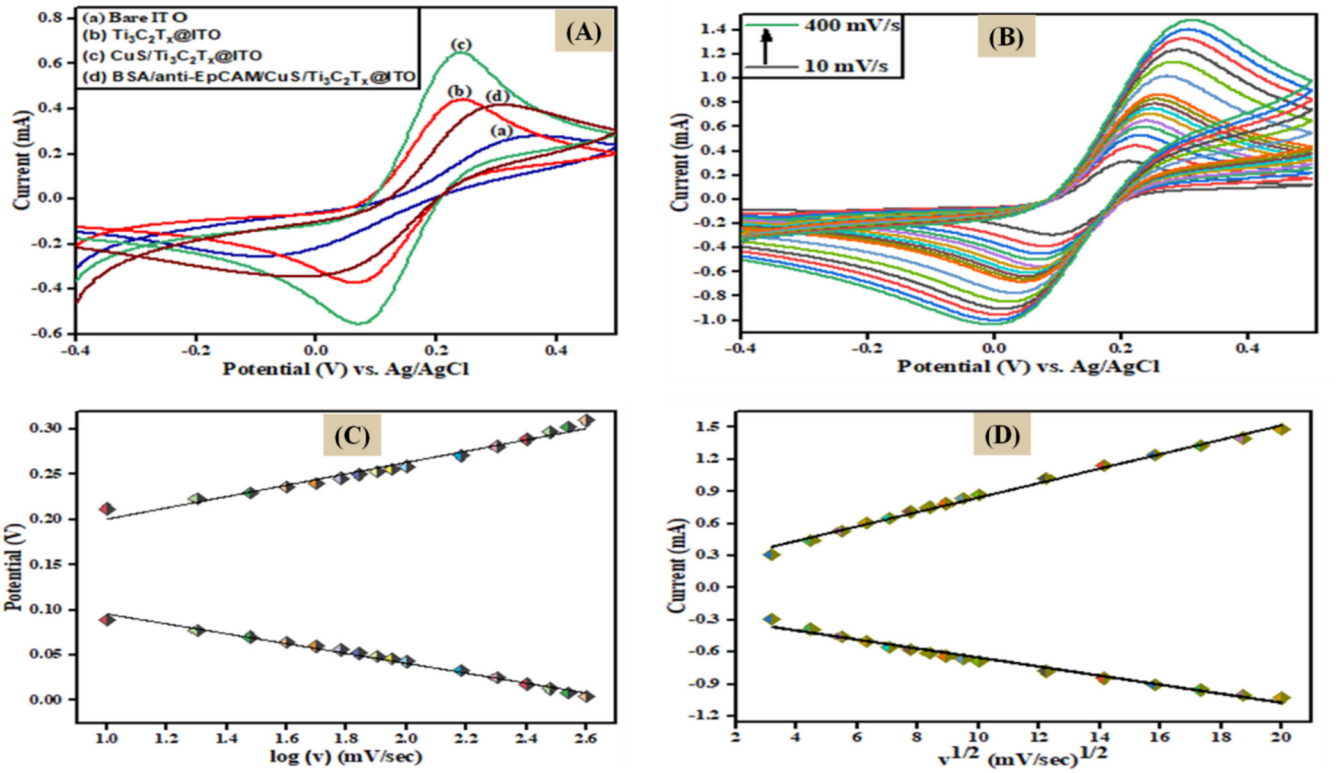



**FIGURE 4** | (A) CV plots of bare ITO (a), $Ti_3C_2T_x$@ITO (b), $CuS/Ti_3C_2T_x$@ITO (c), and BSA/anti-EpCAM/$CuS/Ti_3C_2T_x$@ITO (d); (B) scan rate (10–400 mV/s) study of $CuS/Ti_3C_2T_x$@ITO; (C) potential with log of scan rate for $CuS/Ti_3C_2T_x$@ITO; and (D) the plot of $I_{pa}$, $I_{pc}$ vs. square root of scan rate for $CuS/Ti_3C_2T_x$@ITO.



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7975 by Delhi Technological University, Wiley Online Library on [20/01/2025]. 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

A smooth film appeared on the surface of ITO resulting in the fabrication of CuS/$Ti_3C_2T_x$@ITO electrode.

### 2.4 | Fabrication of Biosensing Platform

Firstly, the CuS/$Ti_3C_2T_x$@ITO electrode was incubated with crosslinker EDC:NHS (20 μL) for 30 min. Subsequently, the fabricated electrode was drop-casted with anti-EpCAM (20 μL) and incubated for 8 h at 4°C. Here, EDC:NHS worked in a reverse manner and activated the carboxylic groups of the antibody, which further makes covalent bonds with the functional groups (−OH) of $Ti_3C_2T_x$. Further, to prevent the non-specific reactions, 1% BSA was applied to the immuno-electrodes for 30 min. Finally, the immuno-electrodes were washed with PBS (7.4 pH) three times for further use.

### 2.5 | Instrumentation

The X-ray diffraction (XRD) spectra were recorded by using a Bruker D-8 X-ray diffractometer. Fourier transforms infrared (FT-IR) spectroscopy was used to study the different types of bonds present in the synthesized materials by using the Perkin Elmer Spectrometer 2. The surface structures of materials were studied by the Scanning Electron Microscope (SEM, ZEISS EVO 18) and Transmission Electron Microscope (TEM, Talos L120C).

All the electrochemical characterization and biosensing studies were carried out by using the Autolab potentiostat/galvanostat (Netherland) in phosphate buffer saline (PBS, 0.2 M, pH = 7.4) having {5 mM $[Fe(CN)_6]^{3-/4-}$}.

## 3 | Results and Discussion

### 3.1 | Fabrication of Sensing Platform

The CuS-modified $Ti_3C_2T_x$ sheets were used as a transducer surface. Initially, the hybrid CuS/$Ti_3C_2T_x$ was synthesized by a one-pot hydrothermal method as shown in Figure 1, Steps 1 and 2. The synthesized hybrid was then electrophoretically deposited onto the ITO electrode and it is depicted in step 3 of Figure 1. The hybrid CuS/$Ti_3C_2T_x$ exhibits a synergistic effect owing to the high metal-like conductivity of $Ti_3C_2T_x$ and the high specific surface area of CuS. Further, the immunosensor was fabricated after immobilization of anti-EpCAM onto the surface of the CuS/$Ti_3C_2T_x$@ITO electrode (Figure 1, Step 4). Finally, the anti-EpCAM/CuS/$Ti_3C_2T_x$@ITO electrode was treated with BSA to block non-specific reactions with EpCAM antigen. This immunosensor was then utilized for the detection of EpCAM antigen (Figure 1, Step 5). A decrease in the current was observed by the electrochemical response study, which may be due to the combination of antibody and antigen onto the surface of immunosensor (Figure 1, Step 6).

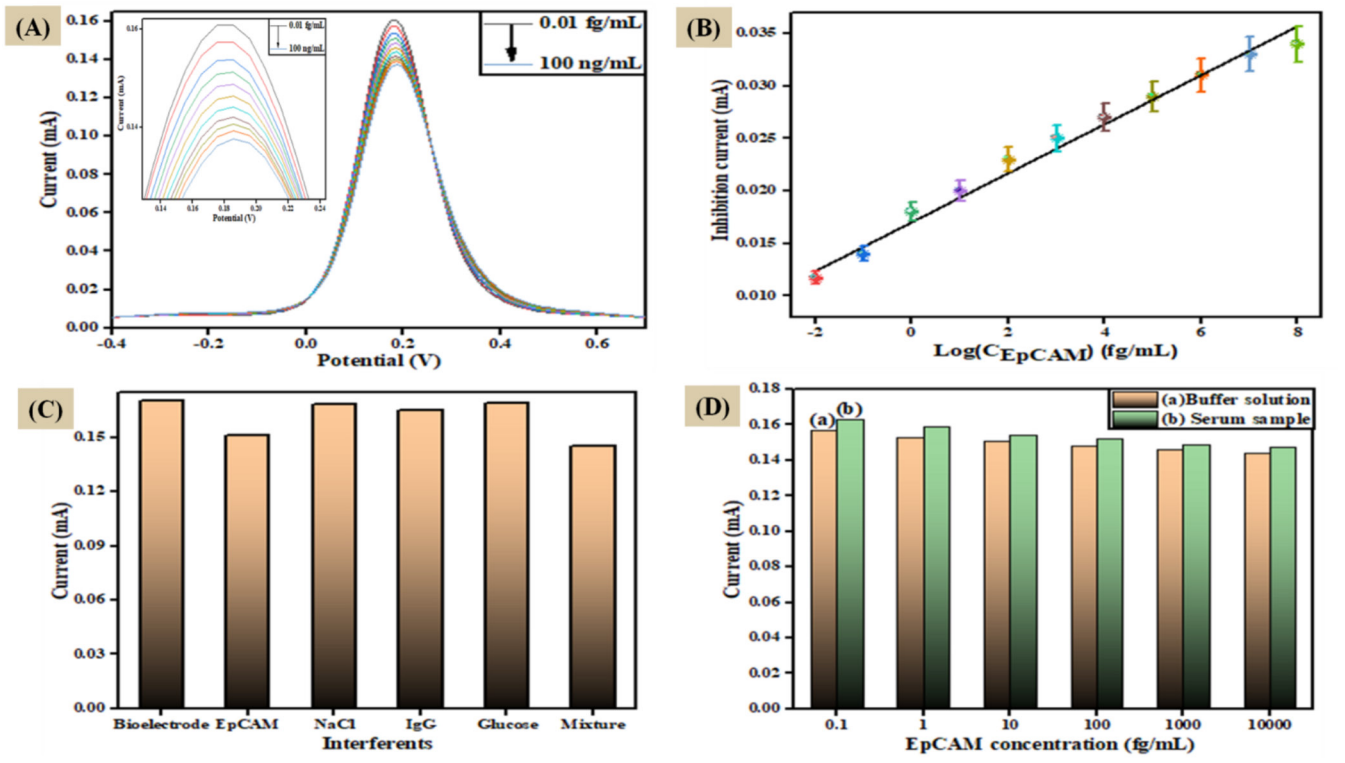



**FIGURE 5** | (A) DPV study showing the response of BSA/anti-EpCAM/CuS/$Ti_3C_2T_x$@ITO electrode with increasing concentration of EpCAM antigen (0.01 fg/mL to 100 ng/mL), (B) linearity plot between peak current change and log of EpCAM antigen concentration, (C) interference study of immunosensor with multiple analytes (100 fg/mL), and (D) comparison of DPV response (a) in buffer sample and (b) in spiked serum sample (0.1–10,000 fg/mL).



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7975 by Delhi Technological University, Wiley Online Library on [20/01/2025]. 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

### 3.2 | Structural Characterization

The XRD studies of $Ti_3C_2T_x$ (curve a), CuS (curve b) and CuS/$Ti_3C_2T_x$ (curve c) are represented in Figure 2A. The broad diffraction peak at 6.5° in the spectra of $Ti_3C_2T_x$ clearly shows the successful synthesis of $Ti_3C_2T_x$ [40]. The XRD curve of CuS shows the characteristic peaks located at 29.5°, 32°, 48.1°, 53°, and 59.5° (JCPDS: 06–0464) [41]. On the other hand, CuS/$Ti_3C_2T_x$ hybrid shows all the characteristic peaks of CuS at 29.3°, 31.9°, and 48° and the main peak of $Ti_3C_2T_x$ shifts from 6.5° to 6.3° indicating an increase in interlayer spacing, which further improves the effective surface area of transducer.

The chemical bonds present in the structure of CuS (curve a), $Ti_3C_2T_x$ (curve b), and CuS/$Ti_3C_2T_x$ hybrid (curve c) are demonstrated by FTIR spectroscopy (Figure 2B). The dominant band at 598 $cm^{-1}$ in FTIR spectra of CuS is attributed due to the stretching vibrations of Cu–S bond. The additional bands seen at 847 and 1084 $cm^{-1}$ are owing to C–H bending vibrations and C–O stretching vibrations of absorbed alcohol, respectively [42]. The FTIR spectrum of $Ti_3C_2T_x$ displays the C–H bonds stretching vibrations at 2980 and 2908 $cm^{-1}$. The vibrations at 1393 and 1059 and 530 $cm^{-1}$ are attributed to the O–H and C–F and Ti–C bonds stretching vibrations, respectively, which further proves the existence of surface groups, –OH and –F [43]. In the FTIR spectra of CuS/$Ti_3C_2T_x$ hybrid, the intensity of all the peaks decreases as shown in the Figure 2B. This intensity decrease is observed because CuS has been entrapped into the layers of $Ti_3C_2T_x$ sheets. The peak owing to the Cu–S stretching vibration (598 $cm^{-1}$) is not visible so clearly because of its combination with Ti–C stretching band (530 $cm^{-1}$) and broadening of peak occurs [35].

### 3.3 | Morphological Characterization

The surface structures of $Ti_3C_2T_x$ and CuS/$Ti_3C_2T_x$ hybrid were analyzed by using the SEM and TEM techniques. The SEM image in Figure 3A represents the 2D-layered accordion-like architecture of $Ti_3C_2T_x$. It is visible that the different layers are separated which are parallel to each other. In the SEM image of CuS/$Ti_3C_2T_x$ hybrid (Figure 3B), the sphere-like morphology of CuS is anchored onto the surface of $Ti_3C_2T_x$ sheets. The spherical structures of CuS in the magnified view (Figure 3C) clearly show a conductive network and space to CuS/$Ti_3C_2T_x$ hybrid for a better immobilization of antibodies. Figure 3D depicts the TEM image of CuS/$Ti_3C_2T_x$ hybrid, in which CuS particles are distributed onto the 2D-sheet of $Ti_3C_2T_x$. The size of spherically shaped CuS was found to be 22–25 nm (Figure 3E).

### 3.4 | Electrochemical Studies of Fabricated Electrodes

The different fabricated electrodes (a) bare ITO, (b) $Ti_3C_2T_x$@ITO, (c) CuS/$Ti_3C_2T_x$@ITO and (d) BSA/anti-EpCAM/CuS/$Ti_3C_2T_x$@ITO were electrochemically characterized in 7.4 pH PBS (0.2 M) containing 5 mM $[Fe(CN)_6]^{3-/4-}$ (Figure 4A) by the cyclic voltammetry (CV) analysis by scanning at 50 mV/s (scan rate [$\nu$]). It is observed that the electrochemical current becomes almost 1.5 times more after the incorporation of CuS, i.e., 0.65 mA for CuS/$Ti_3C_2T_x$@ITO in comparison to 0.442 mA

**TABLE 1** | Comparison of the performance of anti-EpCAM/CuS/$Ti_3C_2T_x$@ITO immunosensor with reported biosensors for EpCAM detection.

| S.No. | Detection method | Linear range | Detection limit | Sensitivity | Labeling | Stability | Ref |
|---|---|---|---|---|---|---|---|
| 1 | Electrochemical microfluidic (AgNPs-PVA) | 2–2000 pg/mL | 0.8 pg/mL | — | No | — | [45] |
| 2 | Differential pulse voltammetry (rGO@$TiO_2$/ITO) | 0.01–60 ng/mL | 6.5 pg/mL | 3.24 μA mL $ng^{-1}$ $cm^{-2}$ | No | 35 days | [46] |
| 3 | Square wave voltammetry (MCH/Hpl/AuE) | 0.1–20 ng/mL | 20 pg/mL | — | Yes | 15 days | [47] |
| 4 | Differential pulse voltammetry ($MoS_2$@rGO/ITO) | 0.001–20 ng/mL | 44.22 fg/mL | 0.104 mA mL $ng^{-1}$ $cm^{-2}$ | No | 45 days | [48] |
| 5 | Anodic stripping voltammetry (GRs–QDs) | 100 fg/mL to 7.5 μg/mL | 100 fg/mL | — | Yes | 21 days | [49] |
| **6** | **Differential pulse voltammetry (CuS/$Ti_3C_2T_x$@ITO)** | **0.01 fg/mL to 100 ng/mL** | **0.0057 fg/mL** | **461.85 μA $fg^{-1}$ mL $cm^{-2}$** | **No** | **42 days** | **Present work** |



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7975 by Delhi Technological University, Wiley Online Library on [20/01/2025]. 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 $Ti_3C_2T_x$@ITO. The electrochemical current decreases to 0.418 mA after immobilization of the antibody owing to the formation of an insulating layer of antigen and antibody onto the surface of transducer. The scan rate (10–400 mV/s) study of CuS/$Ti_3C_2T_x$@ITO electrode (Figure 4B) demonstrates that a linear trend is observed when we plot the peak potential vs. log $\nu$ graph, by using Equations (1) and (2) as shown in Figure 4C.

$$E_{pa}\left[CuS/Ti_3C_2T_x@ITO\right] = 0.0623\ V \log \nu + 0.138\ V;\ R^2 = 0.9703 \quad (1)$$

$$E_{pc}\left[CuS/Ti_3C_2T_x@ITO\right] = -0.0546\ V \log \nu + 0.151\ V;\ R^2 = 0.9871 \quad (2)$$

The values of electron transfer coefficient (α) and apparent charge transfer rate constant ($K_s$) for CuS/$Ti_3C_2T_x$@ITO electrode are calculated by using Laviron's equation [44] as 0.893 and 0.1337 $s^{-1}$, respectively.

Similarly, there is a linear relationship between peak currents and the square root of respective scan rates as depicted in Figure 4D by using Equations (3) and (4), which further demonstrates that the electron transfer is a surface-controlled process.

$$I_{pa}\left[CuS/Ti_3C_2T_x@ITO\right] = 67.53\ \mu A\ (s/mV)^{1/2}\ \nu^{1/2} + 166.5\ \mu A;\ R^2 = 0.9936 \quad (3)$$

$$I_{pc}\left[CuS/Ti_3C_2T_x@ITO\right] = -42.16\ \mu A\ (s/mV)^{1/2}\ \nu^{1/2} - 233.3\ \mu A;\ R^2 = 0.9801 \quad (4)$$

The Randles–Sevick equation ($I_p = 2.69 \times 10^5\ n^{3/2}ACD^{1/2}\nu^{1/2}$, where n is the number of electrons, C [mole/$cm^3$] is the concentration, $\nu$ [V/s] is the scan rate, and $I_p$ [A] is the peak current) and the slopes of Equations (3) and (4) were used to calculate diffusion coefficient (D), and effective surface area (A). The D and A values for CuS/$Ti_3C_2T_x$@ITO electrode are found to be 51.72 × $10^{-6}$ $cm^2/s$ and 0.22 $cm^2$, respectively. Furthermore, an average surface coverage area (τ) for CuS/$Ti_3C_2T_x$@ITO electrode is calculated to be 6.20 × $10^{-8}$ mol/$cm^2$ ($I_p = n^2F^2A\tau\nu/4RT$).

## 3.5 | Electrochemical Immunosensing Study for EpCAM Antigen Detection

The pH of the buffer solution has a remarkable role in the antigen–antibody viability, which further improves the analytical performance of the immunosensor. The pH of the electrolytic solution was taken to be 6.5, as optimized in our previous study [32]. The electrochemical performance of the BSA/anti-EpCAM/CuS/$Ti_3C_2T_x$@ITO electrode was checked by DPV technique, over a wide range of EpCAM antigen concentrations (0.01 fg/mL to 100 ng/mL). The DPV responses of the fabricated immunoelectrode were recorded by incubating different concentrations of EpCAM antigen, with an incubation time of 10 min. The peak current was found to be decreased as the concentration of EpCAM antigen increased (Figure 5A). There is a combination of antigen and antibody on the immunosensing electrode, resulting in the reduction of active sites, which in turn slows down the flow of electrons via the redox probe [27].

Figure 5B shows a calibration curve which has been plotted between the inhibition current (ΔI) and logarithmic scale of EpCAM antigen concentrations. The linearity plot follows the regression equation: (y = 2.336 x + 16.991, where y = current, x = logarithm of the concentration of EpCAM and $R^2$ = 0.9779). By using the slope of above equation, the sensitivity and limit of detection are being calculated as 461.85 μA $fg^{-1}$ mL $cm^{-2}$ and 0.0057 fg/mL, respectively. Various biosensing characteristics of the proposed immunosensor are being compared with reported 2D materials-based biosensors for EpCAM detection and the data are presented in Table 1. The results confirm that the proposed immunosensor outperforms the other biosensors with an extremely low detection limit. The broad linear range and high sensitivity of the fabricated immunosensor may be due to the synergistic effect of CuS and $Ti_3C_2T_x$.

## 3.6 | Selectivity, Reproducibility, and Shelf Life Studies

The interference study of BSA/anti-EpCAM/CuS/$Ti_3C_2T_x$@ITO immunosensor was studied by incubating different interfering species such as sodium chloride (NaCl), immunoglobulin G (IgG), and glucose (each concentration- 100 fg/mL). It is observed that the magnitude of peak current did not show much changes with all these species in comparison to EpCAM antigen (10 fg/mL) (Figure 5C). Furthermore, the specificity of the fabricated immunosensor was tested with a mixture of all the above mentioned analytes in addition to the EpCAM antigen. The electrochemical peak current is decreasing which confirms the specificity of the immunosensor.

The reproducibility of the fabricated immunosensor was examined by DPV technique with four different electrodes prepared under the similar conditions by incubating EpCAM

**TABLE 2** | EpCAM antigen detection in spiked serum samples ($n = 3$).

| S. no. | Approximate added EpCAM concentration (fg/mL) | Approximate recorded EpCAM concentration (fg/mL) | Recovery (%) | RSD (%) |
|---|---|---|---|---|
| 1 | 0.1 | 0.09 | 96.32 | 2.652 |
| 2 | 1 | 0.96 | 96.23 | 2.720 |
| 3 | 10 | 9.80 | 98.05 | 1.391 |
| 4 | 100 | 97.37 | 97.37 | 1.886 |
| 5 | 1000 | 979.90 | 97.99 | 1.438 |
| 6 | 10,000 | 9664.43 | 96.64 | 1.458 |



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7975 by Delhi Technological University, Wiley Online Library on [20/01/2025]. 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

antigen (10 fg/mL) (Figure S1A). A very low relative standard deviation (RSD) of 2.12% was determined with a constant surface area, thereby proving the excellent reproducibility of the immunosensor.

Finally, the shelf life of the fabricated immunosensor was tested by using the DPV technique. This analysis was done by incubating EpCAM antigen (10 fg/mL), every 6 days, for up to 42 days. It is observed that the peak current shows a minimal decrease even after 42 days, hence proving the acceptable stability of the immunosensor (Figure S1B).

### 3.7 | Human Serum Analysis

The clinical applicability of the fabricated immunosensor was analyzed with artificial human serum. The serum was spiked with various EpCAM antigen concentrations (0.1 fg/mL to 10,000 fg/mL), and the electrochemical response was recorded with the BSA/anti-EpCAM/CuS/$Ti_3C_2T_x$@ITO electrode (Figure S2). Figure 5D represents the bar diagram showing a good correlation between the (a) peak current in buffer solution and (b) peak current in spiked serum samples. These results are also presented in Table 2 showing a good recovery (96.23% to 98.05%) with %RSD less than 3%. Thus, these results confirm that the immunosensor could be a reliable biosensing platform for EpCAM detection.

## 4 | Conclusion

In the present work, we have developed a CuS/$Ti_3C_2T_x$ hybrid-based label-free electrochemical immunosensor for the detection of EpCAM antigen. This hybrid works as an excellent transducer surface due to the synergetic effect. The electrochemical study shows an incredibly low LOD—0.0057 fg/mL; for EpCAM antigen, with an ultra-broad detection range of 0.01 fg/mL to 100 ng/mL. Additionally, the fabricated immunosensor delivers an excellent reproducibility, high specificity and selectivity. Furthermore, the serum samples are being used to verify the applicability of the proposed immunosensor. In future, there is a room to modify the $Ti_3C_2T_x$ with other metal or metal oxides to further improve the biosensing results.

#### Author Contributions

*** Sweety:** conceptualization, methodology, software, formal analysis, validation, data curation, writing – original draft, investigation, visualization. **Saroj Paneru:** visualization, methodology, resources, formal analysis. **Devendra Kumar:** supervision, writing – review and editing, project administration.

#### Acknowledgments

Sweety acknowledges the Council of Scientific and Industrial Research (CSIR), Government of India, New Delhi for the financial support (File no: 08/133(0051)/2020-EMR-I).

#### Ethics Statement

Artificial human serum procured from Sigma Aldrich is used.

#### Consent

N.A.

#### Conflicts of Interest

The authors declare no conflicts of interest.

#### Data Availability Statement

Data can be made available on the basis of reasonable request due to the privacy/ethical restrictions.

#### References

1. M. Naguib, O. Mashtalir, J. Carle, et al., "Two-Dimensional Transition Metal Carbides," *ACS Nano* 6 (2012): 1322–1331.

2. S. Muff, J. Signer, and J. Fieberg, "Accounting for Individual-Specific Variation in Habitat-Selection Studies: Efficient Estimation of Mixed-Effects Models Using Bayesian or Frequentist Computation," *Journal of Animal Ecology* 89 (2020): 80–92.

3. I. Ihsanullah, "MXenes (Two-Dimensional Metal Carbides) as Emerging Nanomaterials for Water Purification: Progress, Challenges and Prospects," *Chemical Engineering Journal* 388 (2020): 124340.

4. M. Naguib, V. N. Mochalin, M. W. Barsoum, and Y. Gogotsi, "25th Anniversary Article: MXenes: A New Family of Two-Dimensional Materials," *Advanced Materials* 26 (2014): 992–1005.

5. M. W. Barsoum, "The $M_{N+1}AX_N$ Phases: A New Class of Solids: Thermodynamically Stable Nanolaminates," *Progress in Solid State Chemistry* 28 (2000): 201–281.

6. M. R. Ali, M. S. Bacchu, M. R. Al-Mamun, et al., "Recent Advanced in Mxene Research Toward Biosensor Development," *Critical Reviews in Analytical Chemistry* 54 (2022): 1–18.

7. Z. W. Gao, W. Zheng, and L. Y. S. Lee, "Highly Enhanced Pseudocapacitive Performance of Vanadium-Doped MXenes in Neutral Electrolytes," *Small* 15 (2019): 1902649.

8. M. Boota, B. Anasori, C. Voigt, M. Q. Zhao, M. W. Barsoum, and Y. Gogotsi, "Pseudocapacitive Electrodes Produced by Oxidant-Free Polymerization of Pyrrole Between the Layers of 2D Titanium Carbide (MXene)," *Advanced Materials* 28 (2016): 1517–1522.

9. W. Bao, C. E. Shuck, W. Zhang, X. Guo, Y. Gogotsi, and G. Wang, "Boosting Performance of Na–S Batteries Using Sulfur-Doped $Ti_3C_2T_x$ MXene Nanosheets With a Strong Affinity to Sodium Polysulfides," *ACS Nano* 13 (2019): 11500–11509.

10. W. Huang, L. Hu, Y. Tang, Z. Xie, and H. Zhang, "Recent Advances in Functional 2D MXene-Based Nanostructures for Next-Generation Devices," *Advanced Functional Materials* 30 (2020): 2005223.

11. X. Hui, X. Ge, R. Zhao, Z. Li, and L. Yin, "Interface Chemistry on MXene-Based Materials for Enhanced Energy Storage and Conversion Performance," *Advanced Functional Materials* 30 (2020): 2005190.

12. S. M. George and B. Kandasubramanian, "Advancements in MXene-Polymer Composites for Various Biomedical Applications," *Ceramics International* 46 (2020): 8522–8535.

13. H. L. Chia, C. C. Mayorga-Martinez, N. Antonatos, et al., "MXene Titanium Carbide-Based Biosensor: Strong Dependence of Exfoliation Method on Performance," *Analytical Chemistry* 92 (2020): 2452–2459.

14. Y. Zhang, L. Wang, N. Zhang, and Z. Zhou, "Adsorptive Environmental Applications of MXene Nanomaterials: A Review," *RSC Advances* 8 (2018): 19895–19905.

15. S. Kumar, Y. Lei, N. H. Alshareef, M. Quevedo-Lopez, and K. N. Salama, "Biofunctionalized two-Dimensional $Ti_3C_2$ MXenes for Ultrasensitive Detection of cancer Biomarker," *Biosensors and Bioelectronics* 121 (2018): 243–249.



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7975 by Delhi Technological University, Wiley Online Library on [20/01/2025]. 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

16. Q. Wu, N. Li, Y. Wang, et al., "A 2D Transition Metal Carbide MXene-Based SPR Biosensor for Ultrasensitive Carcinoembryonic Antigen Detection," *Biosensors and Bioelectronics* 144 (2019): 111697.

17. L. Zhu, B. Yang, K. Qian, L. Qiao, Y. Liu, and B. Liu, "Sensitive Electrochemical Aptasensor for Detecting EpCAM With Silica Nanoparticles and Quantum Dots for Signal Amplification," *Journal of Electroanalytical Chemistry* 856 (2020): 113655.

18. S. Imrich, M. Hachmeister, and O. Gires, "EpCAM and Its Potential Role in Tumor-Initiating Cells," *Cell Adhesion & Migration* 6 (2012): 30–38.

19. S. Nagrath, L. V. Sequist, S. Maheswaran, et al., "Isolation of Rare Circulating Tumour Cells in Cancer Patients by Microchip Technology," *Nature* 450 (2007): 1235–1239.

20. T. Ntouroupi, S. Ashraf, S. McGregor, et al., "Detection of Circulating Tumour Cells in Peripheral Blood With an Automated Scanning Fluorescence Microscope," *British Journal of Cancer* 99 (2008): 789–795.

21. T. Skirecki, G. Hoser, J. Kawiak, D. Dziedzic, and J. Domagała-Kulawik, "Flow Cytometric Analysis of CD133-and EpCAM-Positive Cells in the Peripheral Blood of Patients With Lung cancer," *Archivum Immunologiae et Therapiae Experimentalis* 62 (2014): 67–75.

22. C. L. Sawyers, "The Cancer Biomarker Problem," *Nature* 452 (2008): 548–552.

23. L. Qiu, L. Qiu, Z.-S. Wu, G. Shen, and R.-Q. Yu, "Cooperative Amplification-Based Electrochemical Sensor for the Zeptomole Detection of Nucleic Acids," *Analytical Chemistry* 85 (2013): 8225–8231.

24. S. Paneru and D. Kumar, "CuO@PEDOT:PSS-Grafted Paper-Based Electrochemical Biosensor for Paraoxon-Ethyl Detection," *Journal of Applied Electrochemistry* 53 (2023): 1–10.

25. S. Paneru and D. K. Sweety, "$CeO_2$ and PEDOT:PSS Modified Conducting Paper for Organophosphate Pesticide Detection," *Journal of Applied Electrochemistry* 54 (2024): 1–11.

26. S. Chatterjee, H. Singh, D. Hudda, and D. K. Sweety, "A Novel Acetylcholinesterase-Based Electrochemical Biosensor Using g-$C_3N_4$@$MoS_2$ Nanohybrid for the Detection of Trichlorfon," *Applied Organometallic Chemistry* 38 (2024): e7721.

27. W. Hong, S. Lee, E. J. Kim, M. Lee, and Y. Cho, "A Reusable Electrochemical Immunosensor Fabricated Using a Temperature-Responsive Polymer for cancer Biomarker Proteins," *Biosensors and Bioelectronics* 78 (2016): 181–186.

28. J. Yoon, M. Shin, J. Lim, J.-Y. Lee, and J.-W. Choi, "Recent Advances in MXene Nanocomposite-Based Biosensors," *Biosensors* 10 (2020): 185.

29. F. Kareem, M. Rizwan, and M. U. Ahmed, "A Novel Label-Free Electrochemical Immunosensor Based on DCNC@AgNPs/MXene for the Detection of Apolipoprotein A-1 in Human Serum," *Electrochimica Acta* 474 (2024): 143536.

30. Z. Wang, Z. Zhang, Y. Zhang, et al., "MXenes-Au NPs Modified Electrochemical Biosensor for Multiple Exosome Surface Proteins Analysis," *Talanta* 265 (2023): 124848.

31. Sweety and D. Kumar, "Development of $Ti_3C_2T_x$-Based Novel Immunosensor for cancer Biomarker Detection," *Applied Organometallic Chemistry* 38 (2024): e7570.

32. D. Kumar, "Electrochemical Immunosensor Based on Titanium Dioxide Grafted MXene for EpCAM Antigen Detection," *Journal of Colloid and Interface Science* 652 (2023): 549–556.

33. M. Zhang, Z. Shahnavaz, X. Yan, et al., "Mild Synthesis of Flower-Like CuS Modified MXene Nanocomposites for High-Performance Supercapacitors and Efficient Microwave Absorption," *Journal of Alloys and Compounds* 960 (2023): 170914.

34. G. Jamalipour Soufi, P. Iravani, A. Hekmatnia, E. Mostafavi, M. Khatami, and S. Iravani, "MXenes and MXene-Based Materials With Cancer Diagnostic Applications: Challenges and Opportunities," *Comments on Inorganic Chemistry* 42 (2022): 174–207.

35. S. Paneru and D. Kumar, "CuS Modified PEDOT: PSS Grafted Paper-Based Electrochemical Immunosensor for EpCAM Biomarker Detection," *Materials Chemistry and Physics* 313 (2024): 128687.

36. Y. Li, Z. Kang, L. Kong, et al., "MXene-$Ti_3C_2$/CuS Nanocomposites: Enhanced Peroxidase-Like Activity and Sensitive Colorimetric Cholesterol Detection," *Materials Science and Engineering: C* 104 (2019): 110000.

37. Z. Pan, F. Cao, X. Hu, and X. Ji, "A Facile Method for Synthesizing CuS Decorated $Ti_3C_2$ MXene With Enhanced Performance for Asymmetric Supercapacitors," *Journal of Materials Chemistry A* 7 (2019): 8984–8992.

38. C. Peng, P. Wei, X. Chen, et al., "A Hydrothermal Etching Route to Synthesis of 2D MXene ($Ti_3C_2$, $Nb_2C$): Enhanced Exfoliation and Improved Adsorption Performance," *Ceramics International* 44 (2018): 18886–18893.

39. J. Fang, P. Zhang, H. Chang, and X. Wang, "Hydrothermal Synthesis of Nanostructured CuS for Broadband Efficient Optical Absorption and High-Performance Photo-Thermal Conversion," *Solar Energy Materials and Solar Cells* 185 (2018): 456–463.

40. G. B. Tezel, K. Arole, D. E. Holta, M. Radovic, and M. J. Green, "Interparticle Interactions and Rheological Signatures of $Ti_3C_2T_z$ MXene Dispersions," *Journal of Colloid and Interface Science* 605 (2022): 120–128.

41. Z. Li, L. Mi, W. Chen, et al., "Three-Dimensional CuS Hierarchical Architectures as Recyclable Catalysts for Dye Decolorization," *CrystEngComm* 14 (2012): 3965–3971.

42. S. Riyaz, A. Parveen, and A. Azam, "Microstructural and Optical Properties of CuS Nanoparticles Prepared by Sol–Gel Route," *Perspectives in Science* 8 (2016): 632–635.

43. B. Reddy, G.-B. Cho, N. Reddy, et al., "Layered-Like Structure of $TiO_2$-$Ti_3C_2$ Mxene as an Efficient Sulfur Host for Room-Temperature Sodium-Sulfur Batteries," *Journal of Alloys and Compounds* 883 (2021): 160910.

44. K. Li, Y. Li, L. Wang, L. Yang, and B. Ye, "Study the Voltammetric Behavior of 10-Hydroxycamptothecin and Its Sensitive Determination at Electrochemically Reduced Graphene Oxide Modified Glassy Carbon Electrode," *Arabian Journal of Chemistry* 12 (2019): 2732–2739.

45. K. Bravo, F. G. Ortega, G. A. Messina, M. I. Sanz, M. A. Fernández-Baldo, and J. Raba, "Integrated bio-Affinity Nano-Platform Into a Microfluidic Immunosensor Based on Monoclonal Bispecific Trifunctional Antibodies for the Electrochemical Determination of Epithelial Cancer Biomarker," *Clinica Chimica Acta* 464 (2017): 64–71.

46. O. Jalil, C. M. Pandey, and D. Kumar, "Electrochemical Biosensor for the Epithelial Cancer Biomarker EpCAM Based on Reduced Graphene Oxide Modified With Nanostructured Titanium Dioxide," *Microchimica Acta* 187 (2020): 1–9.

47. Q. Chen, W. Hu, B. Shang, et al., "Ultrasensitive Amperometric Aptasensor for the Epithelial Cell Adhesion Molecule by Using Target-Driven Toehold-Mediated DNA Recycling Amplification," *Microchimica Acta* 185 (2018): 1–8.

48. O. Jalil, C. M. Pandey, and D. Kumar, "Highly Sensitive Electrochemical Detection of Cancer Biomarker Based on Anti-EpCAM Conjugated Molybdenum Disulfide Grafted Reduced Graphene Oxide Nanohybrid," *Bioelectrochemistry* 138 (2021): 107733.

49. M. J. Shiddiky, S. Rauf, P. H. Kithva, and M. Trau, "Graphene/Quantum Dot Bionanoconjugates as Signal Amplifiers in Stripping Voltammetric Detection of EpCAM Biomarkers," *Biosensors and Bioelectronics* 35 (2012): 251–257.

## Supporting Information

Additional supporting information can be found online in the Supporting Information section.



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7975 by Delhi Technological University, Wiley Online Library on [20/01/2025]. 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

STM JOURNALS

Journal of
Polymer & Composites

ISSN: 2321-2810 (Online)
ISSN: 2321-8525 (Print)
Volume 13, Special Issue 1, 2025
DOI (Journal): 10.37591/JoPC

https://journals.stmjournals.com/jopc

Research JoPC

# A Review on Effects of Waste Tire Rubber Polymer on Fresh, Mechanical, and Impact Resistance Properties of Rubberized Concrete

Rahul Kumar[1,*], Nirendra Dev [2]

***Abstract***

*According to the collective findings of numerous researchers, waste tire rubber has proven to be a valuable addition to concrete, contributing to advancements in construction. Rubber crumb is artificial rubber polymers, such as butadiene rubber (BR). The chemical composition of tire rubber crumb that is discarded can be represented by the combination of $(C_5H_8)n$, $(C_8H_8)x$-$(C_4H_6)y$, $(C_4H_6)n$, C, $S_8$, ZnO, and complicated hydrocarbons. Various studies have explored the mechanical and impact resistance characteristics of rubberized concrete when rubber crumbs partially replace traditional aggregates. The utility of rubber crumbs in construction is attributed to their favourable attributes, including excellent flexibility and a lower unit weight compared to natural rock aggregates. Consequently, rubberized concrete becomes particularly advantageous in scenarios requiring increased flexibility. The inherent ability of rubber tires to absorb impact energy and exhibit good ductility properties also positions it as a beneficial component, mitigating the structural damage caused by earthquakes. However, it is crucial to note that an excessive increase in the quantity of rubber can have detrimental effects on the mechanical properties of rubberized concrete. Notably, there is a gap in research focusing on the cement rubber matrix to enhance the mechanical properties of rubberized concrete, indicating the need for further studies in this area.*

**Keywords:** Rubber crumbs, rubberized concrete, composite material, mechanical properties, plasticizers, resistance to impact

## INTRODUCTION

The global expansion of the automotive industry has led to substantial environmental concerns regarding the disposal of waste tires. Tires have become a problematic waste source due to two main factors: high production volume and material durability. A tire can stay in a landfill for up to 100 years because it doesn't break down. The correct disposal of used tires has become an increasingly important environmental concern over time. The main components of waste tire rubber crumb are cross-linked polymer chains, specifically styrene-butadiene rubber (SBR) and polyisoprene (natural rubber). Additionally, it comprises a variety of chemicals utilized in the vulcanization process, including sulfur, carbon black, zinc oxide, and processing oils. Furthermore, there could be trace amounts of artificial rubber polymers, such as butadiene rubber (BR). The combination of $(C_5H_8)n$, $(C_8H_8)x$-$(C_4H_6)y$, $(C_4H_6)n$, C, $S_8$, ZnO, and complex hydrocarbons can be used to represent the chemical composition of discarded tire rubber crumb.

Over time, the proper disposal of used tires has grown to be a serious environmental problem.

***Author for Correspondence**
Rahul Kumar

[1]Assistant Professor, Civil Engineering Department, GLA University, Mathura, Uttar Pradesh, India
[2]Professor, Department of Civil Engineering, Delhi Technological University, Delhi, India

Received Date: April 23, 2024
Accepted Date: August 20, 2024
Published Date: December 18, 2024

**Citation**: Rahul Kumar, Nirendra Dev. A Review on Effects of Waste Tire Rubber polymer on Fresh, Mechanical, and Impact Resistance Properties of Rubberized Concrete. Journal of Polymer & Composites. 2025; 13(Special Issue 1): S469–S478p.

© STM Journals 2025. All Rights Reserved 

After being retired from service, rubber tires have little possibility of being recycled and wind up in landfills, degrading the environment. Utilizing rubber in concrete holds significant potential, offering a solution to prevent the area from turning into a landfill and promoting environmental benefits. To increase the impact resistance of Rubberized concrete (RuC), a specific amount of aggregates added to the concrete are replaced with rubber particles. RuC effectively reduces the black pollution caused by waste rubber. Earthquakes and explosions lead to concrete structures collapsing suddenly, causing significant damage to property and human life. Consequently, a new concrete material with strong toughness, blast resistance, and impact resistance is required. Because rubber is both elastic and durable, it can improve RuC's impact resistance [1]. Derived easily from used tires, rubber proves to be a cost-effective alternative to other concrete ingredients. Its incorporation leads to more economical concrete, given its superior flexibility and lower unit weight compared to traditional natural rock aggregates. Consequently, RuC becomes predominantly advantageous in regions demanding greater flexibility. In RuC, the key issue is that it loses mechanical characteristics when rubber crumbs (RC) are added to concrete. It also decreases both static and dynamic MOE, density, and slight workability(Figure.1) The hardness characteristic of the composite was enhanced by the addition of fiber to the rubber matrix [2–6].

Gaining rubber from old used tires is a cost-effective substitute related to other constituents usually used in concrete. Using RC in concrete can lead to a further cost-effective construction material. RC shows greater flexibility and a lower unit weight in comparison to the natural rock aggregates typically employed in traditional concrete. Consequently, the use of RuC becomes advantageous in situations demanding increased flexibility. In RuC, the main problem is that it loses mechanical properties. It also reduced the static and dynamic MOE, density, and slightly workability.

In conclusion, the intersection of waste tire disposal and construction materials, particularly the innovative use of rubberized concrete, represents a crucial step towards addressing environmental challenges associated with tire waste. This sustainable approach not only diverts rubber from landfills but also contributes to the development of more resilient and eco-friendly construction practices. Further research and widespread adoption of such practices hold the potential to usher in a new era of environmentally conscious construction methodologies.

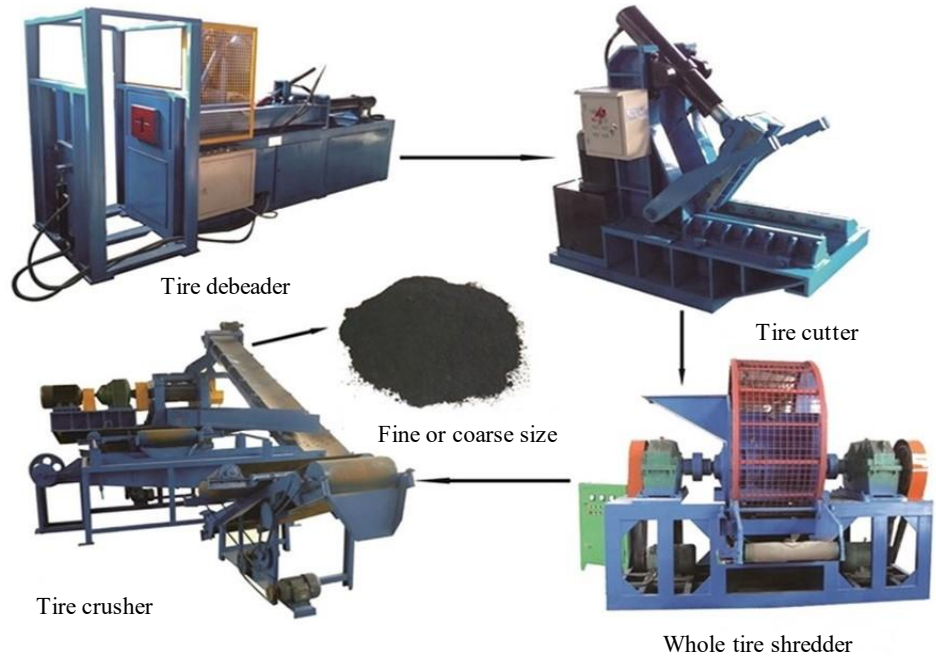
## WASTE RUBBER

Synthetic rubber, natural rubber, or a mix of synthetic and natural rubber are used to make vehicle tires. The tire compound was chosen based on its appropriateness for usage. Tire compositions vary according to vehicle type. Previous investigations have indicated that rubber has the largest proportion of ingredients, and carbon black is an important substance other than rubber [8]. Tire waste rubber is often shredded into various pieces by machines and reused in construction for various uses in order to limit trash generation in landfills and dumping operations, hence preventing harm to people and nature. Chipped rubber (CR), RC, and powder rubber (PR) are the three types of waste rubber that are being used in different studies [9]. The manufacturing process of rubber particles from rubber tyres is shown in Figure 2.



**Figure 1.** Waste tire rubber (7).

© STM Journals 2025. All Rights Reserved 




(a)



(b)

**Figure 2.** (a) Process of making fine and coarse rubber aggregate (10) (b) End product of tire rubber (11).

## PROPERTIES OF RUC

### Workability

Any waste material that is considered for use in concrete must be considered in terms of workability. Lack of concrete compaction may result in air voids, which will compromise the concrete's strength and lifespan. A proper w/c ratio should be adopted to reduce the influence of air voids and provide efficient concrete compaction [12,13]. As the rubber content increases, the workability of RuC tends to decline [8,14–17](Figure.3) The variance in the results was responsible for the differences in particle size. The workability of concrete is influenced significantly by both the size and quantity of rubber [18]. Additionally, when powdered rubber is combined with RC along with an appropriate admixture, favourable results are achieved, especially when compared to using chipped rubber. Due to RC's increased particle size, the overall result revealed a decline in workability, with the slump test showing a 50% drop in RuC compared to control concrete [15]. Treating rubber with a 10% NaOH solution for 120 minutes, followed by a water wash and drying, led to a decline in workability [19].

© STM Journals 2025. All Rights Reserved 

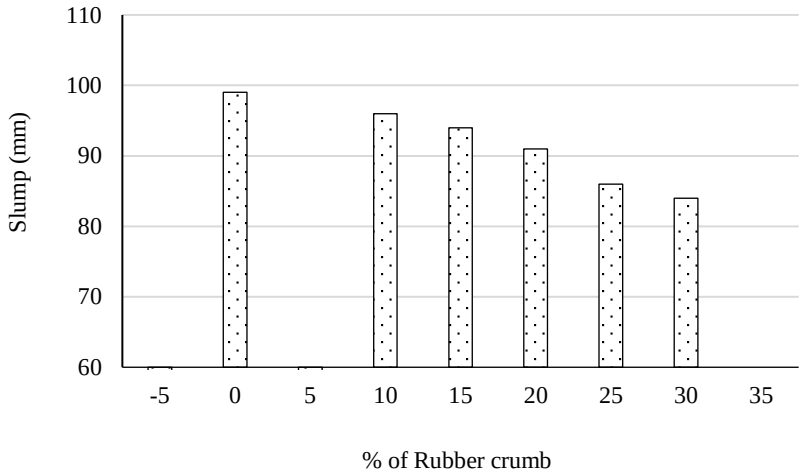



**Figure 3.** Workability of RuC (20).

**Mechanical Characteristics**

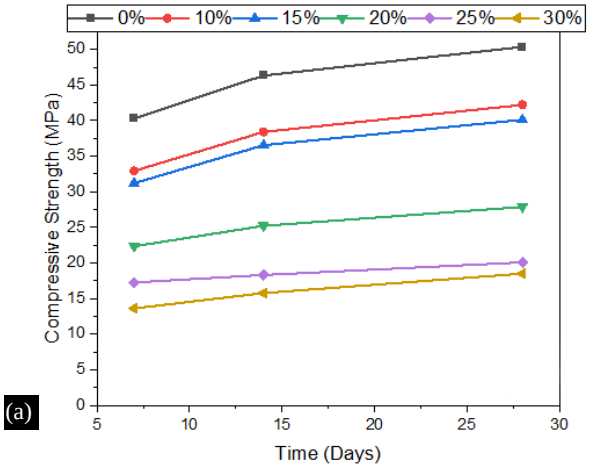
The mechanical characteristics are an essential parameter of the concrete, indicating whether it meets the minimum structural requirements or not. RC has been observed to contribute to the decrease in compressive strength of concrete in several previous studies. Reduction in mechanical properties of RuC depends upon the size and quantity of RC [21,22]. Many reviewers also concluded that replacing coarse aggregate (CA) with rubber chipped exhibits more loss in compressive strength values of RuC than fine aggregate (FA) replacement with RC. When comparing RC to chipped rubber, chipped rubber concrete loses more strength than crumb rubber concrete. The poor bond between the rubber and the other elements in the concrete is the cause of this loss in mechanical characteristics. This inadequate bonding results in a space and trapped air between the RC and the cement paste.

Increasing the rubber percentage by replacing CA or FA, both replacing methods may decrease the concrete strength. Rubber's capacity to serve as a partial substitute for FA is prompted by some research [22,23]. Researchers also discovered that compressive strength decreases when the amount of RC in concrete increases [21,24–26]. The quantity of RC which replaces FA should be restricted to 20%-30%, and the RC should also be surface treated. For RuC to be made, pozzolanic material is necessary since it improves both the mechanical properties of RuC and the bond between rubber and cement paste [25]. According to some studies, a 10% replacement of FA by RC provides the best mechanical characteristics [18]. However, other researchers claim that replacing FA with RC with less than 25% shows greater compressive strength ]27]. “The low RC stiffness, uneven distribution of the rubber aggregate in the concrete mix, lightweight, and weak bonding between the rubber aggregate and the cement paste all contributed to the mechanical characteristics decreasing trend” [18,28].

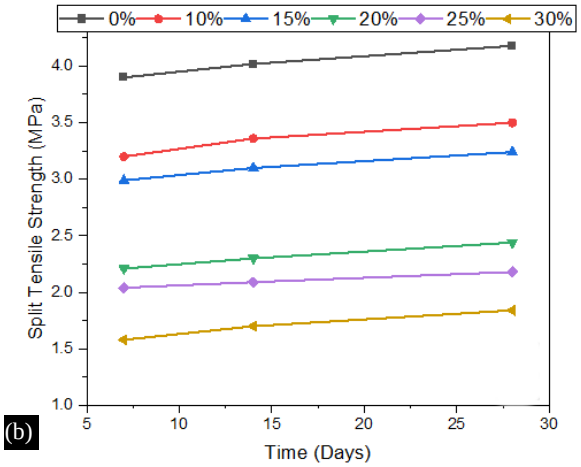
Pre-treating the surface of reinforced carbon (RC) enhances the compressive strength of RuC when incorporated into concrete (19). The mechanical properties of RuC can be improved by treating the RC surface with a solution of NaOH [19,29,30]. It was discovered that soaking RC in water for 24 hours improved the mechanical characteristics significantly [31]. Researchers concluded that soaking rubber particles improves the bonding properties of rubber with cement by removing air trapped in the particles. The influence of pre-treatment of RC by NaOH, $H_2SO_4$, $Ca(OH)_2$, and $CH_3COOH$ solutions on the mechanical characteristics of RuC was investigated in previous research [32]. When compared to untreated rubber concrete, researchers discovered that all treatment methods increased compressive strength after 28 days (32). It happens because of the increasing RC's surface roughness. The compressive strength of RuC was found to be the best after all of the pre-treatment processes. Moreover, pozzolana which contains silica fume, fly ash, or rice husk ash can strengthen RuC [32–34]. Most investigations discovered that tensile strength outcomes followed compressive strength results pattern [34,35].

© STM Journals 2025. All Rights Reserved 

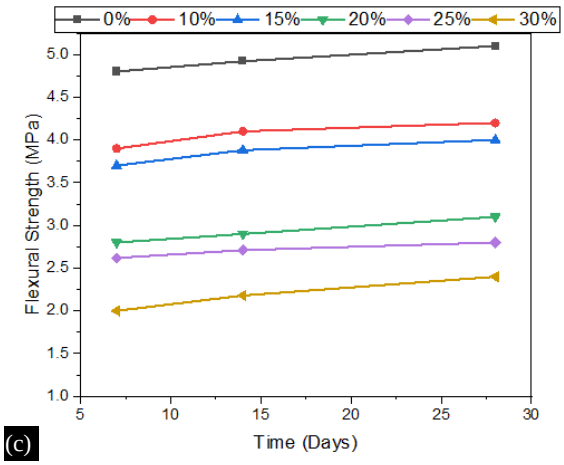
According to the previous study, the modulus of elasticity (MOE) tends to rise as particle size increases [36]. Some researchers agreed that when the rubber content increased, the MOE decreased [30,37]. When compared to natural aggregates for both treated and untreated rubber samples, the rubber aggregate had a low elastic modulus, according to Gupta et al. (38). The partial substitution of stiff FA with rubber crumbs with low elastic modulus decreases the MOE of RuC(Figure.4) The MOE of concrete is shown to be dependent on the elastic modulus of the aggregate [39]. RuC's MOE was enhanced by 24.4 % with the addition of silica fume. As per the author's findings, the inclusion of silica facilitated the filling of gaps between the aggregate and cement paste, thereby augmenting bonding [40].




(a)

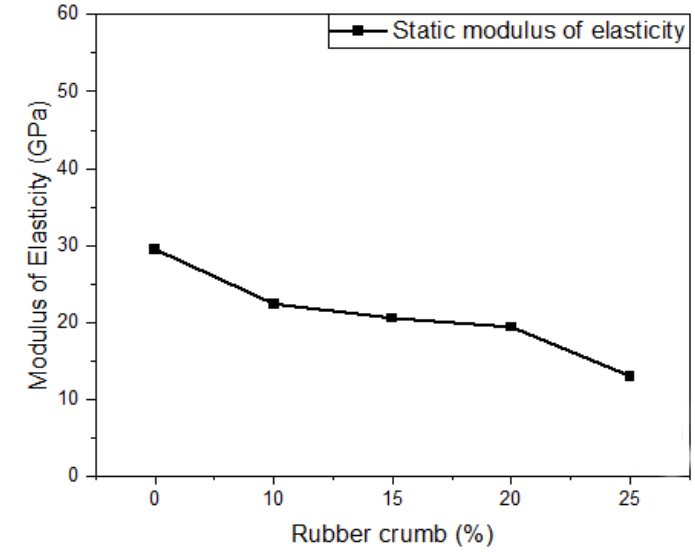



(b)




(c)

© STM Journals 2025. All Rights Reserved 

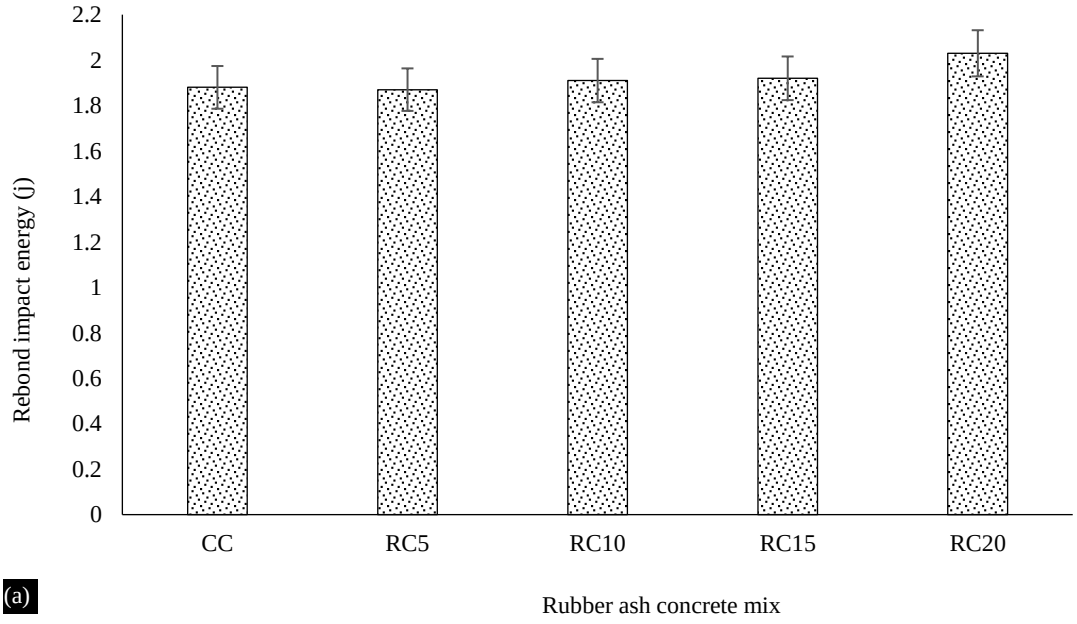



(d)

**Figure 4.** Mechanical properties of RuC (41).

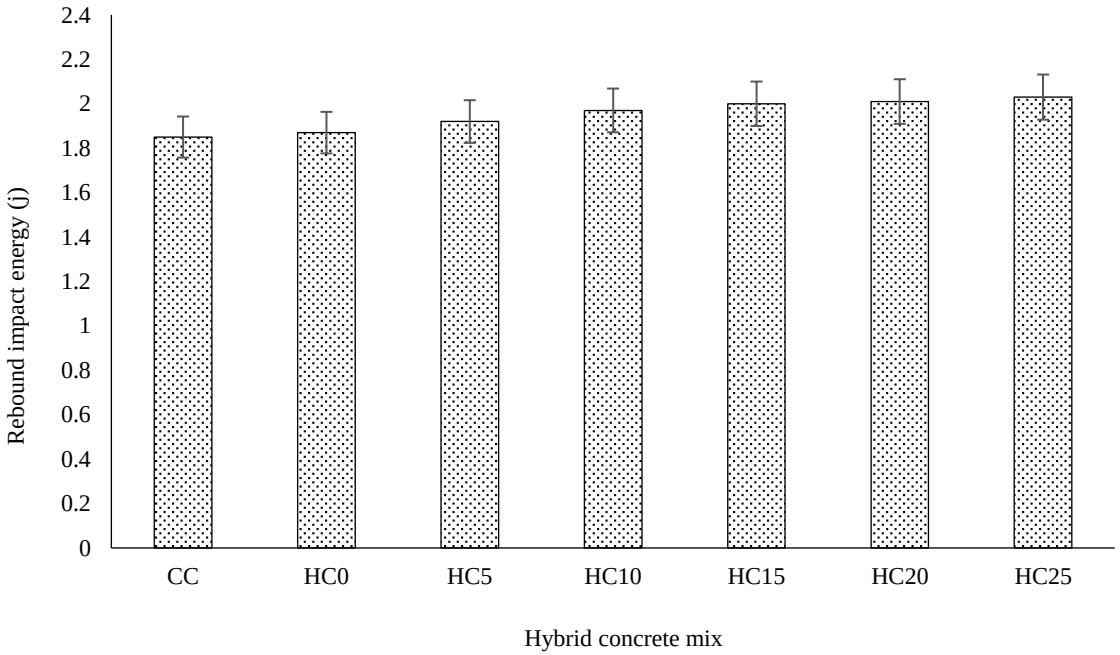
**Energy Absorption**

RuC's ability to absorb energy is associated with its concrete stiffness [42]. Many researchers have suggested that using rubber as a component in concrete to replace partial aggregate can help RuC absorb more impact energy compared to conventional concrete [30,38,43,44]. RuC fracture energy rises to a 20 percent replacement level of RC, according to the study [43]. Furthermore, according to these researchers [37,45,46], the substitution of RC in FA had good results in terms of energy absorption as the proportion of RC increased. Waste rubber RC may also be utilized as a sustainable resource for improving the impact resistance and flexibility of concrete, according to studies [38,47]. Furthermore, research [48,49] revealed that the energy absorption capacity of RC concrete is superior to that of conventional concrete and that the rubber content is more important than the rubber size in improving energy absorption(Figure.5)




(a)

© STM Journals 2025. All Rights Reserved 




(b)

**Figure 5**. (a) Impact energy under rebound of rubber ash concrete mixes (b) Impact energy under rebound of hybrid concrete mixes (50).

## CONCLUSION

After a review of many articles, some conclusions have been introduced. RuC workability is reduced when rubber content is increased. The size of the rubber aggregate also impacts the workability of RuC. The workability of RuC decreases as the size of the rubber increases, while rubber pre-treatment does not affect concrete workability. Mechanical characteristics such as compressive strength, tensile strength, and MOE are reduced due to increased rubber replacement and size. Still, adding pozzolanic material in RuC and the surface treatment of RC helps in increasing the mechanical properties of RuC. Many studies have proposed that NaOH solution be used to treat the surface of rubber aggregate. When compared to chipped rubber for partially replacing CA, RC is a perfect size to replace FA partially. With a rise in rubber content, the outcomes of MOE similarly reduce. Finally, when the proportion of rubber increases, energy absorption capacity increases, indicating that RC is excellent for producing a good energy absorption capacity of RuC. As a result of the studies, it appears that replacing FA with RC in concrete is suitable for use as an energy absorber.

## REFERENCES

1. Xiong Z, Fang Z, Feng W, Liu F, Yang F, Li L. Review of dynamic behaviour of rubberised concrete at material and member levels [Internet]. Vol. 38, Journal of Building Engineering. Elsevier Ltd; 2021. p. 102237. Available from: https://doi.org/10.1016/j.jobe.2021.102237
2. Palanisamy S, Kalimuthu M, Dharmalingam S, Alavudeen A, Nagarajan R, Ismail SO, et al. Effects of fiber loadings and lengths on mechanical properties of Sansevieria Cylindrica fiber reinforced natural rubber biocomposites. Mater Res Express [Internet]. 2023 Aug 1;10(8):085503. Available from: https://iopscience.iop.org/article/10.1088/2053-1591/acefb0
3. Sivasubramanian P, Mayandi K, Santulli C, Alavudeen A, Rajini N. Effect of Fiber Length on Curing and Mechanical Behavior of Pineapple Leaf Fiber (PALF) Reinforced Natural Rubber Composites. J Nat Fibers [Internet]. 2022 Nov 2;19(11):4326–37. Available from: https://doi.org/10.1080/15440478.2020.1856281
4. Murugesan TM, Palanisamy S, Santulli C, Palaniappan M. Mechanical characterization of alkali treated Sansevieria cylindrica fibers–Natural rubber composites. Mater Today Proc [Internet]. 2022;62:5402–6. Available from: https://doi.org/10.1016/j.matpr.2022.03.609

© STM Journals 2025. All Rights Reserved 

5. Palanisamy S, Mani Murugesan T, Palaniappan M, Santulli C. Characterization of an eco-friendly rubber composite material based on Sansevieria cylindrica fibers. Mater Today Proc [Internet]. 2023 Aug;(June):4–8. Available from: https://doi.org/10.1016/j.matpr.2023.08.202
6. Palanisamy S, Mayandi K, Palaniappan M, Alavudeen A, Rajini N, Vannucchi de Camargo F, et al. Mechanical Properties of Phormium Tenax Reinforced Natural Rubber Composites. Fibers [Internet]. 2021 Feb 1;9(2):11. Available from: https://www.mdpi.com/2079-6439/9/2/11
7. Alfayez SA, Suleiman AR, Nehdi ML. Recycling tire rubber in asphalt pavements: State of the art. Sustain. 2020;12(21):1–15.
8. Su H, Yang J, Ling TC, Ghataora GS, Dirar S. Properties of concrete prepared with waste tyre rubber particles of uniform and varying sizes. J Clean Prod [Internet]. 2015;91:288–96. Available from: http://dx.doi.org/10.1016/j.jclepro.2014.12.022
9. Ganjian E, Khorami M, Maghsoudi AA. Scrap-tyre-rubber replacement for aggregate and filler in concrete. Constr Build Mater [Internet]. 2009;23(5):1828–36. Available from: http://dx.doi.org/10.1016/j.conbuildmat.2008.09.020
10. Mhaya AM, Fahim Huseien G, Faridmehr I, Razin Zainal Abidin A, Alyousef R, Ismail M. Evaluating mechanical properties and impact resistance of modified concrete containing ground Blast Furnace slag and discarded rubber tire crumbs. Constr Build Mater [Internet]. 2021;295. Available from: https://doi.org/10.1016/j.conbuildmat.2021.123603
11. Mutalib NANA, Mokhatar SN, Budiea AMA, Jaini ZM, Kamarudin AF, Noh MSM. A review: Study on waste rubber as construction material. IOP Conf Ser Mater Sci Eng. 2021;1144(1):012003.
12. Kumar R, Verma M, Dev N. Investigation on the Effect of Seawater Condition, Sulphate Attack, Acid Attack, Freeze–Thaw Condition, and Wetting–Drying on the Geopolymer Concrete. Iran J Sci Technol Trans Civ Eng [Internet]. 2021 Nov 23; Available from: https://doi.org/10.1007/s40996-021-00767-9
13. Kumar R, Dev N. Effect of acids and freeze–thaw on the durability of modified rubberized concrete with optimum rubber crumb content. J Appl Polym Sci [Internet]. 2022 Jun 5;139(21):52191. Available from: https://onlinelibrary.wiley.com/doi/10.1002/app.52191
14. Rashid K, Yazdanbakhsh A, Rehman MU. Sustainable selection of the concrete incorporating recycled tire aggregate to be used as medium to low strength material. J Clean Prod [Internet]. 2019;224:396–410. Available from: https://doi.org/10.1016/j.jclepro.2019.03.197
15. Holmes N, Browne A, Montague C. Acoustic properties of concrete panels with crumb rubber as a fine aggregate replacement. Constr Build Mater [Internet]. 2014;73:195–204. Available from: http://dx.doi.org/10.1016/j.conbuildmat.2014.09.107
16. Kumar R, Dev N, Ram S, Verma M. Investigation of dry-wet cycles effect on the durability of modified rubberised concrete. Forces Mech [Internet]. 2023 Feb;10(December 2022):100168. Available from: https://doi.org/10.1016/j.finmec.2023.100168
17. Kumar R, Verma M, Dev N. Investigation of fresh, mechanical, and impact resistance properties of rubberized concrete. In: International e-Conference on Sustainable Development & Recent Trends in Civil Engineering [Internet]. 2022. p. 88–94. Available from: https://doi.org/10.1016/j.finmec.2023.100168
18. Elchalakani M. High strength rubberized concrete containing silica fume for the construction of sustainable road side barriers. Structures [Internet]. 2015;1:20–38. Available from: http://dx.doi.org/10.1016/j.istruc.2014.06.001
19. Youssf O, Hassanli R, Mills JE, Skinner W, Ma X, Zhuge Y, et al. Influence of mixing procedures, rubber treatment, and fibre additives on rubcrete performance. J Compos Sci. 2019;3(2).
20. Kumar R, Dev N. Mechanical and Microstructural Properties of Rubberized Concrete After Surface Modification of Waste Tire Rubber Crumb. Arab J Sci Eng [Internet]. 2022 Apr 27;47(4):4571–87. Available from: https://doi.org/10.1007/s13369-021-06154-w
21. Raffoul S, Garcia R, Pilakoutas K, Guadagnini M, Medina NF. Optimisation of rubberised concrete with high rubber content: An experimental investigation. Constr Build Mater [Internet]. 2016;124:391–404. Available from: http://dx.doi.org/10.1016/j.conbuildmat.2016.07.054
22. Si R, Guo S, Dai Q. Durability performance of rubberized mortar and concrete with NaOH-Solution treated rubber particles. Constr Build Mater [Internet]. 2017;153:496–505. Available from: http://dx.doi.org/10.1016/j.conbuildmat.2017.07.085

© STM Journals 2025. All Rights Reserved 

23. He L, Ma Y, Liu Q, Mu Y. Surface modification of crumb rubber and its influence on the mechanical properties of rubber-cement concrete. Constr Build Mater [Internet]. 2016 Sep;120:403–7. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0950061816307462
24. Abdelmonem A, El-Feky MS, Nasr ESAR, Kohail M. Performance of high strength concrete containing recycled rubber. Constr Build Mater [Internet]. 2019;227:116660. Available from: https://doi.org/10.1016/j.conbuildmat.2019.08.041
25. Strukar K, Kalman T, Miličević I, Bušić R. Potential use of rubber as aggregate in structural reinforced concrete element – A review. Eng Struct [Internet]. 2019;188(August 2018):452–68. Available from: https://doi.org/10.1016/j.engstruct.2019.03.031
26. Kumar R, Verma M, Dev N, Lamba N. Influence of chloride and sulfate solution on the long-term durability of modified rubberized concrete. J Appl Polym Sci [Internet]. 2022 Oct 5;139(37):1–15. Available from: https://onlinelibrary.wiley.com/doi/10.1002/app.52880
27. Liu H, Wang X, Jiao Y, Sha T. Experimental investigation of the mechanical and durability properties of crumb rubber concrete. Materials (Basel). 2016;9(3).
28. Thomas BS, Kumar S, Mehra P, Gupta RC, Joseph M, Csetenyi LJ. Abrasion resistance of sustainable green concrete containing waste tire rubber particles. Constr Build Mater [Internet]. 2016;124:906–9. Available from: http://dx.doi.org/10.1016/j.conbuildmat.2016.07.110
29. Najim KB, Hall MR. Crumb rubber aggregate coatings/pre-treatments and their effects on interfacial bonding, air entrapment and fracture toughness in self-compacting rubberised concrete (SCRC). Mater Struct Constr. 2013;46(12):2029–43.
30. Li D, Zhuge Y, Gravina R, Mills JE. Compressive stress strain behavior of crumb rubber concrete (CRC) and application in reinforced CRC slab. Constr Build Mater [Internet]. 2018;166:745–59. Available from: https://doi.org/10.1016/j.conbuildmat.2018.01.142
31. Mohammadi I, Khabbaz H. Shrinkage performance of Crumb Rubber Concrete (CRC) prepared by water-soaking treatment method for rigid pavements. Cem Concr Compos [Internet]. 2015;62:106–16. Available from: http://dx.doi.org/10.1016/j.cemconcomp.2015.02.010
32. Muñoz-Sánchez B, Arévalo-Caballero MJ, Pacheco-Menor MC. Influence of acetic acid and calcium hydroxide treatments of rubber waste on the properties of rubberized mortars. Mater Struct Constr. 2017;50(1).
33. Hadipramana J, Samad AAA, Ibrahim R, Mohamad N, Riza FV. The energy absorption of modified foamed concrete with rice husk ash subjected to impact loading. ARPN J Eng Appl Sci. 2016;11(12):7437–42.
34. Copetti CM, Borges PM, Squiavon JZ, da Silva SR, de Oliveira Andrade JJ. Evaluation of tire rubber surface pre-treatment and silica fume on physical-mechanical behavior and microstructural properties of concrete. J Clean Prod. 2020;256.
35. Jokar F, Khorram M, Karimi G, Hataf N. Experimental investigation of mechanical properties of crumbed rubber concrete containing natural zeolite. Constr Build Mater [Internet]. 2019;208:651–8. Available from: https://doi.org/10.1016/j.conbuildmat.2019.03.063
36. Li L, Ruan S, Zeng L. Mechanical properties and constitutive equations of concrete containing a low volume of tire rubber particles. Constr Build Mater [Internet]. 2014;70:291–308. Available from: http://dx.doi.org/10.1016/j.conbuildmat.2014.07.105
37. Atahan AO, Yücel AÖ. Crumb rubber in concrete: Static and dynamic evaluation. Constr Build Mater [Internet]. 2012;36:617–22. Available from: http://dx.doi.org/10.1016/j.conbuildmat.2012.04.068
38. Gupta T, Chaudhary S, Sharma RK. Assessment of mechanical and durability properties of concrete containing waste rubber tire as fine aggregate. Constr Build Mater [Internet]. 2014 Dec;73:562–74. Available from: http://dx.doi.org/10.1016/j.conbuildmat.2014.09.102
39. Onuaguluchi O, Panesar DK. Hardened properties of concrete mixtures containing pre-coated crumb rubber and silica fume. J Clean Prod [Internet]. 2014;82:125–31. Available from: http://dx.doi.org/10.1016/j.jclepro.2014.06.068
40. Gupta T, Chaudhary S, Sharma RK. Mechanical and durability properties of waste rubber fiber concrete with and without silica fume. J Clean Prod [Internet]. 2016;112:702–11. Available from: http://dx.doi.org/10.1016/j.jclepro.2015.07.081

© STM Journals 2025. All Rights Reserved 

41. Kumar R, Dev N. Assessment of Mechanical and Impact Resistance Properties of Rubberized Concrete After Surface Modification of Rubber Crumb. Iran J Sci Technol Trans Civ Eng [Internet]. 2022 Aug 23;46(4):2855–71. Available from: https://doi.org/10.1007/s13369-021-06154-w
42. Topçu IB. The properties of rubberized concretes. Cem Concr Res. 1995;25(2):304–10.
43. Roychand R, Gravina RJ, Zhuge Y, Ma X, Youssf O, Mills JE. A comprehensive review on the mechanical properties of waste tire rubber concrete. Constr Build Mater [Internet]. 2020;237:117651. Available from: https://doi.org/10.1016/j.conbuildmat.2019.117651
44. Gerges NN, Issa CA, Fawaz SA. Rubber concrete: Mechanical and dynamical properties. Case Stud Constr Mater [Internet]. 2018;9. Available from: https://doi.org/10.1016/j.cscm.2018.e00184
45. Ozbay E, Lachemi M. Compressive strength, abrasion resistance and energy absorption capacity of rubberized concretes with and without slag. Mater Struct. 2011;1297–307.
46. Ganesan N, Raj B, Shashikala AP. Behavior of self-consolidating rubberized concrete beam-column joints. ACI Mater J. 2013;110(6):697–704.
47. Najim KB, Hall MR. Mechanical and dynamic properties of self-compacting crumb rubber modified concrete. Constr Build Mater [Internet]. 2012;27(1):521–30. Available from: http://dx.doi.org/10.1016/j.conbuildmat.2011.07.013
48. XU J, Yao Z, Yang G, Han Q. Research on crumb rubber concrete: From a multi-scale review. Constr Build Mater [Internet]. 2020;232:117282. Available from: https://doi.org/10.1016/j.conbuildmat.2019.117282
49. Issa CA, Salem G. Utilization of recycled crumb rubber as fine aggregates in concrete mix design. Constr Build Mater [Internet]. 2013;42:48–52. Available from: http://dx.doi.org/10.1016/j.conbuildmat.2012.12.054
50. Gupta T, Tiwari A, Siddique S, Sharma RK, Chaudhary S. Response assessment under dynamic loading and microstructural investigations of rubberized concrete. J Mater Civ Eng. 2017;29(8):1–15.

© STM Journals 2025. All Rights Reserved 

AIP Conference Proceedings

RESEARCH ARTICLE | JANUARY 16 2025

# A review on Kerr optical frequency combs

Parth Dipendra ✉; Ashish Kumar Satya; A. S. Rao

*AIP Conf. Proc.* 3233, 020012 (2025)
https://doi.org/10.1063/5.0244603

View Online Export Citation

Articles You May Be Interested In

Optical frequency comb generation by pulsed pumping

*APL Photonics* (May 2017)

Drifting cavity solitons and dissipative rogue waves induced by time-delayed feedback in Kerr optical frequency comb and in all fiber cavities

*Chaos* (October 2017)

Routes to spatiotemporal chaos in Kerr optical frequency combs

*Chaos* (January 2014)

21 January 2025 07:00:08

AIP Publishing

# A Review on Kerr Optical Frequency Combs

Parth Dipendra[1, a)], Ashish Kumar Satya[1], A.S. Rao[1]

[1]Department of Applied Physics, Delhi Technological University, Bawana Road, New Delhi-110042, INDIA.

[a)]Corresponding author Email: parthdipendra_2k19ep067@dtu.ac.in

**Abstract:** Optical frequency combs are structures consisting of numerous, equispaced spectral lines. Its contribution in precision spectroscopy was awarded one half of the Nobel prize in physics in 2005. They find their use in a wide array of fields such as metrology, photonics, optical communication etc. Mode-locked lasers are the conventional technique in producing frequency combs. A new class of combs operating on the principle of parametric frequency conversion inside special microresonators has emerged known as Kerr combs. In this paper, we review the relevant physics behind Kerr combs. We take a look at the Ikeda map, which describes the evolution of the intracavity field. The Lugiato-Lefever equation, a convenient formalization of the field, is covered as well. We provide a review of how the technology of frequency combs has progressed over the last two decades, from its conception, and the theory behind its working along with its applications and future scope in different branches of science.

Keywords: Kerr frequency combs; Microresonators; Four-wave mixing; Lugiato-Lefever equation.

## INTRODUCTION

An optical frequency comb (OFC) is a comb-like spectrum, composed of a large number of fine spectral lines or 'teeth', each having the same spacing from the next in the frequency domain. In addition to being equidistant, all the modes have a defined, constant relation. This property makes them ideal for usage in precision spectroscopy, which led to the Nobel prize in 2005. OFCs have revolutionized precision metrology, spectroscopy, and communication systems since their inception in the 1990s [1].

OFCs have traditionally been realised using Mode-locked lasers (MLLs), a special kind of laser which generates ultrashort pulses by a technique known as mode-locking. In recent years special cavity resonators called microresonators have emerged as a platform for generation of compact and low-power combs [2]. Based on this, a new class of frequency combs, called Kerr combs, generated through utilizing nonlinear effects to produce OFCs inside microresonators has emerged.

In this paper, we review the basic concepts relating to frequency combs and their historical development, followed by a detailed description of the generation of Kerr combs. We discuss the experimental setup for generating Kerr combs, including the different types of microresonators used. Theoretical models of comb generation discussed include the Ikeda map and the Lugiato-Lefever equation. Finally, we discuss the various applications of Kerr combs in fields such as spectroscopy, telecommunications and metrology [3][4][5].

### *BASIC PRINCIPLES OF FREQUENCY COMBS*

OFCs are a series of discrete frequencies with equal spacing between two lines. When translated into the time domain, they are ultrashort pulses with spacing ranging between femtoseconds to attoseconds. The width of the separation is called the repetition rate, $f_r$ (or $f_{rep}$). It is the inverse of the cavity round-trip time, $t_r$, as after each roundtrip a pulse is generated. It is from this repetition rate that one may realize how the modes are linked harmonically.

As the pulses are created, the electric field of the pulses can be expressed as the product of the carrier, having angular frequency $\omega_c$, and the second being the envelope modulating the carrier given as A(t).

$$E(t) = A(t) \cdot \exp(i\omega_c t) \tag{1}$$

However, due to nonlinear effects, dispersion and broadening, causing a difference between the group and phase velocities during transmission, creating an offset between the carrier and the envelope [7]. The phase difference between the phase of carrier and the envelope is defined as the carrier-envelope offset phase ($\varphi_o$). The offset frequency ($f_o$) of the comb is then given as

*Recent Advancements in Artificial Intelligence and its Applications in Engineering*
AIP Conf. Proc. 3233, 020012-1–020012-5; https://doi.org/10.1063/5.0244603
Published under an exclusive license by AIP Publishing. 978-0-7354-5098-1/$30.00



21 January 2025 07:00:08

$$f_o = \frac{1}{2\pi}\frac{d\varphi_o}{dt} \quad (2)$$

We may express the frequency of a mode 'n' of the comb as

$$f_n = f_o + nf_r \quad (3)$$

This equation is known as the comb equation. With knowledge of the two parameters $f_o$ and $f_r$, the frequency of each mode in the comb can be determined. The use of negative feedback can be made for stabilizing both $f_o$ and $f_r$ [8].

*KERR FREQUENCY COMBS*

The generation of frequency combs through the use of MLLs is a well-established technique. However, this method has some drawbacks, including the complexity, cost of setup and bulky size. With increase in research on frequency combs owing to their applications, a new technique for generating OFCs was invented. In 2007, Del'Haye et al. [10] were able to generate a frequency comb by using a continuous-wave (CW) laser pumping laser light inside a microtoroidal cavity. The OFCs generated in this manner was referred to as 'Kerr optical frequency comb'.

*Microresonators*

Kerr combs are formed inside special structures known as microresonators. These are optical microcavities, having dimensions in the micrometer to the millimeter range. Inside these, losses are extremely low, giving them high Quality factors. In addition, they have very small mode volumes and large free spectral range (FSR). Most of the microresonators operate by confining light, where it propagates within the confines of the cavity by the principle of total internal reflection (TIR) at the dielectric-air boundary. Others operate using a distributed Bragg reflector (DBR), formed from an arrangement of periodic structures, by varying the refractive index inside the cavity [11]. The former class of microresonators generally excite a certain kind of optical modes, known as whispering gallery modes (WGM). These modes are present when light travels around by means of TIR inside concave surfaces. Hence these microresonators are often also called WGM resonators (WGMRs). Another class of microresonators that are also popularly used for comb generation are integrated ring resonators [14].

*Kerr Effect and Four-Wave Mixing*

Inside microresonators, frequency combs are generated by the generation of sidebands from a parametric frequency generation process. Conversion of frequency starts with the Kerr effect. The Kerr effect is a nonlinear effect, describing an intensity dependent refractive index of a material due to the presence of an electric field, given by

$$n = n_0 + n_2 I \quad (4)$$

Where $n$ is the refractive index of the medium, $n_0$ is the linear refractive index, $n_2$ is the Kerr coefficient and I the intensity of the field. The Kerr effect is the more dominant of the nonlinear effects arising in centrosymmetric crystals and amorphous media [12]. This effect causes the excitation of WGMs inside the resonator, giving rise to the parametric phenomenon of four-wave mixing (FWM) [4]. FWM is a third order $\chi^{(3)}$ nonlinear effect leading to photons interaction of the type: $\hbar\omega_a + \hbar\omega_b \rightarrow \hbar\omega_m + \hbar\omega_n$. In case of single input frequency, the process is termed degenerate FWM, which is what initially occurs when pump photons get absorbed by the medium, generating two new photons, termed as signal and idler, having the same energy difference and a constant phase relation with respect to the pump, creating sidebands.

As the sideband grows, non-degenerate FWM occurs between the signal and the idler photons, further generating new sidebands. The same FWM occurs for the newly formed sidebands as well, resulting in a cascading process. This is known as cascaded FWM. In this manner, the phenomenon of cascaded FWM gives rise to an OFC, known as Kerr comb.

Kerr combs as compared to the traditional OFCs offer higher repetition rates, going into the GHz region. The other important benefit of Kerr combs is planar chip-scale integration. Kerr combs can be formed from photonic circuits that are compact in nature. This allows OFC to be used directly in any device fabricated with these chips.

21 January 2025 07:00:08

## MODELLING OF KERR COMBS

The modelling of Kerr combs is important for understanding the behavior of these devices and optimizing their performance for specific applications. Various theoretical models based on different approaches have been developed. Perhaps the most fundamental model for describing the generation of an OFC inside microresonators is the Ikeda map.

*Ikeda Map*

Given by Kensuke Ikeda in 1979 [13], the Ikeda map is a theoretical model which describes the evolution of the field envelope propagating through the medium by undergoing TIR inside the microresonator, along with the boundary condition which couples the field inside the resonator after each round trip with the input of the pump field. A mathematical description of the aforementioned is provided by segregating the field's evolution, partitioning it into time intervals corresponding to one round trip of the field inside the resonator. The Ikeda map then may be given as:

$$\frac{\partial E_m(z,\tau)}{\partial z} = -\frac{1}{2}\alpha_i E_m(z,\tau) - \frac{i\beta_2}{2}\frac{\partial^2 E_m(z,\tau)}{\partial \tau^2} + i\gamma|E_m(z,\tau)|^2 E_m(z,\tau) \quad (5)$$

$$E_{m+1}(0,\tau) = \sqrt{\theta}E_{in} + \sqrt{1-\theta}e^{i\phi_0}E_{m+1}(L,\tau) \quad (6)$$

Where, $E_m$ is the field envelope of the m$^{th}$ round trip, more simply referred to as the intracavity field, $\tau$ is the time, $L$ the resonator's roundtrip length, $\alpha_i$ is the cavity absorption loss, $\beta_2$ is the group velocity dispersion (GVD) coefficient, $\gamma$ is the nonlinear coefficient, $\theta$ is the coupling or power transmission coefficient, $E_{in}$ is the pumping fields' amplitude and $\phi_0$ is the linear phase constant. Equation 5 describes the field evolution period of a roundtrip, while equation 6 is the periodic boundary condition which accounts the CW laser field being pumped into the resonator and some of the cavity field exiting the resonator at the point of coupling. Here we also make an approximation that the coupling with the laser field taking place is over a point as opposed to a finite area, which is actually the case.

*Lugiato-Lefever Equation (LLE)*

The Ikeda map is a very useful tool in deriving various other models for comb generation, the most popular of which is the Lugiato-Lefever equation (LLE), by using a mean-field approximation. This is done to find the field in the resonator after one roundtrip which is $E_m(L,\tau)$. In this approximation, the change of the envelope field is slow enough to neglect the variation of $E_m(z,\tau)$ with respect to z on the right-hand side of equation 5. This approximation is also a condition that the detuning must be a small parameter.
A new variable called slow time $t$ is introduced which is related as $E(t = mt_r, \tau) = E_m(0,\tau)$. $t_r$ is the roundtrip time. Time $\tau$ is referred to as the fast time, but remains the same as before.

$$t_r\frac{\partial E(t,\tau)}{\partial t} + \frac{i\beta_2 L}{2}\frac{\partial^2 E(t,\tau)}{\partial \tau^2} - i\gamma L|E(t,\tau)|^2 E(t,\tau) = \sqrt{\theta}E_{in} - (\alpha + i\delta_0)E(t,\tau) \quad (7)$$

Equation 7 is what is known as the LLE. The LLE in the mean field approximation gives a model for the generation of Kerr combs. Here, $\alpha = (\alpha_i + L)/2$ is the total cavity roundtrip loss and $\delta_0$ is the detuning of the pump field from the pump mode. In this approximation, we take the coupling with the laser pump and the cavity losses to be evenly spread around the resonator. Two different time scales, with a fast time $\tau$ and a slow time $t$ are present. While physically time is a single quantity, the fast time describes the profile of the field envelope the same as in the Ikeda map, the slow time does the same but over longer periods of time, at scales of photon lifetimes [14], for a field averaged over a round trip.
Solutions to the LLE can provide the dynamics of the field in the resonator and predict the formation and properties of cavity solitons. Cavity solitons are one of the solutions to the LLE or the Ikeda map, occurring in the form of a localized pulse. Cavity solitons frequency combs have many interesting applications, including in optical frequency synthesis, telecommunications, and sensing.

*APPLICATIONS OF KERR COMBS*

OFCs have been integrated into multiple fields owing to their brilliant spectral characteristics. By lowering the size, complexity, and cost of generation, Kerr frequency combs now have the potential to dramatically increase the usability and range of applications of frequency combs. In this section, we take a look at different uses and applications of Kerr combs and how they have replaced MLL based OFCs for these purposes.

21 January 2025 07:00:08

*Timekeeping*

Perhaps the most interesting and obvious use of Kerr combs lie in timekeeping and atomic clocks. Atomic clocks use atomic transitions as a reference for measuring time with incredibly high precision. OFCs have a much higher frequency range than them, which use microwave transitions, and hence can provide a better time resolution than microwave clocks. OFCs are used in the conversion between these regions, acting as 'light gears'. In a recent study, researchers used a Kerr frequency comb to synthesize an optical frequency with a fractional uncertainty of the order $10^{-18}$ [5].

*Spectroscopy*

Another important use of OFCs lies in the field of spectroscopy. Frequency combs are attractive for investigating atoms and molecules because they consist of a broad bandwidth array of continuous wave oscillators. One approach for spectroscopy involves resolving the modes spectrally and analyzing their amplitudes for investigating the absorptive features of the sample gas [15]. Other alternative approaches include multiheterodyne/dual-comb spectroscopy for measuring the absorption and phase spectrum of a sample using two combs of near equal repetition rates to provide a spectroscopic analysis of the sample [16].

Optical Communication

In the field of optical communication, wavelength division multiplexing (WDM) is a technique that has been employed to increase the transmission capacity of optic fibers. OFCs provide equidistant and phase-locked frequency sources for the purpose of WDM, which are otherwise expensive to generate. Kerr combs can be utilized as small WDM sources. While using OFCs for data transmission bands has been achieved, meeting the strict requirements of channel spacing and data modulation rates in this domain requires adhering to specific mode spacing and power conditions [4] that are difficult to attain using conventional MLLs. However, Kerr combs have higher power per mode and higher repetition rates of several GHz as compared to MLLs. By using Kerr combs, hundreds of channels from a single, low-power, off-chip source can be produced simultaneously, eliminating the need for multiple high-power lasers. It has been shown that crystalline WGMRs with a Q-factor in the order of $10^9$ at 1550 nm allows to reach a transmission capacity of 144 Gbit/s per channel using a 16-quadrature amplitude modulation scheme on three multiplexed carriers [17].

## CONCLUSION

In this paper, we discussed the optical frequency comb technology and its evolution from over 20 years ago. OFCs find their use in frequency metrology, precision spectroscopy and other scientific fields. By leveraging nonlinear effects, namely the Kerr effect and FWM, inside microresonators, a new technique of generating OFCs, termed as Kerr frequency combs, was discovered. We discuss the theoretical models for the generation of Kerr combs, namely, the Ikeda map, and the Lugiato-Lefever equation. A solution to the LLE and the Ikeda map in the form of cavity solitons, an interesting research topic, with soliton combs having several applications, are also discussed briefly. In the final section, we look at the applications of OFCs and the advantages of using Kerr combs in domains of timekeeping, spectroscopy and optical communication. Kerr combs show great promise as a coherent, broadband optical source which will advance the technology of OFCs to new frontiers.

## REFERENCES


1. Fortier, Tara, and Esther Baumann. "20 years of developments in optical frequency comb technology and applications." Communications Physics 2, no. 1 (2019): 153.
2. Alessia Pasquazi, et al. "Micro-combs: A novel generation of optical sources," Physics Reports, Volume 729, 2018, 1-81.
3. Thorpe, M. J., & Ye, J. (2008). "Cavity-enhanced direct frequency comb spectroscopy," Applied Physics B, 91(3-4), 397-414.
4. Kippenberg, Tobias J., Ronald Holzwarth, and Scott A. Diddams. "Microresonator-based optical frequency combs." science 332, no. 6029 (2011): 555-559.
5. Newbury, N. R., & Swann, W. C. (2011). "Low-noise fiber-laser frequency combs" (Invited). Journal of the Optical Society of America B, 28(11), A40-A50.
6. Wiroj Sudatham, Hirokazu Matsumoto, Satoru Takahashi, Kiyoshi Takamasu, "Non-contact measurement technique for dimensional metrology using optical comb." Measurement, Volume 78, 2016, Pages 381-387,


21 January 2025 07:00:08

7. Diddams, Scott A., Kerry Vahala, and Thomas Udem. "Optical frequency combs: Coherently uniting the electromagnetic spectrum." Science 369.6501 (2020): eaay3676.
8. Jones, David J., et al. "Carrier-envelope phase control of femtosecond mode-locked lasers and direct optical frequency synthesis." Science 288.5466 (2000): 635-639.
9. Zhang, Hao, et al. "Optical frequency combs: From principles to applications." Journal of Electronic Science and Technology (2022): 100157.
10. Del'Haye, Pascal, et al. "Optical frequency comb generation from a monolithic microresonator." Nature 450.7173 (2007): 1214-1217.
11. Benson, Trevor M., et al. "Micro-optical resonators for microlasers and integrated optoelectronics." Frontiers in Planar Lightwave Circuit Technology: Design, Simulation, and Fabrication. Springer Netherlands, 2006.
12. Chembo, Yanne K. "Kerr optical frequency combs: theory, applications and perspectives." Nanophotonics 5.2 (2016): 214-230.
13. Ikeda, Kensuke. "Multiple-valued stationary state and its instability of the transmitted light by a ring cavity system." Optics communications 30.2 (1979): 257-261.
14. Hendry, Ian, et al. "Spontaneous symmetry breaking and trapping of temporal Kerr cavity solitons by pulsed or amplitude-modulated driving fields." Physical Review A 97.5 (2018): 053834.
15. Diddams, Scott A., Leo Hollberg, and Vela Mbele. "Molecular fingerprinting with the resolved modes of a femtosecond laser frequency comb." Nature 445.7128 (2007): 627-630.
16. Keilmann, Fritz, Christoph Gohle, and Ronald Holzwarth. "Time-domain mid-infrared frequency-comb spectrometer." Optics letters 29.13 (2004): 1542-1544.
17. Pfeifle, Joerg, et al. "Optimally coherent Kerr combs generated with crystalline whispering gallery mode resonators for ultrahigh capacity fiber communications." Physical Review Letters 114.9 (2015): 093902.

21 January 2025 07:00:08

Multimedia Tools and Applications
https://doi.org/10.1007/s11042-025-20599-3

# A two-stage ensemble approach for the analysis of Parkinson's Disease using speech signals

**Kavita Bhatt**[1] · **N Jayanthi**[1] · **Manjeet Kumar**[1]

Received: 20 July 2023 / Revised: 23 December 2024 / Accepted: 1 January 2025
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2025

**Abstract**
Parkinson's Disease (PD) is a chronic neuro-degenerative disorder that typically targets the neurons in the human brain. In the early stages of PD, speech impairments were discovered in the majority of patients. This study proposes a two-stage ensemble model to diagnose the disease in its initial stage utilizing speech signals. Firstly, an extra tree classifier features selection technique is utilized to discover the most appropriate features from the dataset and the dataset is balanced by the synthetic minority oversampling technique. Finally, the classification problem is addressed by the amalgamation of three base learners, K-nearest neighbor, decision tree, and Naive Bayes, and a new stacked ensemble model is constructed to detect PD. The effectiveness and robustness of the proposed work have been evaluated on accuracy, sensitivity, specificity, and F-1 score. The detection accuracy obtained by the proposed method is 98.31%. The results demonstrate that, in contrast to the traditional feature selection and single classifier technique, the proposed method can effectively decrease the number of features while preserving a higher rate of accuracy of detection. The suggested method surpasses the current cutting-edge studies conducted in this area.

**Keywords** Parkinson's disease · Feature selection · Machine learning · Synthetic minority oversampling technique · Stacking

## 1 Introduction

Parkinson's disease (PD) is a chronic disorder that causes damage to the neurons in the brain which is responsible for the production of the neurotransmitter dopamine. Dopamine is responsible for transferring messages to various areas of the brain, which in turn govern the overall movement of the body parts. As a result of the ongoing deterioration of dopaminergic

✉ N Jayanthi
njayanthi@dce.ac.in

Kavita Bhatt
kavitabhatt57@gmail.com

Manjeet Kumar
manjeetchhillar@gmail.com

[1] Department of Electronics and Communication Engineering, Delhi Technological University (DTU), Shahbad Daulatpur, Delhi 110042, India

Published online: 08 January 2025

neurons, individuals with this disease exhibit a chronic neurological disorder [1]. As PD progresses in the body, a depreciation in the supply of dopamine fluid can be noticed in the body, leaving a person with a disability of movement. Age is considered as the most significant risk factor for PD. With a risk ratio of 1.46, men are at higher risk than women [2]. Generally, disease incidence rises exponentially above the age of 65 [3]. PD is regarded as the second most common brain disease in elderly people after Alzheimer's. Globally, 6.9 million people had PD as of 2015, and 14.2 million are expected to have it by 2040 [4]. The cure for PD is still not available, but a few medicines may help to slow down the pace of spread. Parkinson's symptoms are classified as motor and non-motor symptoms. The motor symptoms include bradykinesia, stiffness, and tremors, and the non-motor symptoms include cognitive impairment, sleep issues, anxiety, depression, speech difficulties, hallucinations, etc. In recent years, non-motor symptoms have become crucial while having a major influence on quality-of-life measures as well as the societal costs of PD [5].

As PD is a progressive disease, its symptoms start getting worse with time. Gradually, people start having trouble walking and even doing some of the simplest daily tasks. They may also face nocturnal problems. Depression is also seen in more than half of the PD Patients [6]. Apart from these symptoms, speech disorder is also faced by 90% of patients in the early stage [7]. The speech impairment that PD patients have is typically referred to as "hypokinetic dysarthria," and it mainly involves issues with articulation, breathing, and voice quality. Recently, many studies have been conducted utilizing speech datasets to detect PD in the initial stage.

In recent years, machine learning (ML) and deep learning (DL) methods have garnered significant attention in diagnosing and treating brain diseases [8–13]. Avci et al. [8] developed a genetic algorithm-based extreme learning algorithm for PD diagnosis using speech signals. A wavelet kernel is used to map the non-linear features of speech signals for this purpose. To categorize the healthy voice and the voice of a PD patient, a twin support vector machine (TSVM) is used in [9]. To extract the necessary speech characteristics from the speech dataset, correlation-based feature selection is performed. Rueda et al. [10] used a two-stage filter wrapper feature selection approach to minimize the number of features and applied different ML methods to detect PD. PD patients and healthy individuals are distinguished using a model based on SVM and convolutional neural networks (CNN) in [14]. SVM is utilized for classification, and CNN is employed to extract the features.

Dastjerd et al. [15] suggested a predictive model to discriminate PD patients from healthy people using inductive fuzzy classification and fuzzy rough classification. The fuzzy classifiers were tested on speech data. In order to account for PD-related alterations in the voice signal, a time-frequency characteristics model is used in [16]. To perform PD classification, a weighted local discriminative preserving projection technique (wLDPP) is used with an ensemble voting mechanism to create subspaces in [17]. Xue et al. [18] developed a local dynamic feature selection fusion (LDFSF) approach that selects features for PD recognition and symptom severity prediction tasks. The LDFSF technique based on self-organizing map clustering uses the maximal information coefficient and selects feature subsets with a higher level of relevance, minimal redundancy, and higher complementarity. Li et al. [19] used a multitype transformation algorithm for PD speech recognition by taking the speech samples.

Different ML classifiers, including logistic regression (LR), random forest (RF), and artificial neural network (ANN), are used in [20] for predicting the occurrence of PD. Recursive feature elimination (RFE) is applied for PD detection in [21]. A total of 13 features out of 22 are selected for classification. Maximum relevance minimum redundancy feature selection (mrMR) method with K-nearest neighbor (KNN) is used in [22]. Combining the Gaussian process with automatic relevance determination (ARD) leads to a better output while taking

good note of the features of PD patients [23]. Mathur et al. [24] performed ten-fold cross-validation on speech datasets using various ML algorithms, out of which a combination of KNN and ANN algorithms prevails. In [25], a gradient boosting algorithm is applied for PD diagnosis on the voice dataset. Adaptive wolf search optimization and sparse autoencoder neural network are employed to extract the vocal features in [26]. Murugappan et al. [27] employed tunable Q wavelet transform (TQWT) and probabilistic neural networks to distinguish healthy control (HC) and PD patients. They have taken 20 PD patients and 20 HC. The features are acquired from the TQWT sub-bands, and classification is performed by the probabilistic neural network.

Chintalapudi et al. [28] applied three different deep-learning architectures utilizing vocal biomarkers to identify PD. Chen et al. [29] suggested a PD detection approach based on KNN and the Hilbert-Huang transform (HHT). HHT is used to examine the 12 characteristics of each recorded voice. The experimental findings show that the KNN classifier outperforms the other classifiers, with an accuracy of 93.3%. Lamba et al. [30] developed a hybrid MIRFE feature selection strategy based on recursive feature elimination and mutual information gain. This study made use of a PD speech dataset that included 756 speech samples from 252 individuals. The MIRFE approach selects 40 features out of 754. RF and XGBoost classifiers are incorporated to assess the effectiveness of hybrid feature selection methods.

Bchir et al. [31] suggested a classification model based on a Gaussian mixture model to categorize the voice recordings of PD patients. Out of 752 features, the 50 best features are selected using the minimum redundancy maximum relevance method. A chaotic bacterial foraging optimization with an enhanced fuzzy KNN classifier is utilized for early detection of PD employing the UCI PD dataset [32]. Quan et al. [33] created an end-to-end deep-learning model for PD detection. 2D-CNN is used to extract time-series features, which are then categorized using 1D-CNN. This research makes use of two datasets. Database-1 is based on Chinese dialogues, while database-2 is based on Spanish speech data. Yaman et al. [34] utilized a statistical pooling method to detect PD. By applying the statistical pooling approach, the attributes of the dataset are improved, and then the Relief method is used to produce weighted features. SVM and KNN classifiers are utilized to categorize the obtained weighted feature vectors to PD and HC groups.

Rahman et al. [35] explored different ML methods and deep neural networks (DNN) to predict and categorize PD. Harithha et al. [36] presented a fuzzy cognitive map (FCM) feature selection approach to select relevant features in different big data classification applications. For PD classification, ten features obtained using the FCM approach achieved a classification accuracy of 91.83%. A hyperparameter tuning technique based on Bayesian optimization (BO) was used to obtain better PD detection accuracy in [37]. Shyamala and Navamani [38] developed an interpretable feature ranking XGBoost model based on RFE and XGB classifier for PD detection utilizing speech features. A Parkinson classification neural network (PCNN) was developed in [39]. The optimal set of features was selected using a multi-agent salp swarm (MASS) algorithm. The integrated approach MASS-PCNN achieved 96.77% accuracy. Saha et al.[40] developed a Euclidean distance-based graph convolutional network (GCN) method for PD classification.

The majority of earlier research focused on single classifiers. But the problem with single classifiers is, at times, they may lack accuracy and precision which can be achieved using a stacking-based method. Sometimes, these classifiers adversely affect the errors and downgrade their performance with respect to overfitting, generalization, and bias of the model, hence resulting in poor prediction. Also, most of the research has not dealt with class imbalance problems, which could result in biased model performance favoring the majority class and decreasing diagnosis reliability. To address all these problems, this paper proposes a two-

stage PD detection approach. In the first stage, an extra tree classifier (ETC)- based features selection approach has been applied in order to reduce the non-required speech features, and in the second stage, a new model is built by stacking three ML algorithms, which are Naive Bayes (NB), KNN, and DT. The suggested method provides better results by simply taking the speech features of PD patients.

The following are the paper's main contributions:

- A stacked-based ensemble ML approach has been proposed for the early detection of PD.
- This method leverages an ETC-based feature selection approach for determining the optimum features and thereby reducing the model's training time and computing complexity.
- The UCI speech dataset is balanced using the synthetic minority oversampling (SMOTE) method in order to improve the prediction accuracy of PD.
- The proposed method's reliability and efficiency are validated on the UCI PD speech dataset.
- A comparative analysis is performed between the proposed approach and existing methodologies for detecting PD. The proposed approach has produced improved PD detection results.

The remaining part of the paper is structured as follows: An elucidation of the dataset utilized in the paper is provided in Section 2. Section 3 details the proposed methodology. In this part, the feature selection strategy and classifiers used in the proposed methodology are discussed. Results and discussion are presented in Section 4 and Section 5 respectively. Limitations and future work are discussed in Section 6. Finally, in Section 7, the conclusion is covered.

## 2 Dataset description

This paper utilizes the dataset from the renowned UCI ML repository [41]. It is the speech dataset of 31 people whose age ranges from 46 to 85, of which 23 are diagnosed with PD while the rest are HC. The speech signals are sampled at 44.1 kHz and 16 bits of resolution. Amplitude normalization is performed before calculating the features. The dataset comprises 195 sustained vowel phonations that are extracted from these 31 people's speech signals. An average of six phonation recordings are provided by each person. Dysphonia measurements are utilized to distinguish between HC and PD patients. Table 1 portrays the features retrieved from speech signals. The dataset contains 22 features derived from speech signals using the Pratt software.

Figure 1 depicts the histogram representation of status distribution with dysphonia measurements.

## 3 Proposed methodology

In this paper, to identify and forecast PD, stacked-based ensemble learning is used. This work aims to boost the accuracy and speed of PD diagnosis in the context of an intelligent network by a stacked ensemble model. Figure 2 illustrates the block diagram of the proposed methodology. The acquired PD speech data is used to train the classifier using the following steps. Firstly, the raw data is pre-processed, and then the ETC-based feature selection technique is used to acquire the optimum set of features. Afterward, SMOTE is applied to overcome

**Table 1** List of dysphonia features extracted from speech signals for PD detection

| Feature Name | Feature Description |
|---|---|
| Shimmer: APQ3 | 3- Point amplitude perturbation quotient. |
| Shimmer: APQ5 | 5- Point amplitude perturbation quotient. |
| Shimmer: DDA | Difference of amplitude differences. |
| Jitter: DDP | Variation in fundamental frequency. |
| MDVP: Jitter(%) | Jitter value in percentage. |
| MDVP: Jitter(Abs) | Absolute value. |
| MDVP: Fo (Hz) | Average vocal fundamental frequency. |
| MDVP: Fhi (HZ) | Maximum vocal fundamental frequency. |
| MDVP: Flo (HZ) | Minimum vocal fundamental frequency. |
| MDVP: PPQ | Five point period perturbation quotient. |
| MDVP: APQ | Variation in amplitude. |
| MDVP: RAP | Relative average perturbation. |
| MDVP: Shimmer | Amplitude variation. |
| MDVP: Shimmer (dB) | Amplitude variation in dB. |
| Spread 1 | Nonlinear measure of fundamental frequency. |
| Spread 2 | Nonlinear measure of fundamental frequency. |
| HNR | Harmonic to Noise ratio. |
| NHR | Noise to Harmonic ratio. |
| DFA | Detrended fluctuation analysis. |
| D2 | Correlation dimension. |
| PPE | Pitch period entropy. |
| RPDE | Recurrence period density entropy. |
| Status | Health status of the person<br>1: Person with PD.<br>0: Healthy person. |

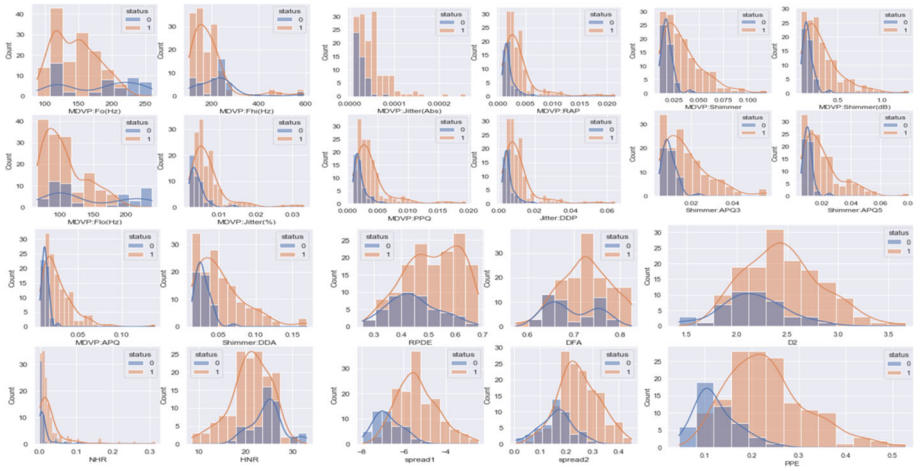



**Fig. 1** Histogram representation of status distribution with dysphonia measurements

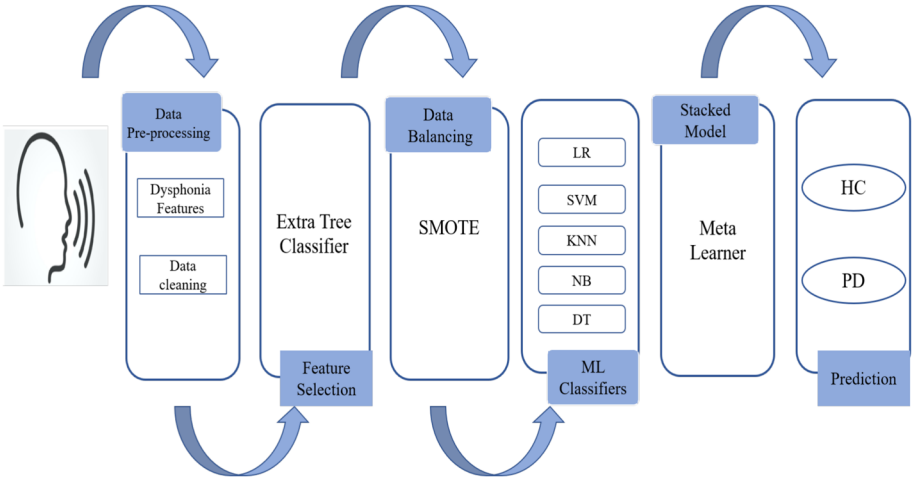



**Fig. 2** Block Diagram of the proposed methodology for PD detection

the class imbalance problem. Different base learners and meta-learners are used separately to evaluate their performance.

### 3.1 Data pre-pocessing

The most important aspect of any ML investigation is data pre-processing. The primary purpose of data cleaning is to ensure data accuracy by searching the dataset for duplicated or inaccurate information, as well as missing values present in the dataset. The data set contains dysphonia measurements taken from speech signals. The dataset includes 195 sustained vowel phonations. The data does not contain any missing values. There are 147 PD phonetics and 48 healthy phonetics. The class is specified by the status column in the data, which is 0 represents HC and 1 represents PD patients.

### 3.2 Feature selection

Feature selection aims to remove features that are insignificant or irrelevant. This is a technique for selecting a subset of crucial characteristics that accurately identify the problem. It becomes a necessity to determine the most important features as datasets having insignificant features result in the reduction of the model accuracy. Feature selection proceeds through a process of selecting requisite features from the prime features [42]. Feature selection techniques are, thus, applied to showcase the most relevant features. It comprises selecting a subset of features and, then, in turn, feature transformation. Selection of a subset of features contains both feature ranking, as well as frequency calculation of features. Feature selection contains two pivotal objectives, which are to enhance the accuracy of the classification and feature reduction [43]. In this work, an ETC-based feature selection approach is utilized for selecting the optimal set of features.

### 3.2.1 ETC-based feature selection

ETC is a tree-based supervised machine learning technique for extracting the most crucial features. The ETC is an ensemble algorithm that works on datasets by developing multiple tree models, which are constructed at random and sort out the most important features that are also voted the most [44]. It splits the nodes by fitting each decision tree and determining a split point. Nodes are divided based on the measure of random sub-nodes. These nodes are split on available variables and the obtained split gained in the results is selected in the constituent tree models, which helps to make the model less prone to overfitting by lowering the variance.

ETC calculates feature importance based on the average impurity reduction caused by a feature across all decision trees in the ensemble. The most relevant feature based on the Gini index (GI) is used by the corresponding decision tree to split the data. GI measures the probability of a specific feature being incorrectly classified when it is chosen randomly and is represented as follows:

$$GI = 1 - \sum_{i=1}^{n} p_i^2 \tag{1}$$

Where $n$ denotes the number of data points, and $p_i$ denotes the probability that a feature will be assigned to a specific class.

The feature importance scores obtained from individual decision trees in the ensemble are aggregated using a weighted voting scheme. Each decision tree contributes its importance score based on performance and accuracy. After obtaining the feature importance scores, a ranking is typically established by sorting the features in descending order based on their scores. For this dataset, the top ten features are selected by ETC. Figure 3 represents the ranking of the top ten features generated by ETC.

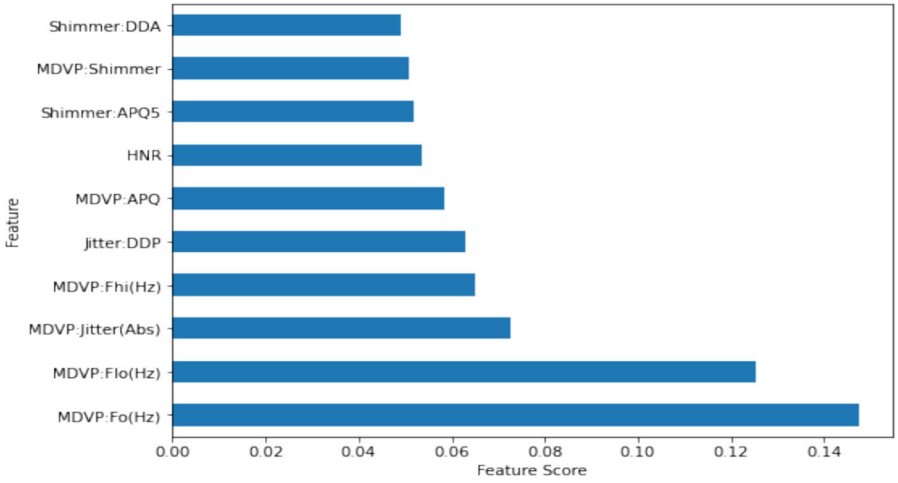



**Fig. 3** Important features using ETC-based feature selection

### 3.3 Data balancing

One of the main issues with the healthcare dataset is the imbalance of class instances. This indicates the instance is not being appropriately distributed among the different classes. As a result, the classification results from the imbalanced class data are biased, favoring the dominating class. So, SMOTE is used to balance the dataset. SMOTE is an oversampling technique where the minority class generates the synthetic samples [45]. The problems like overfitting, posed by random oversampling, can be overcome by this algorithm. It primarily focuses on the feature space, that, using interpolation between the positive instances, goes side by side, resulting in bringing out new instances. SMOTE sorts out examples that have occurred in proximity to a feature space. It then picks up a new example at a certain point along with the segment, while also joining other adjacent examples. SMOTE then, at random, opts for an instance from the minority class and then computes its KNN. After that, a random neighbor is chosen. Then, at a point selected at random between the two examples, a synthetic example is created. This process can help in generating as many synthetic examples as are needed for any particular minority class to create a balance.

### 3.4 Classifiers

In this work, various types of prominent supervised ML algorithms are used for PD detection. The labeled training dataset is initially utilized in supervised ML algorithms to evaluate the basic algorithm, and in the testing phase, these classifiers are used to identify whether a test instance is HC or PD patient.

#### 3.4.1 Logistic Regression (LR)

LR is a supervised learning algorithm commonly utilized in classification as well as predicting the probability of a certain variable. The variable is further divided into two possible classes, which are also referred to as dichotomous. The nature of the variable is binary, which codes the data in the binary form (i.e., 1,0) where 1, denotes a true, and 0 denotes a false. In a mathematical representation, an LR model predicts $P(Y = 1)$ as a function of $X$. This ML algorithm is generally used for several classification problems like breast cancer detection, heart disease prediction, Diabetes prediction, etc. The sigmoid function is utilized for mapping predicted values of probabilities into another value between 0 and 1. It is distinguished from linear regression for providing a binary outcome variable [46].

#### 3.4.2 K-nearest neighbor (KNN)

KNN algorithm is a comprehensible and easily implementable supervised learning algorithm. In general, a person, before creating a bond with people in his surroundings and forming a relationship, looks out for qualities on certain parameters like behavior, ideology, character, psychology, etc. Then, after measuring all these parameters according to his interests, he finds the best-suited people around him to be part of his social circle. In the same way, KNN works in ML. The objective of the KNN algorithm is to figure out the class to which it belongs after locating all its closest neighbors around unknown data. It is also known as the distance-based approach. To calculate the distance, the most used approach is Euclidean distance (ED). It is given as,

$$ED = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2} \tag{2}$$

Here, $(X_1, Y_1)$ = Coordinate of the first point, and $(X_2, Y_2)$= Coordinate of the second point.

In this method, choosing the value of $k$ is the most important part. If a small value of k is selected, it results in an overfitting problem, and the model will take all the training dataset, including the noise, and perform poor for the test dataset, while, if a bigger value of $k$ is chosen, it results in underfitting problems and becomes difficult to train the model. Thus, the appropriate selection of $k$ is required to train the model to obtain good results. Though there is no particular approach for determining the value of $k$, multiple values of $k$ can be taken, and accuracy can be assessed; whichever value provides higher accuracy should be selected. The value of $k$ is chosen as an odd number to ensure that equality between two classes never occurs. For a wide range of samples, it is preferable to use the vote of the nearest $k$ neighbors [47].

### 3.4.3 Support vector machine (SVM)

SVM has always been a prominent supervised learning approach that classifies linear and non-linear data. It functions by projecting the input vector into a more dimensional feature space and identifying the hyperplane that divides the data points into two groups. It is generally categorized into two types: support vector classification and support vector regression [48]. It allows classification tasks by maximizing the marginal distance that separates both classes while minimizing classification errors. The marginal distance in this context refers to the separation between the decision hyperplane and its closest instance, which is also a member of that particular class, i.e., each data point is plotted with each feature.

Thus, the classification is performed by finding the hyperplane for the differentiation of two classes. Figure 4 depicts an SVM model for PD and HC classes.

The hyperplane is defined as,

$$w^T x = b \tag{3}$$

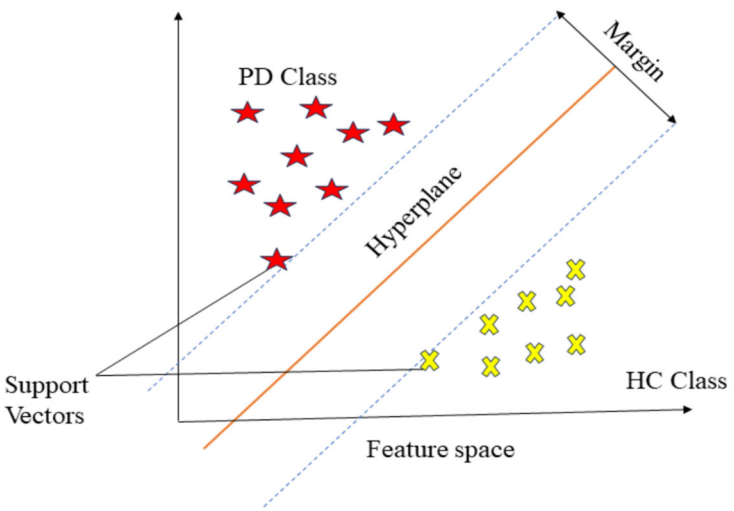



**Fig. 4** SVM Classifier

Where $w$ is the weighted vector, $b$ is bias and $x$ is input. For the PD class the hyperplane is defined as,

$$w^T x + b \geq 0 \tag{4}$$

For the HC class the hyperplane is defined as,

$$w^T x + b < 0 \tag{5}$$

#### 3.4.4 Decision tree (DT)

DT can be defined as a graphical representation of viable outcomes pertaining to a decision and obtained following certain conditions. The adequate number of attributes plays a pivotal role in constructing a decision tree that classifies each object in the training set correctly [49]. With a tree-like complex and certain data structure, it contains internal nodes that serve as the input variables or attributes, and the decision outcomes are represented by leaves. These nodes and their leaves are then used to execute a plan to obtain the classification goal. These nodes can also be classified into several levels, and the leaves are placed on the bottom level of the corresponding branch. The root node is the primary node of the tree structure. It is related to a flowchart that represents non-leaves as a test on a particular attribute, branches as a result of the test conducted, and leaves denote a class label. Methods like CART, id3, and C4.5 are predominantly used to find the best split for constructing a DT.

In the DT algorithm, the variable with maximum information gain or minimum entropy is considered for splitting the dataset. Entropy and information gain are defined as,

$$E(D) = \sum p_i \log_2 \frac{1}{p_i} \tag{6}$$

$$I_G(D, x) = E(D) - \sum \frac{|D_v|}{|D|} E(D_v) \tag{7}$$

Here, $D$ represents the data point for which entropy is calculated, $x$ represents a specific attribute, represents the proportion of the values in $D_v$ to the number of values in dataset D, and $E(D_v)$ is the entropy of dataset of $D_v$.

#### 3.4.5 Naive Bayes (NB)

It is a probabilistic ML classifier based on the Bayes theorem [50]. NB is recognized as the most straightforward and efficient classifier approach for building ML models for predicting outcomes quickly. According to the NB assumption, the attribute value of one class is independent of another class.

Assuming the dataset contains a number of features $x_i = (x_1, x_2, \ldots, x_n)$ and labelled as $l_i = (l_1, l_2, \ldots, l_n)$. Then according to Bayes theorem, the probability of the predictor class having $x_i$ attributes to be classified is given as,

$$P(L/x_i) = P(L) \prod \frac{P(x_i/L)}{P(L)} \tag{8}$$

Where $P(L/x)$ represents the probability of instance $x$ being in label $L$, and $P(L)$ represents the probability of occurrence of label $L$.

### 3.5 Stacked-based model

Stacking is one of the most renowned ensemble modeling techniques used in ML. To predict better predictions, a number of weak learners are ensembled in a parallel manner, such as to make a combination with meta learners. Algorithm 1 describes the stages involved in the stacked model.

**Algorithm 1** Algorithm of stacked model for PD prediction.

**Input:** Speech dataset, $D = \{d_1, d_2, \ldots, d_n\}$,
Base learners = DT, NB, KNN
Base learner's predictions, $BP = \{BP_1, BP_2, BP_3\}$.
**Output:** Classification of speech signals [PD or HC].

1: Data collection.
2: Data distribution into train and test data.
Training data=0.8
Testing data=0.2
3: **for** each learner $BP = \{BP_1, BP_2, BP_3\}$ **do**
4: Train base learners using training data (D),
5: Optimize the base learners until minimum error is achieved.
6: **end for**
7: Integrate the predictions from base learners.
8: Training of Meta learner.
9: Apply test data on meta learner
10: Return Predicted Results.

This ensemble technique is used to achieve an improved model for output prediction by applying the input of meta-learners and combining multiple weak learners' predictions. While performing stacking, the outputs of sub-models are taken as input and combined to anticipate the desired outcome. Stacking, at times, is also called a stacked generalization. Multiple base learners are trained on the same data and then using meta learner to keep the predictions synthesized. Figure 5 portrays the flow chart of the stacked model.

### 3.6 Performance metrics

The proposed work's performance is assessed using a variety of performance metrics, like accuracy, selectivity, sensitivity, and Area Under Curve (AUC).

#### 3.6.1 Accuracy ($A_c$)

Accuracy is the percentage of correct predictions that the classification model makes. It tells us how often our classifier is right.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{9}$$

Here, *TP* (True Positive) indicates that the patients with PD have been correctly identified; FP (False Positive) depicts that the HCs have been mistakenly identified as patients; TN (True Negative) indicates that the correctly identified HCs are healthy; and *FN* (False Negative) indicates that the patients have been mistakenly identified as HCs.

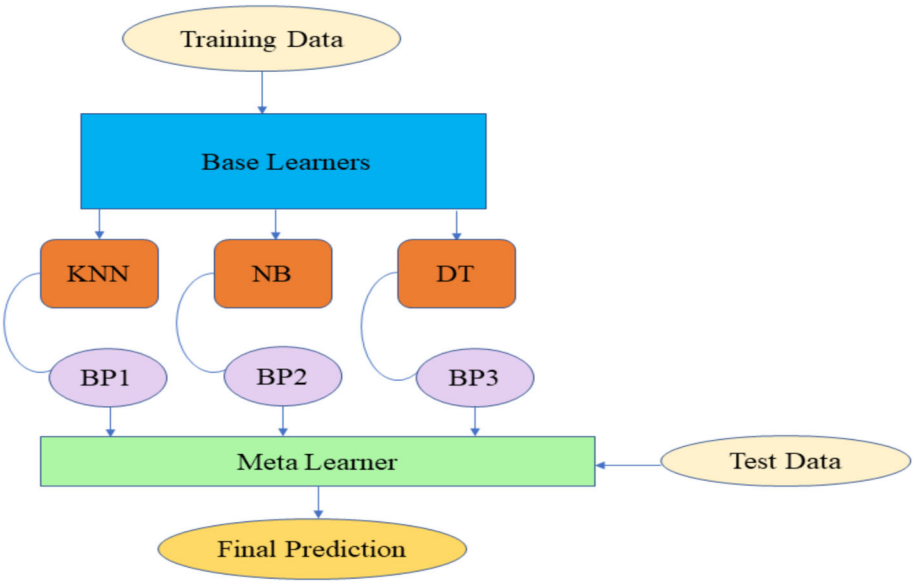



**Fig. 5** Flow chart of stacked-based ensemble model

### 3.6.2 Sensitivity or true positive rate ($S_n$)

Sensitivity is defined as the proportion of PD patients who have been diagnosed as true positives to all PD patients by the classifier.

$$Sensitivity = \frac{TP}{TP + FN} \tag{10}$$

### 3.6.3 Specificity or false positive rate ($S_p$)

It gives an account of those people who are healthy and got predicted the same.

$$Specificity = \frac{TN}{TN + FP} \tag{11}$$

### 3.6.4 Area under the curve (AUC)

The plot between the true positive rate and the false positive rate is known as the Receiver Operating Curve (ROC). A model's ability to accurately predict whether a classification belongs to one of two classes is measured by the AUC- ROC curve. The classifier performs more effectively in terms of prediction when the value of AUC is higher.

### 3.6.5 Precision ($P_r$)

The quality of a successful forecast produced by the model is known as precision. It is calculated by taking the percentage of the positive predicted class and dividing the total

number of positive predictions.

$$Precision = \frac{TP}{TP + FP} \tag{12}$$

#### 3.6.6 F-1 score ($F1_{sr}$)

It is represented as a function of both sensitivity and precision taken together. It is frequently employed when striking a balance between sensitivity and precision is required.

$$F1_{sr} = \frac{2 \times \ P_r \times S_n}{P_r + S_n} \tag{13}$$

## 4 Experimental results

The effectiveness of the proposed work is evaluated in this section. Extensive performance analyses are performed on the UCI PD speech dataset containing the HC and PD classes. The entire experimental work is performed in a Python environment with a CPU running at 3.8GHz, 16 GB of RAM, and a 64-bit operating system.

***Impact of feature selection:*** Prior to feature selection, the dataset had 22 features. To acquire the strongly linked features, the ETC-based feature selection approach is used. The top ten optimal features are selected based on their scores. As the dataset was imbalanced, SMOTE is used to balance the dataset. The dataset is divided into training and testing sections. 80% of the data is used for training, while 20% is used for testing purposes. In order to analyze the performance, popular ML classifiers are utilized.

To increase the performance of classification models, feature selection plays a crucial role by removing extraneous attributes from the dataset. The dataset is condensed using feature selection into a subset that retains the same degree of classification or information as the data in its initial form. The ranking of the most important attributes is determined using the inferences drawn from the proposed method. The ETC-based feature selection procedure is carried out to determine the top ten attributes. These features are used as input to different ML classifiers, which distinguish them into speech samples from healthy or PD patients. Each classifier included in the model has undergone extensive training using several feature combinations. The results of the single classifiers without feature selection are presented in Table 2.

**Table 2** Performance comparison of different ML algorithms without feature selection for PD detection

| Classifier | $A_c$ (%) | $S_n$ (%) | $S_p$ (%) | $AUC$ (%) | $P_r$ (%) | $F1_{sr}$ (%) |
|---|---|---|---|---|---|---|
| LR | 71.79 | 78.12 | 42.85 | 60.49 | 86.21 | 81.97 |
| KNN | 76.92 | 84.38 | 42.85 | 63.61 | 87.10 | 85.71 |
| SVM | 74.35 | 81.25 | 42.85 | 62.05 | 86.67 | 83.87 |
| DT | 79.49 | 90.62 | 28.57 | 59.60 | 85.29 | 87.88 |
| NB | 61.54 | 56.25 | 85.70 | 70.98 | 94.74 | 70.59 |

**Table 3** Performance comparison of different ML algorithms and proposed stacked model with ETC-based feature selection method for PD detection

| Classifier | $A_c$ (%) | $S_n$ (%) | $S_p$ (%) | $AUC$ (%) | $P_r$ (%) | $F1_{sr}$ (%) |
|---|---|---|---|---|---|---|
| LR | 79.67 | 82.76 | 76.67 | 79.71 | 77.42 | 80 |
| KNN | 89.83 | 79.31 | 100 | 89.65 | 100 | 88.46 |
| SVM | 79.67 | 75.86 | 83.34 | 79.59 | 81.48 | 78.57 |
| DT | 96.61 | 93.10 | 100 | 96.55 | 100 | 96.43 |
| NB | 83.05 | 68.97 | 96.67 | 82.81 | 95.24 | 80 |
| Stacked Model | 98.31 | 96.55 | 100 | 98.27 | 100 | 98.25 |

***Impact of stacked-based ensemble model:*** The stacked model combines the advantages of several distinct models to improve prediction accuracy and performance parameters. A stacked-based ensemble method that consists of DT, NB, and KNN is used after selecting the most relevant features. The selection of the classifier is done on the basis of their individual performance. The top ten features based on their feature score are taken into account for PD detection. Table 3 depicts the proposed model's detection results in the UCI PD speech dataset. The results of the stacked model for classification are compared to various classifiers, such as LR, NB, DT, KNN, and SVM.

As demonstrated in Table 3, the stacked model outperformed the distinct classifiers, with accuracy, sensitivity, specificity, and AUC values of 98.31%, 96.55%, 100%, and 98.27%, respectively. Furthermore, selecting features based on ETC not only improves the performance of the model but also considerably upgrades the overall performance of conventional ML methods. Figure 6 depicts the proposed method's performance metrics comparison with several classifiers.

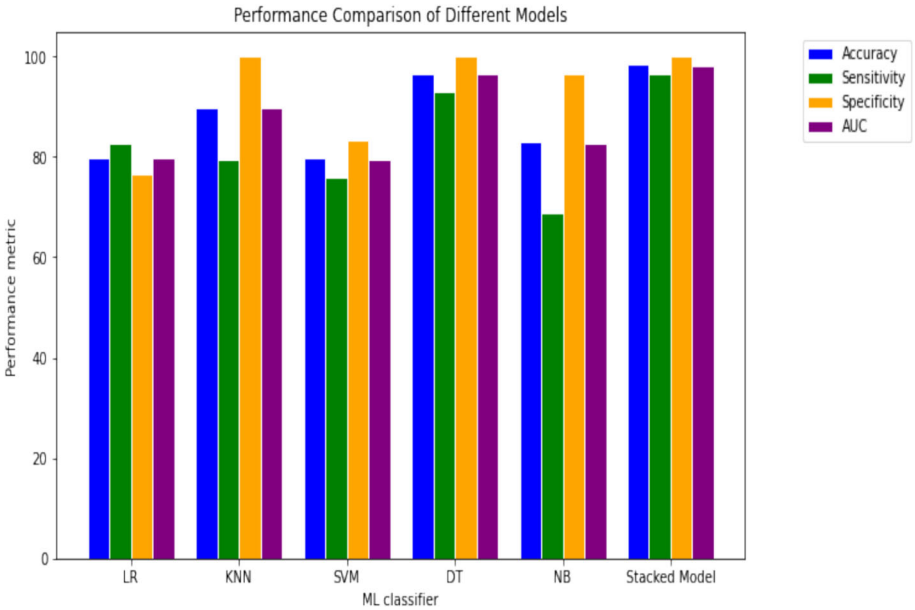


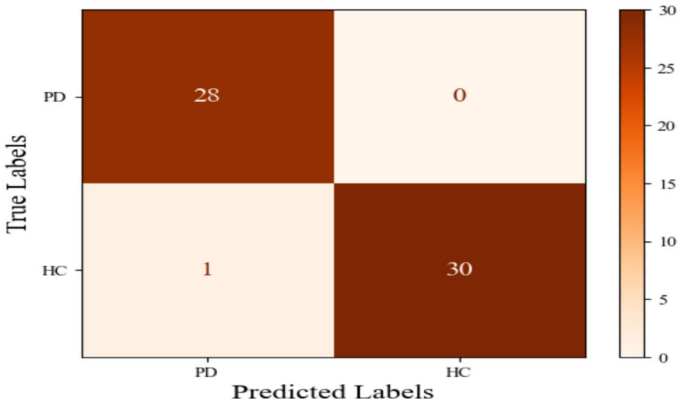

**Fig. 6** Performance metric comparison for PD detection

The confusion matrix and the ROC-AUC curve for the proposed model are illustrated in Fig. 7.

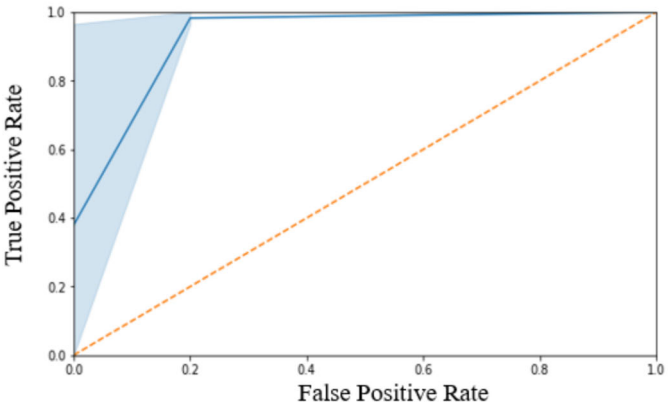
## 5 Discussion

Speech signals are widely used to analyze and detect PD. To detect PD using speech signals, a two-stage stacked-based ensemble technique has been presented in this work. The results of the suggested framework are compared with existing approaches on the same dataset. Table 4 shows the performance comparison of the suggested framework with other existing methods.

In this paper, a feature selection-based decision support system is developed by extracting features from PD patients and HC's speech signals. Distinct feature selection approaches were




(a) Confusion matrix




(b) ROC curve

**Fig. 7** (a) Confusion matrix and (b) ROC curve of the proposed method for PD detection

**Table 4** Performance evaluation of the proposed method in comparison to other existing methods

| Year | Authors | Feature selection | Number of features | Classifier | Results $A_c$ (%) | $S_n$ (%) | $S_p$ (%) | $AUC$ (%) |
|---|---|---|---|---|---|---|---|---|
| 2020 | Thapa et al. [9] | Correlation | 13 | TSVM | 93.90 | 88 | 99 | 93.9 |
| 2020 | Senturk et al. [21] | RFE | 13 | SVM | 93.84 | – | – | – |
| 2020 | Despotovic et al. [23] | ARD | – | Gaussian Process | 96.92 | 90 | 99.29 | – |
| 2021 | Liu et al. [17] | wLDPP | – | ELM | 89.17 | 90.69 | 95 | – |
| 2021 | Ouhmida et al. [22] | mrMR and Relief | 8 | KNN | 97.92 | 97.92 | 100 | 98.26 |
| 2022 | Li et al. [19] | LDA | – | RF | 97.50 | 98.33 | 95 | – |
| 2022 | Haritha et al. [36] | FCM | 10 | RF | 91.83 | – | – | – |
| 2023 | Xue et al. [18] | LDFSF | 5 | Gaussian process | 98.20 | 92 | 99.50 | – |
| 2023 | Rahman et al. [35] | – | – | DNN | 95.41 | 95 | – | – |
| 2023 | Elshewey et al.[37] | – | – | Bo-SVM | 92.3 | 92.3 | – | – |
| 2024 | Shyamala and Navamani [38] | RFE | 15 | XGB | 96.61 | 97.73 | – | 97.73 |
| 2024 | Akila and Nayahi [39] | MASS | – | PCNN | 96.77 | – | – | – |
| 2024 | Saha et al. [40] | – | – | GCN | 97.4 | 97.4 | – | – |
| | **Proposed** | **ETC** | **10** | **Stacked model** | **98.31** | **96.55** | **100** | **98.27** |

To emphasis the values of performance parameters of the proposed method

applied by various researchers. Thapa et al. [9] utilized a correlation-based feature selection approach and attained an accuracy of 93.90% using the TSVM classifier. Senturk et al. [21] incorporated RFE feature selection method in conjunction with SVM classifier and obtained 93.84% accuracy. Authors in [17, 19, 22, 23, 36, 38, 39] used different feature selection approaches such as ARD, wLDPP, mrMR, Relief, LDA, RFE, FCM, and MASS with different ML models achieved an accuracy of 96.92%, 89.17%, 97.92%, 97.50, 96.61%,91.83%, and 96.77% respectively. Recently, Xue et al. [18]. utilized the LDFSF-based feature selection method and obtained an accuracy of 98.20%. Saha et al. [40] developed a Euclidean distance-based- GCN and achieved 97.4% classification accuracy.

Existing studies employed single classifiers along with various feature selection strategies. To enhance the overall performance of PD detection, this paper introduces a two-stage stacked-based approach. The proposed technique differs from the literature in terms of the feature selection method and stacked-based classifier. The suggested model effectively reduces computation time by taking less number of features and outperforms the current methods with a maximum accuracy of 98.31% for PD detection.

## 6 Limitations and future work

The proposed technique employs a stacked ensemble model, which prevents overfitting on training data while providing good accuracy at an accessible computational cost. The optimal features reduce the time required to train the ensemble model. The proposed method only considers cepstrum coefficients and special features retrieved from audio signals; however, other features, such as acoustic and statistical data, can be investigated. The proposed method is intended for three basis learners, which can be expanded, and alternative combinations of base learners can be tested to improve classification performance. Furthermore, the suggested method may be used for bigger datasets containing voice signals that vary in tone, pitch, ethnicity, language, and environmental noise. Also, the proposed approach can potentially be used in PD severity classification.

## 7 Conclusion

This paper proposes a stacked ensemble model to identify PD in its early stage using the UCI PD speech dataset. This model amalgamates the strengths of base learners, thereby overcoming the limitations encountered by single models during classification. The base models in the stack were chosen to be the best fitted for the binary classification tasks, and the meta-learner enhanced predicted accuracy while controlling overfitting. ETC is used to determine the optimal feature subsets, and oversampling using SMOTE helps to address the class imbalance issue, resulting in a more balanced dataset and better model performance. To evaluate the efficacy and reliability of the proposed endeavor, evaluation metrics like accuracy, sensitivity, specificity, AUC, precision, and F-1 score have been calculated. The proposed model exhibits better classification performance compared to the existing state-of-the-art methods.

# Appendix

## 8 A List of Abbreviations

| Abbreviation | Full Form |
|---|---|
| PD | Parkinson's disease |
| ML | Machine learning |
| DL | Deep learning |
| CNN | Convolutional neural network |
| TSVM | Twin support vector machine |
| LR | Logistic regression |
| RF | Random forest |
| ANN | Artificial neural network |
| RFE | Recursive feature elimination |
| mrMR | Maximum relevance minimum redundancy |
| KNN | K-nearest neighbor |
| TQWT | Tunable Q wavelet transform |
| HC | Healthy control |
| HHT | Hilbert-Huang transform |
| DT | Decision tree |
| ETC | Extra tree classifier |
| LDA | Linear discriminant analysis |
| LDFSF | Local dynamic feature selection fusion |
| ARD | Automatic relevance determination |
| NB | Naive Bayes |
| SMOTE | Synthetic minority oversampling technique |
| MDVP | Multidimensional voice program |
| HNR | Harmonic to noise ratio |
| NHR | Noise to harmonic ratio |
| DFA | Detrended fluctuations analysis |
| PPE | Pitch period entropy |
| RPDE | Recurrence period density entropy |
| APQ | Amplitude perturbation quotient |
| DDP | Degree of deviation of periodicity |
| wLDPP | Weighted local discriminative preserving projection |
| GI | Gini Index |
| ED | Euclidean distance |
| FCM | Fuzzy cognitive map |
| MASS | Multi-agent salp swarm |
| PCNN | Parkinson classification neural network |
| BO-SVM | Bayesian optimization-support vector |
| TP | True Positive |
| TN | True Negative |
| FP | False Positive |
| FN | False Negative |
| $A_c$ | Accuracy |
| $S_n$ | Sensitivity |
| $S_p$ | Specificity |
| $P_r$ | Precision |
| $AUC$ | Area under the curve |
| $F1_{sr}$ | F-1 score |

**Funding:** No funding is received for this work.

**Availability of Data and Materials:** The dataset used in this paper is available on the University of California, Irvine (UCI) machine learning dataset repository website. The URL is as follows: https://archive.ics.uci.edu/ml/datasets/parkinsons.

## Declarations

**Conflict of Interest:** All Authors of this work declare no conflict of interest.

**Ethics Approval:** This article does not contain any studies with human participants or animals performed by any of the authors.

## References

1. Naranjo L, Perez CJ, Martin J, Campos-Roca Y (2017) A two-stage variable selection and classification approach for parkinson's disease detection by using voice recording replications. Comput Methods Prog Biomed 142:147–156
2. Earhart GM, Falvo MJ (2013) Parkinson disease and exercise. Comprehensive. Physiology 3(2):833–848
3. Surmeier DJ, Guzmán JN, Sánchez-Padilla J, Goldberg JA (2010) What causes the death of dopaminergic neurons in parkinson's disease? Prog Brain Res 183:59–77
4. Dorsey ER, Bloem BR (2018) The parkinson pandemic-a call to action. JAMA Neurol 75(1):9–10
5. Politis M, Wu K, Molloy S, G. Bain P, Chaudhuri KR, Piccini P (2010) Parkinson's disease symptoms: the patient's perspective. Mov Disord 25(11):1646–1651
6. Anguenot A, Loll P, Neau J, Ingrand P, Gil R (2002) Depression and parkinson's disease: study of a series of 135 parkinson's patients. The Canadian journal of neurological sciences. J Can Sci Neurol 29(2):139–146
7. Sakar CO (2010) Kursun O: Telediagnosis of parkinson's disease using measurements of dysphonia. J Med Syst 34:591–599
8. Avci D, Dogantekin A (2016) An expert diagnosis system for parkinson disease based on genetic algorithm-wavelet kernel-extreme learning machine. Parkinson's disease f2016
9. Thapa S, Adhikari S, Ghimire A, Aditya A (2020) Feature selection based twin-support vector machine for the diagnosis of parkinson's disease. In: 2020 IEEE 8th R10 Humanitarian Technology Conference (R10-HTC), pp 1–6. IEEE
10. Rueda A, Vásquez-Correa JC, Rios-Urrego CD, Orozco-Arroyave JR, Krishnan S, Nöth E (2019) Feature representation of pathophysiology of parkinsonian dysarthria. In: INTERSPEECH, pp 3048–3052
11. Qureshi SA, Raza SEA, Hussain L, Malibari AA, Nour MK (2022) Rehman Au, Al-Wesabi FN, Hilal AM: Intelligent ultra-light deep learning model for multi-class brain tumor detection. Appl Sci 12(8):3715
12. Qureshi SA (2023) Hussain L, Ibrar U, Alabdulkreem E, Nour MK, Alqahtani MS, Nafie FM, Mohamed A, Mohammed GP, Duong TQ: Radiogenomic classification for mgmt promoter methylation status using multi-omics fused feature space for least invasive diagnosis through mpmri scans. Sci Rep 13(1):3291
13. Bhatt K, Jayanthi N, Kumar M (2023) High-resolution superlet transform based techniques for parkinson's disease detection using speech signal. Appl Acoust 214:109657
14. Arias-Vergara T, Vasquez-Correa JC, Orozco-Arroyave JR, Klumpp P, Nöth E (2018) Unobtrusive monitoring of speech impairments of parkinson's disease patients through mobile devices. In: 2018 IEEE international conference on acoustics, speech and signal processing (ICASSP), pp 6004–6008. IEEE
15. Dastjerd NK, Sert OC, Ozyer T, Alhajj R (2019) Fuzzy classification methods based diagnosis of parkinson's disease from speech test cases. Curr Aging Sci 12(2):100–120
16. Karan B, Sahu SS, Orozco-Arroyave JR, Mahto K (2021) Non-negative matrix factorization-based time-frequency feature extraction of voice signal for parkinson's disease prediction. Comput Speech & Lang 69:101216
17. Liu Y, Li Y, Tan X, Wang P, Zhang Y (2021) Local discriminant preservation projection embedded ensemble learning based dimensionality reduction of speech data of parkinson's disease. Biomed Signal Process Control 63:102165
18. Xue Z, Lu H, Zhang T, Xu J, Guo X (2023) A local dynamic feature selection fusion method for voice diagnosis of parkinson's disease. Computer Speech & Language, pp 101536

19. Li Y, Liu C, Wang P, Zhang H, Wei A, Zhang Y (2022) Envelope multi-type transformation ensemble algorithm of parkinson speech samples. Appl Intell, pp 1–22
20. Marar S, Swain D, Hiwarkar V, Motwani N, Awari A: Predicting the occurrence of parkinson's disease using various classification models. In: 2018 International conference on advanced computation and telecommunication (ICACAT), pp 1–5 (2018). IEEE
21. Senturk ZK (2020) Early diagnosis of parkinson's disease using machine learning algorithms. Med Hypotheses 138:109603
22. Ouhmida A, Raihani A, Cherradi B, Terrada O (2021) A novel approach for parkinson's disease detection based on voice classification and features selection techniques. Int J Online Eng 17:111
23. Despotovic V, Skovranek T, Schommer C (2020) Speech based estimation of parkinson's disease using gaussian processes and automatic relevance determination. Neurocomputing 401:173–181
24. Mathur R, Pathak V, Bandil D: Parkinson disease prediction using machine learning algorithm. In: Emerging trends in expert applications and security: proceedings of ICETEAS 2018, pp 357–363 (2019). Springer
25. Vigneswari DA, Aravinth J (2021) Parkinson's disease diagnosis using voice signals by machine learning approach. In: 2021 International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT), pp 869–872. IEEE
26. Xiong Y, Lu Y (2020) Deep feature extraction from the vocal vectors using sparse autoencoders for parkinson's classification. IEEE Access 8:27821–27830
27. Murugappan M, Alshuaib W, Bourisly AK, Khare SK, Sruthi S, Bajaj V (2020) Tunable q wavelet transform based emotion classification in parkinson's disease using electroencephalography. Plos One 15(11):0242014
28. Chintalapudi N, Battineni G, Hossain MA, Amenta F (2022) Cascaded deep learning frameworks in contribution to the detection of parkinson's disease. Bioengineering 9(3):116
29. Chen L, Wang C, Chen J, Xiang Z, Hu X (2021) Voice disorder identification by using hilbert-huang transform (hht) and k nearest neighbor (knn). J Voice 35(6):932–1
30. Lamba R, Gulati T, Jain A (2022) A hybrid feature selection approach for parkinson's detection based on mutual information gain and recursive feature elimination. Arab J Sci Eng 47(8):10263–10276
31. Bchir O (2020) Parkinson's disease classification using gaussian mixture models with relevance feature weights on vocal feature sets. International Journal of Advanced Computer Science and Applications 11(4)
32. Cai Z, Gu J, Wen C, Zhao D, Huang C, Huang H, Tong C, Li J, Chen H (2018) An intelligent parkinson's disease diagnostic system based on a chaotic bacterial foraging optimization enhanced fuzzy knn approach. Computational and mathematical methods in medicine 2018
33. Quan C, Ren K, Luo Z, Chen Z, Ling Y (2022) End-to-end deep learning approach for parkinson's disease detection from speech signals. Biocybernetics Biomed Eng 42(2):556–574
34. Yaman O, Ertam F, Tuncer T (2020) Automated parkinson's disease recognition based on statistical pooling method using acoustic features. Med Hypotheses 135:109483
35. Rahman S, Hasan M, Sarkar AK, Khan F (2023) Classification of parkinson's disease using speech signal with machine learning and deep learning approaches. Eur J Electr Eng Comput Sci 7(2):20–27
36. Haritha K, Judy M, Papageorgiou K, Georgiannis VC, Papageorgiou E (2022) Distributed fuzzy cognitive maps for feature selection in big data classification. Algorithms 15(10):383
37. Elshewey AM, Shams MY, El-Rashidy N, Elhady AM, Shohieb SM, Tarek Z (2023) Bayesian optimization with support vector machine model for parkinson disease classification. Sensors 23(4):2085
38. Shyamala K, Navamani T (2024) Design of an efficient prediction model for early parkinson's disease diagnosis. IEEE Access
39. Akila B, Nayahi JJV (2024) Parkinson classification neural network with mass algorithm for processing speech signals. Neural Computing and Applications, pp 1–17
40. Saha U, Ahamed IU, Ahamed IU, Hossain A-A (2024) Graph convolutional network-based approach for parkinson's disease classification using euclidean distance graphs. In: 2024 7th International conference on informatics and computational sciences (ICICoS), pp 532–537. IEEE
41. Little M, Mcsharry P, Roberts S, Costello D, Moroz I (2007) Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection. Nature Precedings, pp 1–1
42. Liu H, Yu L (2005) Toward integrating feature selection algorithms for classification and clustering. IEEE Trans Knowl Data Eng 17(4):491–502
43. Xue B, Zhang M, Browne WN, Yao X (2015) A survey on evolutionary computation approaches to feature selection. IEEE Trans Evol Comput 20(4):606–626
44. Bhatt K, Jayanthi N, Kumar M (2023) Machine learning based optimal feature selection technique for early stage prediction of alzheimer's disease. In: 2023 Third International Conference on Secure Cyber Computing and Communication (ICSCCC), pp 715–719. IEEE

45. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) Smote: synthetic minority over-sampling technique. J Artif Intell Res 16:321–357
46. Sperandei S (2014) Understanding logistic regression analysis. Biochemia medica 24(1):12–18
47. Cover T, Hart P (1967) Nearest neighbor pattern classification. IEEE Trans Inf Theory 13(1):21–27
48. Gholami R, Fakhari N (2017) Support vector machine: principles, parameters, and applications 1:515–535
49. Quinlan JR (1986) Induction of decision trees. Mach Learn 1:81–106
50. Saritas MM, Yasar A (2019) Performance analysis of ann and naive bayes classification algorithm for data classification. Int J intell Syst Appl Eng 7(2):88–91

**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.

2024 2nd DMIHER International Conference on Artificial Intelligence in Healthcare, Education and Industry (IDICAIEI) | 979-8-3315-2871-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/IDICAIEI61867.2024.10842728

# A Novel Hybrid Model for Efficient Prediction of Diabetes Using Machine Learning

1st Akshit Jain
*Department of Production and Industrial Engineering*
*Delhi Technological University*
Delhi, India
akshitj285@gmail.com

2nd Shagun Jain
*Department of Software Engineering*
*Delhi Technological University*
Delhi, India
shagunjain191172@gmail.com

***Abstract*—Diabetes mellitus, a long-term metabolic disease marked by elevated blood glucose levels, has a substantial negative impact on public health worldwide. Accurate prediction is essential for effective diabetes treatment due to its increasing prevalence. Machine Learning (ML) addresses the issues of this chronic metabolic condition, which provides a promising answer by enabling individualized therapies. This paper presents a comprehensive study on diabetes prediction using ML techniques, focusing on a hybrid approach stacked ensemble models. The research employs the Indian Pima dataset from Kaggle. The primary objective is to assess various ML models with and without preprocessing techniques. Preprocessing involves data cleaning, normalization, and feature engineering to enhance predictive accuracy. Overall, the results underscore the importance of preprocessing in improving predictive performance and highlight the efficacy of hybrid ensemble techniques in diabetic prediction tasks. Moreover, the hybrid approach utilizes a stacked ensemble comprising Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Multilayer Perceptron (MLP) as base models and Logistic Regression as a meta-model. The hybrid approach achieves superior performance compared to individual models with 95.1% accuracy. This methodology effectively combines the strengths of diverse algorithms, mitigating their weaknesses and yielding more robust predictions. It provides substantial benefits by integrating ML algorithms, statistical techniques, and expert knowledge.**

***Keywords*—Machine Learning, Diabetic Prediction, Stacking Ensemble, Artificial Intelligence, Healthcare**

## I. Introduction

Over 422 million people worldwide have diabetes, a chronic illness that causes excessively high blood sugar levels. With its lengthy shadow and potential for problems including heart disease, blindness, and kidney failure, this quiet monster is a hazard [1]. Clinical tests, which are frequently costly, intrusive, and provide results too late, are the standard method to diagnose diabetes. However, there is a ray of light amid the shadows. As a potent tool for diabetes prediction, Machine Learning (ML) provides a proactive strategy for handling this worldwide health emergency [2]. ML algorithms are used to sort vast amounts of data, including genetic information, lifestyle traits, and medical records. They can forecast a person's chance of acquiring diabetes even without symptoms by examining these complex patterns.

Although promising, current diabetes prediction techniques frequently encounter major obstacles that reduce their efficacy. The intricacy and high dimensionality of medical data present challenges for conventional statistical methods and more straightforward machine learning models, which can result in imprecise forecasts and restricted applicability. Researchers have applied techniques such as Support Vector Machines and Logistic Regression, but these methods generally need to improve in identifying non-linear relationships in the data. Furthermore, a lot of earlier research needs to do a better job of addressing the significance of data preprocessing, which is essential for improving model performance and lowering overfitting.

Envision detecting high-risk patients before the onset of the disease and providing them with lifestyle modifications and preventive treatments that may enable them to avoid diabetes altogether. ML can customize treatment regimens by identifying each patient's best action, maximizing results, and reducing needless stress. By concentrating diagnostic tools and therapies on people most likely to benefit, ML can optimize the distribution of healthcare resources and make them more widely available [3]. Though ML has much promise, there are still difficulties. Essential topics to handle include data quality, model interpretability, and guaranteeing forecast fairness. However, pursuing responsible development and research can overcome these obstacles and fully utilize ML to transform diabetes care, sparing millions of lives from the disease's shadow.

By putting forth a novel hybrid model that combines several machine learning techniques to increase prediction accuracy and robustness, this research seeks to close these gaps. The suggested model makes use of a stacked ensemble technique, combining Logistic Regression as a meta-model with Base Models Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Multilayer Perceptron (MLP). By utilizing the advantages of each unique algorithm while minimizing its drawbacks, this hybrid approach aims to produce a prediction system that is more accurate and dependable.

When using ML models, preprocessing is essential because it improves the performance and interpretability of the models by cleaning, normalizing, and engineering features [4]. Preprocessing the data reduces noise, removes superfluous features, highlights significant patterns, and improves prediction accuracy for ML algorithms [5]. Preprocessing also ensures

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:44:12 UTC from IEEE Xplore. Restrictions apply.

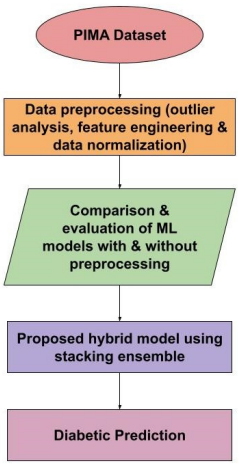



Fig. 1. Research methodology flowchart

that models apply to previously unseen data by reducing the possibility of overfitting and enhancing their capacity for generalization [6]. Once appropriately preprocessed, leveraging ensemble models harnesses the collective strength of diverse predictions, ultimately fortifying the reliability and efficacy of ML applications by further optimizing predictive power [7]. Moreover, ensemble models offer a natural mechanism for handling uncertainty and variability in the data. By considering multiple viewpoints, ensemble methods are less susceptible to overfitting and more resilient to noise in the data [8]. It ensures the model's predictions are more stable and generalizable across different datasets and conditions. The main objectives of this study are:

- To assess various ML models' efficacy in identifying early risk and laying the groundwork for individualized preventative actions, this study compares and evaluates them with and without preprocessing for diabetes prediction.
- In addition to potentially overcoming the shortcomings of current models and enhancing risk identification, this research provides a novel hybrid model using a stacking ensemble.

This study aims to advance ML-driven approaches in diabetic prediction, ultimately paving the way for improved patient outcomes and more effective management of this prevalent chronic condition. Fig 1 shows the methodology of this study. It outlines the step-by-step process followed to achieve the research objectives.

The following sections make up this manuscript: The related work is shown in Section 2. The methodology is outlined in Section 3. The discussions or results are shown in Section 4. The paper is finally concluded in Section 5.

## II. Related Work

Table I provides a comparative overview of previous work on PIMA Dataset using ML models applied for a particular task, likely classification. Additionally, it includes some of the work that has not been addressed in previous studies. Febrian et al. [10] focused on K-nearest Neighbours (KNN) and Naive Bayes, achieving 73.33% accuracy. Their study needed to have comparative analysis among ML models and overlooked ensemble learning techniques, missing insights into model strengths and weaknesses. Similarly, Reza et al. [11] achieved 85.5% accuracy with Support Vector Machine (SVM) but also missed model comparisons and ensemble learning exploration. However, like other studies, they omitted model comparisons and ensemble methods, limiting their analysis and potential for enhanced model robustness. Conversely, Bhoi et al. [12], and Chang et al. [13] achieved lower accuracies but failed to apply ensemble learning methods, potentially improving performance.

With these limitations, this research thoroughly evaluates preprocessing techniques to enhance model performance and generalizability. Additionally, comparative analyses among various ML models provide insights into selecting the most suitable approach. Furthermore, incorporating ensemble learning techniques improves predictive accuracy by leveraging the strengths of multiple models. Overall, by addressing these limitations and integrating advancements in ML methodologies, this study strives to enhance diabetic prediction.

In contrast to earlier research, this study compares the effects of several preprocessing methods on model performance methodically and covers a broad range of preprocessing techniques. This research seeks to improve predictive accuracy and robustness by combining a variety of machine-learning algorithms. The study also deviates from earlier research in that it uses hybrid modelling techniques to maximize prediction and incorporates more recent, cutting-edge machine learning algorithms. By taking a comprehensive approach, the suggested model not only meets but may even surpass the performance of earlier models, offering a more complete tool for diabetes prediction.

## III. Materials and Methods

### A. Dataset Description

The dataset pertains to the diagnosis of diabetes in female Pima Indian adults aged 21 years or older. The National Institute of Diabetes and Digestive and Kidney Diseases owns it. This dataset can be accessed at https://www.kaggle.com/datasets/uciml/pima-indians-diabetes-database. It was obtained from the Kaggle repository. It is composed of multiple medical predictor factors and a target variable called outcome. The goal of the predictor factors is to forecast the presence or absence of diabetes based on diagnostic readings. Among these are age, blood pressure, glucose concentration, skin thickness, insulin levels, BMI, number of pregnancies, and diabetes pedigree function. With 768 observation units, the dataset is a valuable resource for developing diagnostic models and understanding the factors influencing diabetes in this specific population.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:44:12 UTC from IEEE Xplore. Restrictions apply.

TABLE I
PREVIOUS WORK FOR DIABETIC PREDICTION USING PIMA DATASET

| Authors | Year | Model | Accuracy | Comments |
|---|---|---|---|---|
| Khaleel et al. [9] | 2023 | KNN, Naive Byes and Logistic Regression | Logistic Regression 94% | 1. Limit the assessment of the preprocessing techniques. |
| Febrian et al. [10] | 2023 | KNN and Naive Byes | KNN 73.33% | 1. No comparison of various ML models.<br>2. Ensemble learning not applied. |
| Reza et al. [11] | 2023 | SVM | 85.5 | 1. No comparison of various ML models.<br>2. Ensemble learning not applied |
| Bhoi et al. [12] | 2021 | CT, SVM, KNN, Naive Byes, RF, ANN, AdaBoost, Logistic Regression | 76.8% | 1. Ensemble learning is not applied. |
| Chang et al. [13] | 2021 | RF, Naive Byes and DT | RF 79.57% | 1. Ensemble learning is not applied. |

### *B. Data Preprocessing*

Preprocessing is essential to ensure the accuracy and dependability of the analysis in any data-driven study. Cleaning, converting, and arranging raw data into an appropriate format are all part of data preprocessing. It aids in handling missing values, eliminating anomalies, and normalizing or standardizing features. Preprocessing procedures that apply to this study are included in this section.

*1) Outlier Analysis:* This paper uses Local Outlier Factor (LOF) to identify outliers. It determines the local density deviation of a data point from its neighbours. Outliers are defined as points with a noticeably lower density than their neighbours. It is widely used in many domains, including fraud detection, network security, and sensor data analysis.

*2) Feature Engineering:* In feature engineering, researchers apply two techniques: hot encoding and feature aggregation. Through feature aggregation, researchers create multiple categories of various columns and then use one-hot encoding, transforming categorical variables into numerical values. This technique protects from the Dummy variable trap.

*3) Data Normalization:* The researchers in this paper normalize data using Standard Scaler. It is used to transform the data so that its mean is zero and its standard deviation is one. This process is essential when working with algorithms that depend on the size of the input features, such as neural networks, k-nearest neighbours, and support vector machines. The standard scaler's formula is displayed in Equation 1.

$$z = \frac{x - \mu}{\sigma} \tag{1}$$

In this case, the scaled data is denoted by z and x, and the training sample mean and standard deviation are denoted by $\mu$ and $\sigma$, respectively.

### *C. Predictive models*

This section lists all ML models for Diabetic Prediction using the PIMA dataset. Table II highlights the strengths and weaknesses of existing ML models compared to a proposed hybrid model.

1) ***Random Forest (RF):*** a powerful method that creates multiple decision trees during training and yields the class mode (classification) or average prediction (regression) for each tree. It is well known for its robustness against overfitting and strong accuracy across many datasets. [13].
2) ***Support Vector Machine (SVM):*** To maximize the margin between classes, SVM determines the ideal hyperplane in the feature space for class separation. Its versatility and effectiveness in high-dimensional spaces come from its ability to use various kernel functions [11].
3) ***Logistic Regression (LR):*** Despite its name, it is a linear model for binary classification. Using a logistic function, it determines the probability that an input belongs to a given class. It is easy to understand, straightforward, and effective with linearly separable data [9].
4) ***K Nearest Neighbour (KNN):*** A non-parametric technique applied to regression and classification problems. Using a distance measure like the Manhattan or Euclidean distance, KNN determines the majority class of a data point's k nearest neighbours to predict the class of the data point [10].
5) ***Decision Tree (DT):*** A flexible algorithm for supervised learning that can handle tasks involving both regression and classification. It divides the dataset into subsets according to the feature values to maximize homogeneity within the subsets [13].
6) ***Naive Bayes (NB):*** A Bayes-based probabilistic classifier that makes the "naive" assumption that features are independent. Naive Bayes is a simple algorithm that often works well in practice, especially when dealing with high-dimensional datasets [9].
7) ***Extreme Gradient Boosting (XGBoost):*** It constructs a sequence of decision trees, each of which fixes mistakes made by the one before it. It is renowned for its efficiency, quickness, and capacity to manage intricate data relationships [12].
8) **Multilayer Perceptron (MLP):** One kind of feedforward artificial neural network is the MLP, which comprises several layers of nodes (neurons), each connected to the layer above it. It uses nonlinear activation functions to help the model identify intricate patterns in the

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:44:12 UTC from IEEE Xplore. Restrictions apply.

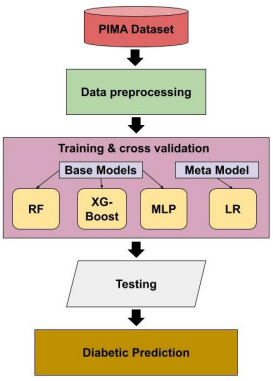



Fig. 2. Flowchart of hybrid model

data, like sigmoid or ReLU.

### D. K-Fold Cross Validation

When applying ML techniques to forecast diabetes, K-fold cross-validation is necessary to evaluate model efficacy reliably. The dataset is divided into K "folds," or roughly equal-sized subsets, to apply the K-fold cross-validation technique. The model is then trained using K-1 folds and validated using the last fold. During the K iterations of this operation, each fold acts as the validation set precisely once. By averaging the performance metrics' overall folds, K-fold cross-validation produces a more accurate measure of the model's generalization performance than a single train-test split. K-fold cross-validation further guarantees that the model's performance is evaluated on multiple different subsets of the data, thereby lowering the risk of overfitting.

### E. Proposed model

To produce a more reliable and accurate final prediction, this study presents a hybrid model that uses a stackingable approach to combine the predictions of multiple base models. In this scenario, three diverse models are employed as base models: RF, XGBoost, and MLP. Each of these models brings unique strengths to the ensemble. Gradient boosting is how XGBoost captures subtle patterns and interactions, while RF is famous for handling complex relationships in data and minimizing overfitting. On the other hand, MLP, as a neural network, can learn intricate nonlinear relationships in the data. This process involves aggregating the predictions from each model and applying a meta-learner to the logistic regression model to combine these predictions effectively. By reducing the shortcomings of individual models, this method produces a predictive model that is more reliable and accurate. Fig 2 shows the flowchart of the proposed hybrid model. The procedure encompasses data preprocessing, model training and validation, and final testing for predicting diabetes. Every individual stage is essential in guaranteeing the dependability and efficiency of the hybrid model in precisely detecting cases of diabetes.

### F. Rationale of Selecting Algorithms for Proposed Hybrid Model

The study selected three models— RF, XGBoost, and MLP—from the pool of possible study algorithms due to their outstanding performance metrics and chose these models after they underwent extensive hyperparameter tuning and training on the training dataset, as they consistently showed the highest accuracy rates throughout the experimentation stages. Their excellent accuracy, precision, and recall metrics across a range of datasets and scenarios demonstrated their suitability for additional research in this study. They further discovered that their interpretability and computational efficiency were advantageous for the intended use, which is in complete accordance with our goal of attaining high accuracy and usefulness.

## IV. Results and Discussion

### A. Parameter Settings

The study uses hyperparameter optimization for each classification model to get the optimum parameters of each attribute corresponding to the model. The Grid Search CV method is used to get the parameters grid. Table III shows the selected hyperparameters used by various ML classifiers for diabetic prediction. This information offers a deeper understanding of the particular configuration choices used for each classifier, such as the learning rate, regularisation parameters, kernel type, or maximum depth of decision trees. The hyperparameters are crucial as they directly impact the model's performance and generalization capacity.

### B. Experimental Results and Analysis

A thorough summary of the accuracy scores attained by different ML models in predicting diabetes, both before and after data preprocessing, is given in Table IV. Table V shows the result of the overall performance metric for diabetic prediction, and Table VI shows the comparison of previous studies with the proposed hybrid model for diabetic prediction. Initially, without preprocessing, the models exhibit a diverse range of accuracies. RF leads the pack with an accuracy of 75.9%, closely trailed by XGBoost at 75.6%. However, Naive Bayes falls short with an accuracy of 76.4%, indicating its susceptibility to the inherent complexities of the raw data. Upon implementing preprocessing techniques, all models showcase significant enhancements in accuracy. This improvement can be attributed to several preprocessing techniques, such as feature scaling, handling missing values, and possibly feature engineering. By refining the data quality, preprocessing optimizes the models' learning process. Notably, both RF and XGBoost experience substantial accuracy boosts post-preprocessing. Random Forest's accuracy surges to an impressive 88.5%, while XGBoost achieves an even more commendable accuracy of 89.2%. The standout performer, however, is the proposed hybrid model, which integrates RF, XGBoost, and MLP. This hybrid approach achieves an accuracy of 84.5% without preprocessing and an outstanding 95.1% with preprocessing. What sets this hybrid model apart is its ability to leverage the unique strengths of each constituent

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:44:12 UTC from IEEE Xplore. Restrictions apply.

TABLE II
STRENGTHS AND WEAKNESSES OF EXISTING ML MODELS AND PROPOSED HYBRID MODEL

| Model | Strength | Weakness |
|---|---|---|
| Random Forest | 1. Robust to overfitting.<br>2. Handles large datasets well.<br>3. Provides feature importance. | 1. It can be prone to overfitting with noisy data.<br>2. It is less interpretable compared to single decision trees. |
| SVM | 1. Effective in high-dimensional spaces.<br>2. Robust to overfitting.<br>3. Works well for clear margin separation. | 1. Requires careful parameter tuning.<br>2. Computationally intensive for large datasets.<br>3. Performance depends heavily on the chosen kernel. |
| Logistic Regression | 1. Simple and easy to interpret.<br>2. Fast training and prediction.<br>3. Effective for binary classification.<br>4. Provides probability estimates. | 1. Assumes a linear relationship between features and log odds.<br>2. Sensitive to multicollinearity.<br>3. Limited ability to capture complex patterns. |
| KNN | 1. Simple and easy to implement.<br>2. No training phase.<br>3. Effective with small datasets. | 1. Sensitive to irrelevant features.<br>2. High computational cost during prediction.<br>3. Poor performance with high-dimensional data. |
| Decision Tree Classifier | 1. Easy to interpret and visualize.<br>2. Handles both numerical and categorical data.<br>3. No need for feature scaling. | 1. Prone to overfitting.<br>2. It can be unstable with small variations in data.<br>3. Less effective for complex patterns compared to ensemble methods. |
| Naive Bayes | 1. Fast and efficient.<br>2. Works well with small datasets.<br>3. Handles irrelevant features well. | 1. Assumes feature independence.<br>2. Limited expressiveness for complex relationships. |
| XGBoost | 1. High predictive accuracy.<br>2. Effective handling of missing data.<br>3. Robust to overfitting. | 1. Computationally intensive.<br>2. Requires careful parameter tuning.<br>3. It can be sensitive to noisy data. |
| MLP | 1. Capable of learning complex relationships.<br>2. Flexible and can model non-linear relationships. | 1. Requires large datasets.<br>2. Prone to overfitting.<br>3. Computationally intensive. |
| **Hybrid model** | 1. Combines strengths of multiple models.<br>2. Higher predictive accuracy.<br>3. Improved robustness and generalizability. | 1. Increased computational complexity.<br>2. Requires careful design and parameter tuning.<br>3. Potentially longer training times. |

TABLE III
HYPERPARAMETERS FOR VARIOUS ML CLASSIFIERS USED FOR DIABETIC PREDICTION

| ML Model | Hyperparameters |
|---|---|
| Random Forest | n_estimators : 500, max_depth : 8, min _samples_split : 2, max_features : 7 |
| SVM | Kernel: 'linear,' 'C: 100, gamma: 'scale' |
| LR | Default Parameters |
| KNN | Default Parameters |
| Decision Tree | Default Parameters |
| Naive Bayes | Default Parameters |
| XGBoost | learning_rate : 0.1, n_estimators : 100, subsample : 1.0, max_depth : 5, min_samples_split : 0.1 |
| MLP | hidden_layer_sizes : (64, 32), activation : 'relu', max_iter : 100, solver : 'adam', random_state : 42 |

TABLE IV
RESULTS OF VARIOUS ML MODELS AND PROPOSED HYBRID MODEL WITH AND WITHOUT PREPROCESSING FOR DIABETIC PREDICTION

| Model | Without preprocessing Accuracy | With preprocessing Accuracy |
|---|---|---|
| Random Forest | 75.9 | 88.5 |
| SVM | 79.13 | 84 |
| Logistic Regression | 78.7 | 84.6 |
| KNN | 72.4 | 84.4 |
| Decision Tree Classifier | 67.7 | 82.5 |
| Naive Bayes | 76.4 | 82.2 |
| XGBoost | 75.6 | 89.2 |
| MLP | 76.8 | 86.6 |
| **Hybrid model** | **84.5** | **95.1** |

model. RF contributes its robustness, XGBoost brings its boosting capabilities, and MLP contributes its DL prowess. Combining these strengths, the hybrid model surpasses the performance of individual models, demonstrating the power of ensemble techniques in diabetes prediction. This significant advancement underscores the efficacy of ensemble methods in tackling complex prediction tasks like diabetes prediction. Integrating diverse models enhances predictive accuracy and offers a more comprehensive understanding of the underlying patterns within the data. Thus, the hybrid model represents a notable breakthrough in predictive analytics, paving the way for more effective and accurate healthcare interventions.

TABLE V
RESULT METRIC OF ML MODELS AND PROPOSED HYBRID MODEL FOR DIABETIC PREDICTION

| Model | Accuracy | Precision | Recall | F1 Score |
|---|---|---|---|---|
| Random Forest | 88.5 | 72 | 60.8 | 65.9 |
| SVM | 84 | 77.5 | 63.9 | 70.1 |
| Logistic Regression | 84.6 | 77.2 | 62.9 | 69.3 |
| KNN | 84.4 | 66.7 | 55.7 | 60.7 |
| Decision Tree Classifier | 82.5 | 58.1 | 55.7 | 56.8 |
| Naive Bayes | 82.2 | 72.8 | 60.8 | 66.3 |
| XGBoost | 89.2 | 69.1 | 67.3 | 68.2 |
| MLP | 86.6 | 84.4 | 75.8 | 79.9 |
| **Hybrid Model** | **95.1** | **93.5** | **87.3** | **90.3** |

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:44:12 UTC from IEEE Xplore. Restrictions apply.

TABLE VI
COMPARISON OF PREVIOUS STUDIES WITH A PROPOSED HYBRID MODEL FOR DIABETIC PREDICTION

| Reference | Model | Accuracy |
|---|---|---|
| Khaleel et al. [9] | Logistic Regression | 94% |
| Febrian et al. [10] | KNN | 73.33% |
| Reza et al. [11] | SVM | 85.5 |
| Bhoi et al. [12] | Logistic Regression | 76.8% |
| Chang et al. [13] | RF | 79.57% |
| **This Study** | **Proposed Hybrid Model** | **95.1** |

### C. Discussion And Application Scenarios

The results suggest that applying preprocessing approaches greatly enhances the predictive accuracy of ML models for diabetic prediction. In addition, the hybrid model performs better than individual models, demonstrating the effectiveness of ensemble learning in improving prediction accuracy. The findings of this study have substantial ramifications for healthcare professionals, researchers, and politicians. Precise diabetic prediction facilitates proactive healthcare management, decreases healthcare expenses linked to problems, and enhances patients' overall quality of life. Moreover, the effective implementation of ensemble techniques in healthcare highlights the possibility of employing comparable methods in other areas of medicine. The proposed hybrid model has a wide range of application scenarios that have significant effects. Precise diabetes prediction in healthcare can result in timely interventions, customized treatment strategies, and enhanced patient outcomes. Moreover, the suggested methodology can be implemented in telemedicine apps to oversee patients' health conditions and offer immediate responses remotely. Through ongoing analysis of patient data, including blood glucose levels and lifestyle factors, the model can provide tailored suggestions for managing and preventing diabetic complications.

## V. CONCLUSION

Early detection of diabetes enables timely intervention, allowing individuals to adopt preventive measures and lifestyle changes that can significantly mitigate the progression of the disease. The study utilizes ML models in predicting diabetes using the PIMA dataset and has demonstrated promising results, particularly with implementing a hybrid approach employing stacking ensemble techniques. This study underscores the significance of preprocessing data before applying ML algorithms, as it significantly impacts the model's accuracy and robustness. Integrating multiple models into a stacked ensemble not only enhances prediction accuracy but also mitigates the limitations of individual models, thus providing more reliable predictions. The thorough incorporation of many methodologies, encompassing data preprocessing, model training, and cross-validation, has significantly enhanced the hybrid model's accuracy and resilience. The results highlight the efficacy of ensemble learning in healthcare contexts, namely in predicting diabetes. Moreover, this research sheds light on the prospects of diabetic prediction using ML. As technology advances and more sophisticated algorithms emerge, predictive models have vast potential for improving accuracy and efficiency. Furthermore, exploring additional datasets and incorporating more diverse features could further enhance the performance of predictive models. Focusing on real-time implementation and integrating with healthcare systems could also facilitate early diagnosis and intervention, ultimately improving patient outcomes. This study is a foundation for future research endeavours in diabetic prediction using ML, emphasizing the importance of continuous refinement in algorithmic techniques and data preprocessing methodologies to achieve more accurate and reliable predictions.

## REFERENCES


[1] V. Jaiswal, A. Negi, and T. Pal, "A review on current advances in machine learning based diabetes prediction," *Primary Care Diabetes*, vol. 15, no. 3, pp. 435–443, 2021. [Online]. Available: https://doi.org/10.1016/j.pcd.2021.02.005

[2] S. C. Mackenzie, C. A. Sainsbury, and D. J. Wake, "Diabetes and artificial intelligence beyond the closed loop: a review of the landscape, promise and challenges," *Diabetologia*, vol. 67, no. 2, pp. 223–235, 2024.

[3] A. Vellido, "The importance of interpretability and visualization in machine learning for applications in medicine and health care," *Neural computing and applications*, vol. 32, no. 24, pp. 18 069–18 083, 2020. [Online]. Available: https://doi.org/10.1007/s00521-019-04051-w

[4] C. V. G. Zelaya, "Towards explaining the effects of data preprocessing on machine learning," in *2019 IEEE 35th international conference on data engineering (ICDE)*. IEEE, 2019, pp. 2086–2090. [Online]. Available: https://doi.org/10.1109/ICDE.2019.00245

[5] S. Alam and N. Yao, "The impact of preprocessing steps on the accuracy of machine learning algorithms in sentiment analysis," *Computational and Mathematical Organization Theory*, vol. 25, pp. 319–335, 2019.

[6] D. H. Deshmukh, T. Ghorpade, and P. Padiya, "Improving classification using preprocessing and machine learning algorithms on nsl-kdd dataset," in *2015 International Conference on Communication, Information & Computing Technology (ICCICT)*. IEEE, 2015, pp. 1–6. [Online]. Available: https://doi.org/10.1109/ICCICT.2015.7045674

[7] O. Sagi and L. Rokach, "Ensemble learning: A survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 4, p. e1249, 2018. [Online]. Available: https://doi.org/10.1002/widm.1249

[8] M. Di Napoli, F. Carotenuto, A. Cevasco, P. Confuorto, D. Di Martire, M. Firpo, G. Pepe, E. Raso, and D. Calcaterra, "Machine learning ensemble modelling as a tool to improve landslide susceptibility mapping reliability," *Landslides*, vol. 17, no. 8, pp. 1897–1914, 2020. [Online]. Available: https://doi.org/10.1007/s10346-020-01392-9

[9] F. A. Khaleel and A. M. Al-Bakry, "Diagnosis of diabetes using machine learning algorithms," *Materials Today: Proceedings*, vol. 80, pp. 3200–3203, 2023. [Online]. Available: https://doi.org/10.1016/j.matpr.2021.07.196

[10] M. E. Febrian, F. X. Ferdinan, G. P. Sendani, K. M. Suryanigrum, and R. Yunanda, "Diabetes prediction using supervised machine learning," *Procedia Computer Science*, vol. 216, pp. 21–30, 2023. [Online]. Available: https://doi.org/10.1016/j.procs.2022.12.107

[11] M. S. Reza, U. Hafsha, R. Amin, R. Yasmin, and S. Ruhi, "Improving svm performance for type ii diabetes prediction with an improved non-linear kernel: Insights from the pima dataset," *Computer Methods and Programs in Biomedicine Update*, vol. 4, p. 100118, 2023. [Online]. Available: https://doi.org/10.1016/j.cmpbup.2023.100118

[12] S. K. Bhoi, "Prediction of diabetes in females of pima indian heritage: a complete supervised learning approach," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 10, pp. 3074–3084, 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:236616594

[13] V. Chang, J. Bailey, Q. A. Xu, and Z. Sun, "Pima indians diabetes mellitus classification based on machine learning (ml) algorithms," *Neural Computing and Applications*, vol. 35, no. 22, pp. 16 157–16 173, 2023. [Online]. Available: https://doi.org/10.1007/s00521-022-07049-z


Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:44:12 UTC from IEEE Xplore. Restrictions apply.

Microsystem Technologies
https://doi.org/10.1007/s00542-024-05829-2

TECHNICAL PAPER

# Advancements in and multi-objective transportation problems: techniques and comparative analysis

Rajpal Rajbhar[1] · Laxminarayan Das[1]

Received: 31 December 2023 / Accepted: 4 December 2024
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2025

**Abstract**
Optimising the allocation of resources from "m" supply sources to "n" demand locations is the goal of a Transportation Problem (TP), which tries to maximise operational efficiency while considering variable costs attached to delivering commodities. Many studies in this field have focused on finding ways to reduce shipping costs, delivery times, and distance. The decision-making situations in multi-objective transportation problems (MOTP) are more intricate since they incorporate numerous objectives. Looking at three case studies with one, two, and three objectives, this paper covers several strategies applied in MOTP. The results show that the suggested method achieves a transport cost of 2968, which is far better than both contemporary single-objective and multi-objective traffic algorithms. Additionally, as compared to other approaches, ours not only improves logistics operations' overall efficiency by cutting down on transit time and fuel usage, but it also provides decision-makers with a wider range of ideal solutions.

## 1 Introduction

There is a special kind of decision-making problem called the Transportation Problem (TP), which in turn has been considered the central nervous system to keep the balance in the economic world from ancient days to today. Shipment of goods from different origins to different destinations can be modeled using this linear programming problem, which seeks to minimize transportation costs while meeting demand at each destination and ensuring that consumers pay a low price per unit (Maity et al. 2016). Finding the appropriate amounts of consignment that should move from one place to another, bearing in mind both supply and demand constraints is what brings about the solution to the TP. The essence is to decrease the total cost of transportation which is summative of all shipment costs (Niluminda and Ekanayake 2022; Guo and Obłój 2019). North-West Corner Method, Least Cost Method, and Vogel's Approximation Method are only a few of the optimization techniques that have been developed for the TP (Marwan 2022; Mishra 2017). There are primarily two categories of transportation problems:

A. **Single-objective transportation problem:** A single-objective transportation problem is one in which the optimization process focuses just on one goal.
B. **Multi-objective transportation problem:** The term "MOTP" describes a transportation problem that involves the optimization of several goals at once (Kacher and Singh 2021).

## 2 Review of literature

Gütmen et al. (2024) stated that the term "multi-objective transportation problem" (MOTP) refers to a specific class of vector maximum (minimum) linear programming problems with multiple competing and incompatible objective functions. In this article, we summarise GP and weighted goal programming (WGP) in terms of MOTP. We then prove that WGP is superior to GP and use it to formulate a solution procedure for MOTP. Finally, we present an application on mobile device screen-panels to demonstrate the applicability.

Chhibber et al. (2021) stated that the notion of fuzzy transportation problem gained prominence as a means to address the inherent uncertainty in various transportation-related parameters, such as fuel prices, weather patterns, product demand and supply, etc. Several researchers have found it to be effective and expanded upon it in their

✉ Rajpal Rajbhar
rajpal_2k18phdam506@dtu.ac.in

[1] Department of Applied Mathematics, Delhi Technological University, New Delhi, India

Published online: 13 January 2025

writings. The current approaches to solving various kinds of fuzzy transportation problems and their extensions are examined in this article. This research has concentrated on intuitive fuzzy transportation problems with either one or more objectives. To learn about the latest development, a look back at the last decade has been conducted. To help understand the pros and cons of various approaches, we have included tabular and graphical comparisons.

Roy et al. (2017) examined a MOFCTP with random rough supply, demand, and objective function parameters. Market globalisation, unexpected developments, etc. may misstate multi-objective fixed-charge transportation problem parameters. For random and difficult scenarios, the multi-objective fixed-charge transportation problem is suggested. The proposed approach uses an anticipated value operator for rough and random parameters. There is also a deterministic model method for the uncertain multi-objective fixed-charge transportation problem. The author employed fuzzy programming, global criterion, and ε-constrained methods to find the best compromise solutions for the model. Compare ideal solutions from several ways for the best response. The ε-constrained method produces optimal solutions and a precise Pareto front. Actual multi-objective fixed-charge transportation scenarios demonstrate the model's applicability and viability. The study improved MOFCTP decision-making with two uncertainty.

Roy et al. (2018) researched on the transportation problem (TP) with interval goals in a context with numerous objectives is detailed in this work. The goal programming (GP) method is generally effective in resolving MOTPs. When the presented problem includes an interval-valued ambition level, the decision maker (DM) may not always be happy with the MOTP solution when using GP. Here, the author offered the methods of utility function into the MOTP and revised multi-choice goal programming (RMCGP) as solutions to this problem, and we compare their results. To demonstrated the practicality and value of this research, study concluded by taking into account a real-life issue on TP.

**Table 2** Initial MOTP table with cost

| Cost | $D_1$ | $D_2$ | $D_3$ | Supply |
|---|---|---|---|---|
| $S_1$ | 4 | 8 | 8 | 76 |
| $S_2$ | 16 | 24 | 16 | 82 |
| $S_3$ | 8 | 16 | 24 | 77 |
| Demand | 72 | 102 | 41 | |

**Table 3** Unbalanced matrix

| Cost | $D_1$ | $D_2$ | $D_3$ | $D_{dummy}$ | Supply |
|---|---|---|---|---|---|
| $S_1$ | 4 | 8 | 8 | 0 | 76 |
| $S_2$ | 16 | 24 | 16 | 0 | 82 |
| $S_3$ | 8 | 16 | 24 | 0 | 77 |
| Demand | 72 | 102 | 41 | | |

Abdelati et al. (2023) emphasis on taking into account supply, demand, and transportation constraints, the transportation problem in operations research seeks to optimise the allocation of goods from various sources to destinations in order to minimise costs. This paper seeks to identify the optimal allocation of truck feet for a private sector Egyptian company by applying the multi-dimensional solid transportation problem approach. Using a multi-objective approach allows decision-makers to have a full range of options to minimise total transportation time or reduce fuel consumption costs. Using three multi-objective techniques—the Zimmermann Programming Technique, the Global Criteria Method, and the Minimum Distance Method—the study investigates the optimal compromise solution. Optimal solutions are derived for time and fuel consumption objectives, offering decision-makers a broad range to make informed decisions.

**Table 1** A general tabular depiction of the MOTP with annotations (Afwat et al. 2018)

| Destination → Source ↓ | $D_1$ | $D_2$ | ... | $D_n$ | Supply ($a_i$) |
|---|---|---|---|---|---|
| $S_1$ | $C_{11}^1, C_{11}^2, \ldots C_{11}^t$ | $C_{12}^1, C_{12}^2, \ldots C_{12}^t$ | ... | $C_{1n}^1, C_{1n}^2, \ldots C_{1n}^t$ | $a_1$ |
| $S_2$ | $C_{21}^1, C_{21}^2, \ldots C_{21}^t$ | $C_{22}^1, C_{22}^2, \ldots C_{22}^t$ | ... | $C_{2n}^1, C_{2n}^2, \ldots C_{2n}^t$ | $a_2$ |
| ⋮ | ⋮ | ⋮ | ⋯ | ⋮ | ⋮ |
| $S_m$ | $C_{m1}^1, C_{m1}^2, \ldots C_{m1}^t$ | $C_{m2}^1, C_{m2}^2, \ldots C_{m2}^t$ | ... | $C_{mn}^1, C_{mn}^2, \ldots C_{mn}^t$ | $a_m$ |
| Demand($(b_j)$ | $b_1$ | $b_2$ | ... | $b_n$ | |

$C_{ijMAX}$ = individually columns or row's maximum $C_{ij}$ value, $C_{ijMIN2}$ = individually columns or row's second smallest $C_{ij}$ value

**Table 4** Balanced matrix

| Cost | $D_1$ | $D_2$ | $D_3$ | $D_{dummy}$ | Supply |
|---|---|---|---|---|---|
| $S_1$ | 4 | 8 | 8 | 0 | 76 |
| $S_2$ | 16 | 24 | 16 | 0 | 82 |
| $S_3$ | 8 | 16 | 24 | 0 | 77 |
| Demand | 72 | 102 | 41 | 20 | |

**Table 5** $D_{dummy}$ row allocation

| Cost | $D_1$ | $D_2$ | $D_3$ | $D_{dummy}$ | Supply |
|---|---|---|---|---|---|
| $S_1$ | 4 | 8 | 8 | 0 20 | ~~76~~ 56 |
| $S_2$ | 16 | 24 | 16 | 0 | 82 |
| $S_3$ | 8 | 16 | 24 | 0 | 77 |
| **Demand** | 72 | 102 | 41 | ~~20~~ 0 | |

## 3 Research methodology

In this study, we proposed MOTP using different methods. The study follows the following structure: In Sect. 2, the authors provide a little explanation of MOTP. The proposed novel algorithm is explained in Sect. 3, and the latter section is the basis for the anticipated findings and comparison presented in Sect. 4. The report then draws a final judgment and makes recommendations for further study.

### 3.1 Multi-objective transportation problem (MOTP)

In real-world scenarios, many transportation issues do not have a single objective goal. Here, we take into consideration the transportation issues, which are distinguished by the presence of multiple objective functions on the list of issues (Akram et al. 2023).

Amount $x_{ij}$ in the MOTP is to be carried from sources $i$ to destinations $j$ at an expense $C_{ij}$. Here $i = 1,2,\ldots..m$ and $j = 1,2,\ldots...,n$ and also $C_{ij}^{k}$ might represent shipping costs/times/distances, transport risk, or power consumption, among other things. If the problem has a "$t$" number of objectives to minimize transport expenses, then it can be written as (Nomani et al. 2017)

$$\begin{aligned} minZ_r(x) &= (Z_1(x), Z_2(x), \ldots\ldots\ldots, Z_r(x)), x \\ &= (x_1, x_2, \ldots...x_r) \end{aligned}$$

Below is a mathematical formulation of the MOTP with constraints (Maity and Roy 2016):

**Table 6** Remove $D_{dummy}$ row

| Cost | $D_1$ | $D_2$ | $D_3$ | $D_{dummy}$ | Supply |
|---|---|---|---|---|---|
| $S_1$ | 4 | 8 | 8 | ~~0~~ 20 | ~~76~~ 56 |
| $S_2$ | 16 | 24 | 16 | ~~0~~ | 82 |
| $S_3$ | 8 | 16 | 24 | ~~0~~ | 77 |
| **Demand** | 72 | 102 | 41 | ~~20~~ 0 | |

Objective 3 function:

$$\text{Obj. no. } 1 = Z_1(x) = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}^{1}.x_{ij}$$

$$\text{Obj. no. } 2 = Z_2(x) = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}^{2}.x_{ij}$$

$$\text{Obj. no. r} = Z_r(x) = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}^{r}.x_{ij}$$

Constraints:

Supply constraints (Yano 2015):

$$\sum_{j=1}^{n} x_{ij} = a_i \quad i = 1,2,\ldots\ldots\ldots,m$$

Demand constraints:

$$\sum_{j=1}^{n} x_{ij} = b_j \quad j = 1,2,\ldots\ldots\ldots,n$$

$$x_{ij} \geq 0; \quad \forall\, i = 1,2,\ldots\ldots\ldots,m \quad j = 1,2,\ldots\ldots\ldots,n$$

where $x_{ij}$: the allocation from source i to destination j and $x_{ij} \geq 0 for\ all\ i,j$. $c_{ij}$: the distance from source i to destination j. $a_i$: the commodity available on sources 1,2, and 3. $b_j$: the commodity required on demand at 1, 2, and 3 (Irine et al. 2018) (Table 1).

### 3.2 The proposed algorithm

We provide two approaches and a step-by-step explanation of each algorithm in this work. You may use these methods to find the best or almost the best solution to balanced or unbalanced multi-objective transportation issues. Firstly,

**Table 7** New generated matrix

| Cost | $D_1$ | $D_2$ | $D_3$ | Supply |
|---|---|---|---|---|
| $S_1$ | 4 | 8 | 8 | 56 |
| $S_2$ | 16 | 24 | 16 | 82 |
| $S_3$ | 8 | 16 | 24 | 77 |
| **Demand** | 72 | 102 | 41 | |

**Table 8** $S_1$ allocated cell

| Cost | $D_1$ | $D_2$ | $D_3$ | Supply |
|---|---|---|---|---|
| $S_1$ | 4 [56] | 8 | 8 | ~~56~~ 0 |
| $S_2$ | 16 | 24 | 16 | 82 |
| $S_3$ | 8 | 16 | 24 | 77 |
| Demand | ~~72~~ 16 | 102 | 41 | |

**Table 9** New generated matrix

| Cost | $D_1$ | $D_2$ | $D_3$ | Supply |
|---|---|---|---|---|
| $S_2$ | 16 | 24 | 16 | 82 |
| $S_3$ | 8 | 16 | 24 | 77 |
| Demand | 16 | 102 | 41 | |

**Table 10** $D_1$ became allocated cell

| Cost | $D_1$ | $D_2$ | $D_3$ | Supply |
|---|---|---|---|---|
| $S_2$ | 16 | 24 | 16 | 82 |
| $S_3$ | 8 [16] | 16 | 24 | ~~77~~ 61 |
| Demand | ~~16~~ 0 | 102 | 41 | |

**Table 11** New generated matrix

| Cost | $D_2$ | $D_3$ | Supply |
|---|---|---|---|
| $S_2$ | 24 | 16 | 82 |
| $S_3$ | 16 | 24 | 61 |
| **Demand** | 102 | 41 | |

**Table 12** $S_3$ became allocated cell

| Cost | $D_2$ | $D_3$ | Supply |
|---|---|---|---|
| $S_2$ | 24 | 16 | 82 |
| $S_3$ | 16 [61] | 24 | ~~61~~ 0 |
| **Demand** | ~~102~~ 41 | 41 | |

the following steps may be used to show the least cost method:

Step 1: Verify that the problem that has been provided is balanced.

Step 2: We determine whether or not the number of rows and columns is equal.

If there is a difference between the number of rows and the number of columns, and vice versa. For our matrix to be transformed into a square matrix, the dummy row or dummy column must be included in the equation along with items that have zero cost or time and zero demand or supply.

Step 3: The lowest cost from the complete matrix should be chosen, and the minimal amount of supply or demand should be assigned to it.

Step 4: The row or column whose supply or demand has been fulfilled should be removed, and a new matrix should be prepared.

Step 4: The process should be repeated until all of the allocations have been over.

Step 5: When all of the allocations have been completed, the next step is to write down the allocations and determine the cost of transportation.

**Table 13** New generated matrix

| Cost | $D_2$ | $D_3$ | Supply |
|---|---|---|---|
| $S_2$ | 24 | 16 | 82 |
| **Demand** | 41 | 41 | |

**Table 14** $D_3$ became allocated cell

| Cost | $D_2$ | $D_3$ | Supply |
|---|---|---|---|
| $S_2$ | 24 | 16 [41] | ~~82~~ 41 |
| **Demand** | 41 | ~~41~~ 0 | |

**Table 15** New generated matrix

| Cost | $D_2$ | Supply |
|---|---|---|
| $S_2$ | 24 | 41 |
| Demand | 41 | |

**Table 16** Final allocated cell

| Cost | $D_2$ | Supply |
|---|---|---|
| $S_2$ | 24 | ~~41~~ 0 |
| **Demand** | ~~41~~ 0 | |

Secondly, the steps of the Geometric Mean and Penalty approaches may be shown as follows:

**Table 17** Balanced matrix with allocations

| Cost | $D_1$ | $D_2$ | $D_3$ | $D_{d/d4}$ | Supply |
|---|---|---|---|---|---|
| $S_1$ | 4 [56] | 8 | 8 | 0 [20] | 76 |
| $S_2$ | 16 | 24 [41] | 16 [41] | 0 | 82 |
| $S_3$ | 8 [16] | 16 [61] | 24 | 0 | 77 |
| Demand | 72 | 102 | 41 | 20 | 235 |

Step 1: It is important to check that your MOTP is balanced. A dummy row or column should be added to the MOTP table to make it balanced if it is not already balanced.

Step 2: Calculate the Geometric Mean (GM) of each cell by using Eq. (1) which is shown below, and then generate a new table that contains the GM values.

$$M = \left(\prod_{i=1}^{n} x_i\right)^{1/n} = \sqrt[n]{x_1 \times x_2 \times x_3 \times \ldots\ldots \times x_n} \quad (1)$$

Here; $x_i$ is the unit cost of each objective in each cell.

Step 3: To determine the penalty value for each row and column, the following formula should be utilized:

$$\left|C_{ijMAX} - C_{ijMIN_2}\right| \quad (2)$$

Step 4: Select the maximum penalty value for each row and column, and then assign the applicable minimum (Supply, Demand) to the cell that contains the minimum $C_{ij}$ value information.

Step 5: If there is a tie between the maximum penalty value for a column or row, the penalty value that resembles the lowest $C_{ij}$ value of that column or column should be selected so that the tie could be resolved.

Step 6: If the allocation from the preceding row matches the supply at the origin, then cross out that row. Cross out the appropriate column if it fulfills the condition there.

Step 7: If there is sufficient supply at each origin and sufficient demand at each destination, the process should be stopped. If not, go back and do the previous steps.

Step 8: Determine the applicable effective cost values for each goal using a MOTP allocation table.

## 4 Result and analysis

This section uses the proposed methods, to provide an example of an imbalanced transportation problem.

Example 1: Unbalanced transportation problem using LCM.

The example below shows a MOTP with a single goal: cost (Table 2).

First, we find out the total supply and total demand

$$Total\ supply = 76 + 82 + 77 = 235$$

$$Total\ Demand = 72 + 102 + 41 = 215$$

Here we see our total supply is greater than the total demand which means our matrix is unbalanced. So, we now follow step 2 (Table 3).

Here we don't have a value of Demand in $D_{dummy}$ Column for balance the matrix. In this case, we can effectively control supply and demand by reducing total demand to total supply, because our total supply exceeds total demand. The result of such an operation is a row value of 20. Table 4 is now balanced.

We will find the cell with the lowest value in this whole matrix because we are utilizing the least-cost method.

The lowest value in this matrix is 0 in this case. So, take that cell and assign the least of demand or supply, i.e. 20 here (since supply is 76 and demand is 20). Checking the first row rather than the final column since we are assigning 20 in the cell for demand because it is the smallest. Subtraction of the assigned value (i.e. 20) from the associated supply and demand (Table 5).

This row or column should be removed from the matrix and a new matrix should be prepared, as shown below, since we fulfilled the demand or supply for that row or column, respectively (Tables 6, 7):

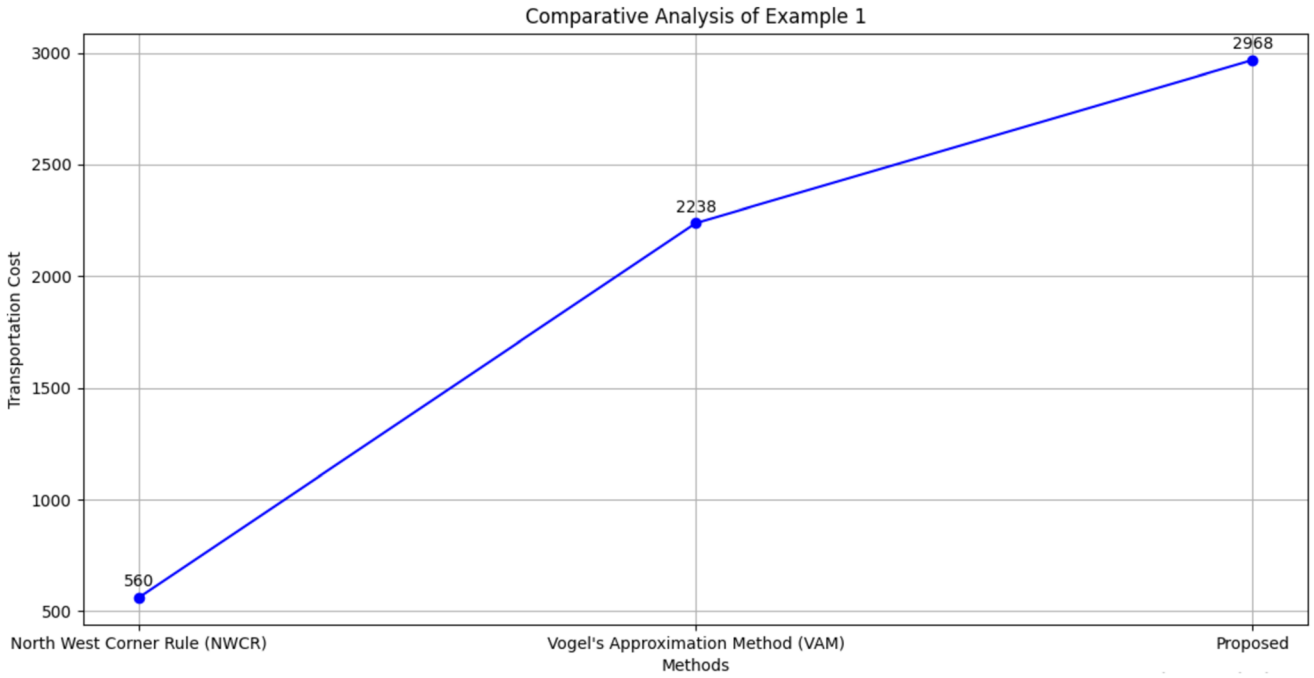



**Fig. 1** Analyzing Example 2 with various approaches

**Table 18** Comparison of Example 1 using various approaches

| Methods | Transportation cost |
|---|---|
| North West Corner Rule (NWCR) (Niluminda and Ekanayake 2023) | 560 |
| VAM (Alardhi et al. 2022) | 2238 |
| Proposed | 2968 |

**Table 19** Initial MOTP table with cost and time

| Time, cost | **D1** | **D2** | **D3** | Supply |
|---|---|---|---|---|
| **S1** | (13,14) | (15,15) | (16,10) | 17 |
| **S2** | (7,21) | (11,13) | (2,19) | 12 |
| **S3** | (19,17) | (20,26) | (9,9) | 16 |
| Demand | 14 | 8 | 23 | |

**Table 20** GM of objectives (cost, time)

| Geometric mean | **D1** | **D2** | **D3** | Supply |
|---|---|---|---|---|
| **S1** | 13.49 | 15.00 | 12.65 | 17 |
| **S2** | 12.12 | 11.96 | 6.16 | 12 |
| **S3** | 17.97 | 22.80 | 9.00 | |
| Demand | 14 | 8 | 23 | |

As soon as that is done, continue the same technique of allocating the lowest value in the newly formed matrix (Table 8), and then check out demand or supply depending on the smallest value (of demand or supply), as shown below, until all allocations have been completed. The minimum cell value in the matrix below is 4.

**Table 21** Penalty method for Example 2

| | D1 | D2 | D3 | Supply | | P1 | P2 | P3 | P4 | P5 |
|---|---|---|---|---|---|---|---|---|---|---|
| **S1** | 13.49 (14) | 15.00 | 12.65 (3) | 17 | | 1.51 | 0.84 | 0.84 | 0.84 | 12.65 |
| **S2** | 12.12 | 11.96 (8) | 6.16 (4) | 12 | 0.16 | | 5.96 | 5.96 | – | – |
| **S3** | 17.97 | 22.80 | 9.00 (16) | 16 | 4.83 | | 8.97 | – | – | – |
| Demand | | | 14 | | | 8 | | | | 23 |
| **P1** | | | 4.48 | | | 7.80 | | | | 3.65 |
| **P2** | | | 4.48 | | | – | | | | 3.65 |
| **P3** | | | 1.37 | | | – | | | | 6.49 |
| **P4** | | | 13.49 | | | – | | | | 12.65 |
| **P5** | | | – | | | – | | | | 12.65 |

**Table 22** Last allocation table of Example 2

| Time, cost | D1 | D2 | D3 | Supply |
|---|---|---|---|---|
| **S1** | 13, 14<br>Maity and Roy 2016) | 15, 15 | 16, 10 | 17 |
| **S2** | 7, 21 | 11, 13<br>Chhibber et al. 2021) | 2, 19<br>Marwan 2022) | 12 |
| **S3** | 19, 17 | 20, 26 | 9, 9 | 16 |
| Demand | 14 | 8 | 23 | |

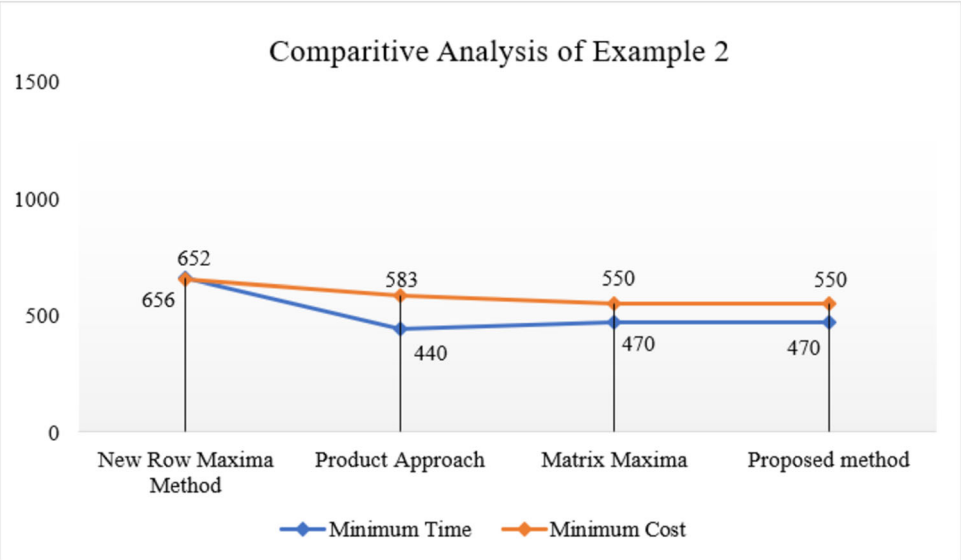



**Fig. 2** Analyzing Example 2 with various approaches

**Table 23** Analysis of methods used in Example 2

| Comparison analysis | Minimum time | Minimum cost |
|---|---|---|
| New Row Maxima Method (Abdelwali et al. 2019) | 656 | 652 |
| Doke's Method (Ekanayake 2023) | 335 | 305 |
| Matrix Maxima (Kumar et al. 2017) | 470 | 550 |
| Proposed method | 470 | 560 |

**Table 24** Primary MOTP table with cost, time, and distance

| Cost, time, distance | D1 | D2 | D3 | D4 | Supply |
|---|---|---|---|---|---|
| **S1** | 21 | 16 | 15 | 13 | 11 |
| | 1 | 2 | 1 | 4 | |
| | 11 | 13 | 17 | 14 | |
| **S2** | 17 | 18 | 24 | 23 | 13 |
| | 3 | 3 | 2 | 1 | |
| | 16 | 18 | 14 | 10 | |
| **S3** | 32 | 27 | 18 | 41 | 19 |
| | 4 | 2 | 5 | 9 | |
| | 21 | 24 | 13 | 10 | |
| Demand | 6 | 10 | 12 | 15 | |

**Table 25** GM of objects (cost, time, distance)

| Geometric mean | D1 | D2 | D3 | D4 | Supply |
|---|---|---|---|---|---|
| **S1** | 6.13 | 7.46 | 6.34 | 8.99 | 11 |
| **S2** | 9.34 | 9.90 | 8.75 | 6.13 | 13 |
| **S3** | 13.90 | 10.90 | 10.53 | 15.45 | 19 |
| Demand | 6 | 10 | 12 | 15 | |

**Table 26** Penalty method for example 3

| | D1 | D2 | D3 | D4 | Supply | P1 | P2 | P3 | P4 | P5 | P6 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **S1** | 6.13 (6) | 7.46 | 6.34 (3) | 8.9 (2) | 11 | 2.65 | 2.65 | 1.53 | – | – | – |
| **S2** | 9.34 | 9.90 | 8.75 | 6.13 (13) | 13 | 1.15 | – | – | – | – | – |
| **S3** | 13.90 | 10.90 (10) | 10.53 (9) | 15.45 | 19 | 4.55 | 4.55 | 4.55 | 0.37 | 0.37 | 10.53 |
| Demand | | 6 | | 10 | | | | 12 | | | 15 |
| **P1** | | 4.56 | | 1.00 | | | | 1.78 | | | 6.46 |
| **P2** | | 7.77 | | 3.44 | | | | 4.19 | | | 6.46 |
| **P3** | | – | | 3.44 | | | | 4.19 | | | 6.46 |
| **P4** | | – | | 3.44 | | | | 4.19 | | | – |
| **P5** | | – | | 10.90 | | | | 10.53 | | | – |

**Table 27** Last allocation table of Example 3

| Cost, time, distance | D1 | D2 | D3 | D4 | Supply |
|---|---|---|---|---|---|
| **S1** | (21, 1, 11) (Kacher and Singh 2021) | (16, 2, 13) | (15, 1, 17) (Guo and Obłój 2019) | (13, 4, 14) (Niluminda and Ekanayake 2022) | 11 |
| **S2** | (17, 3, 16) | (18, 3, 18) | (24, 2, 14) | (23, 1, 10) (Nomani et al. 2017) | 13 |
| **S3** | (32, 4, 21) | (27, 2, 24) (Roy et al. 2018) | (18, 5, 13) (Roy et al. 2017) | (41, 9, 10) | 19 |
| Demand | 6 | 10 | 12 | 15 | |

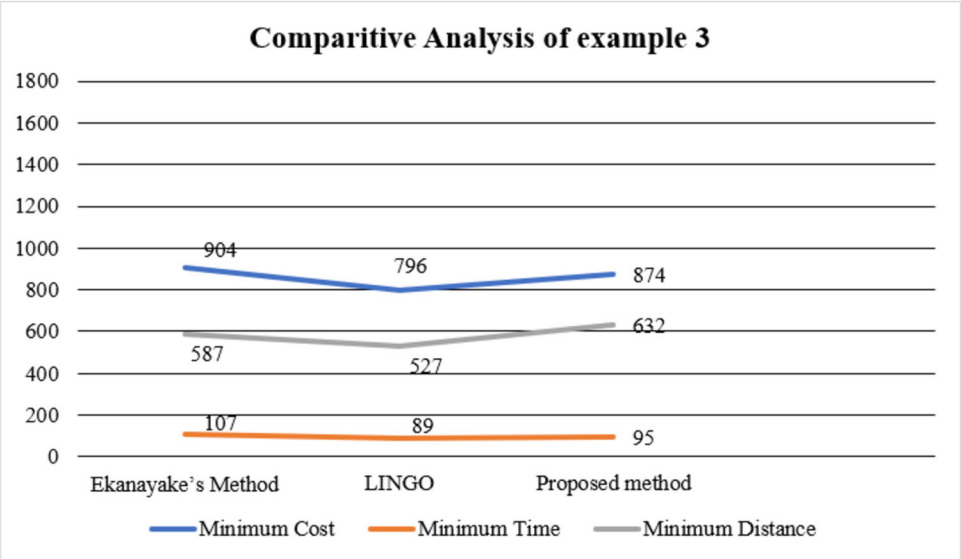



**Fig. 3** Examining Example 3 with various approaches

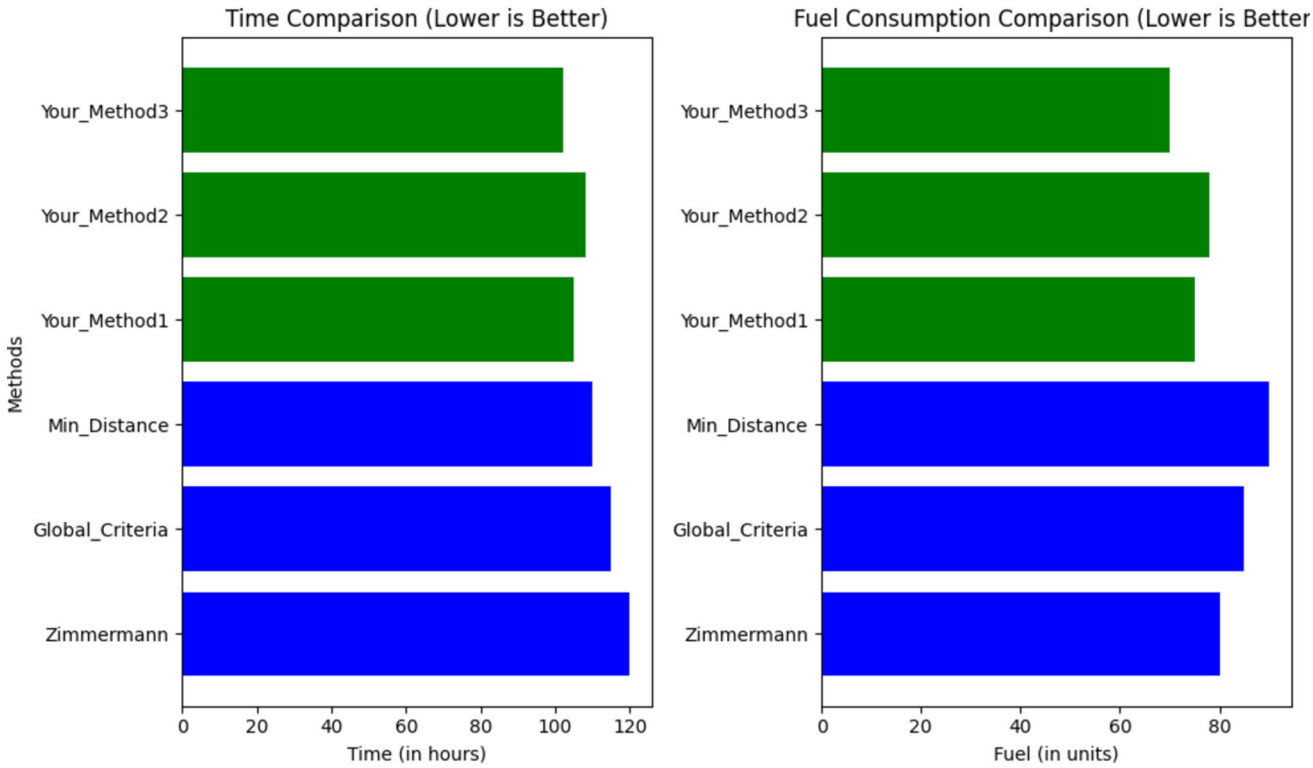



**Fig. 4** Comparison analysis with background layout

**Table 28** Evaluation of Example 3 using various approaches

| Methods | Minimum cost | Minimum time | Minimum distance |
|---|---|---|---|
| Ekanayake's Method (Bhadane et al. 2021) | 904 | 107 | 587 |
| LINGO | 796 | 89 | 527 |
| Proposed method | 874 | 95 | 632 |

Now our $S_1$ became an allocated cell, so we removed that raw. Now our next generated matrix is (Table 9):

Now the minimum cell value is 8, so again repeat the same above step (Table 10):

Now $D_1$ became an allocated cell, so we removed that column. Now our next generated matrix is (Table 11):

In the above table minimum cell value is 16 (Table 12).

Now $S_3$ became an allocated cell, so we removed that raw. Now our next generated matrix is (Table 13):

Now in the above table minimum cell is 16 (Table 14).

Now $D_3$ became an allocated cell, so we removed that column. Now our next generated matrix is (Table 15):

Now in the above table minimum cell value is 24, so repeat the same step and generate the final matrix table (Table 16):

Now finally $D_2$ became an allocated cell. The technique that we have been following is now complete; after all, allocations have been completed, you should build the table with all allocations indicated and compute the transportation cost using the following formula (Table 17):

The solutions are given by,

$x_{11} = 56, \quad x_{14} = 20, \quad x_{22} = 41, \quad x_{23} = 41, x_{31} = 16, x_{32} = 61.$

The total minimum cost is

$$= 4 \times 56 + 0 \times 20 + 24 \times 41 + 16 \times 41 + 8 \times 16 + 16 \times 61$$

$$Transportation\ cost = Rupees2968$$

There is only one objective (Cost) in Example 1 TP. We are aiming to lower the total cost. Figure 1 is the visual depiction of the comparison analysis of example 1, and Table 18 summarizes the results. Our suggested approach performs better than the NWCR and VAM in this comparison (Tables 19, 20, 21, 22).

Example 2 This example represents a MOTP with two objects such as cost and time.

$$Minimum\ Time = (14 * 13) + (3 * 16) + (8 * 11) + (4 * 2) + (16 * 9) = 470$$

$$Minimum\ Cost = (14 * 14) + (3 * 10) + (8 * 13) + (4 * 19) + (16 * 9) = 550$$

The goals of Example 2 of the Transportation Problem (TP) are to reduce expenses as well as time. A graphical depiction of the comparison study is shown in Fig. 2, and the outcomes are compiled in Table 23. We find that our method works better than the Matrix Maxima Method, the New Row Maxima Method, and Doke's technique. Interestingly, there is no difference in the outcomes obtained with our suggested method and the Matrix Maxima Method (Tables 24, 25, 26, 27).

Example 3: This instance represents a MOTP with three objects such as cost, time, and distance.

Minimum Cost = $(6 \times 21) + (3 \times 15) + (2 \times 13) + (13 \times 23) + (10 \times 27) + (9 \times 12) = 874.$

Minimum Time = $(6 \times 1) + (3 \times 1) + (2 \times 4) + (13 \times 1) + (10 \times 2) + (9 \times 5) = 95.$

Minimum Distance = $(6 \times 11) + (3 \times 17) + (2 \times 14) + (13 \times 10) + (10 \times 24) + (9 \times 13) = 632.$

Example 3 highlights the three objectives of MOTP. The main intention of that problem is to cut down on costs, time, and distance. In Table 28, we have analyzed to compare our method with Ekanayake's Method and the Optimum solution obtained by LINGO. Figure 3 represents a line graph of this comparison. Compared to all other existing methods, the proposed method provides a solution that is better in terms of its cost, time, and distance objectives (Fig. 4).

Transportation time and fuel consumption can be significantly reduced by using the strategies suggested here, as contrasted with the findings of Abdelati et al. (2023). For the Zimmermann Programming Technique, our Method 1 reduced transportation time by around 12.5% and fuel consumption by 6.25%. In a similar vein, our Method 2 cut fuel usage by 8.24% and improved time by 6.1% when compared to the Global Criteria Method. With a time reduction of 7.27% and a fuel consumption decrease of 22.2%, Method 3 finally beat the Minimum Distance Method. Better and more cost-effective solutions are provided to decision-makers by these improvements, which demonstrate the excellence of our multi-objective approach.

## 5 Conclusion

One kind of linear programming problem that seeks to optimise costs in response to changes in supply and demand is known as the Transportation Problem (TP). Researchers have looked into ways to optimise shipping prices, time, and distance, but minimising expenses remains the main objective. As a result, new types of transportation problems called Multi-Objective Transportation Problems (MOTP) have emerged, which aim to solve numerous problems at once. Several authors have put

out and discussed several ways to solve MOTP. By comparing and contrasting different optimisation methods, this study reveals notable progress in solving transportation problems with multiple objectives. In comparison to more conventional methods such as the Minimum Distance Method, the Zimmermann Programming Technique, and the Global Criteria Method, the suggested ways significantly reduce transportation time and fuel consumption, according to the results. Decisions in transportation logistics and operations, especially in cases involving multi-objective optimisation, can be improved with the help of this research, which provides better solutions with quantifiable efficiency gains. These results demonstrate the promise of multi-objective strategies for addressing difficult transportation issues, which may benefit from additional development and implementation.

**Author contributions** Rajpal Rajbhar: manuscript design, Laxminarayan Das: supervision.

**Funding** None of the authors received any funding for this research work.

**Data availability** No datasets were generated or analysed during the current study.

## Declarations

**Conflict of interest** The authors declare no competing interests.

## References

Abdelati MH, Abd-El-Tawwab AM, Ellimony EEM, Rabie M (2023) Solving a multi-objective solid transportation problem: a comparative study of alternative methods for decision-making. J Eng Appl Sci 70(1):82

Abdelwali HA, Swilem SM, El Shiaty R, Murad MM (2019) Solving a transportation problem actual problem using excel solver. Int J Eng Techn Res 869(12):13–17

Afwat M, Salama AAM, Farouk N (2018) A new efficient approach to solve multi-objective transportation problem in the fuzzy environment (product approach). Int J Appl Eng Res 13(18):13660–13664

Akram M, Shahzadi S, Shah SMU, Allahviranloo T (2023) An extended multi-objective transportation model based on Fermatean fuzzy sets. Soft Comput. https://doi.org/10.1007/s00500-023-08117-9

Alardhi M, Abdelwali HA, Khalfan AM, Abdelati MH (2022) Using the minimize distance method to find the best compromise solution of multi-objective transportation problem with case study. Int J Traffic Transport Eng 11(2):24–30

Bhadane AP, Manjarekar SD, Dighavkar CG (2021) APBs method for the IBFS of a transportation problem and comparison with North West Corner Method. J Res Appl Math 6(3):109–114

Chhibber D, Srivastava PK, Bisht DCS (2021) From fuzzy transportation problem to non-linear intuitionistic fuzzy multi-objective transportation problem: a literature review. Int J Model Simul 41(5):335–350

Doke DM, Jadhav VA (2015) A solution to three objective transportation problems using fuzzy compromise programming approach. Int J Mod Sci Eng Technol 9(2):9–13

Ekanayake EMUSB (2023) To use an ant colony technique to solve a crispy type bi-and tri-objective transportation problem. J Comput Cogn Eng 2(4):294–303

Ekanayake EMUSB, Daundasekara WB, Perera SPC (2022) New approach to obtain the maximum flow in a network and optimal solution for the transportation problems. Mod Appl Sci 16(1):30

Goel P (2021) New row maxima method to solve multi-objective transportation problem using C-programme and fuzzy technique. Int J Eng Sci Math 10(4):27–36

Guo G, Obłój J (2019) Computational methods for martingale optimal transport problems. Ann Appl Prob 29(6):3311–3347

Gütmen S, Roy SK, Weber G-W (2024) An overview of weighted goal programming: a multi-objective transportation problem with some fresh viewpoints. Central Eur J Oper Res 32(3):557–568

Irine J, Sophia Porchelvi R, Regupathi R, Balasubramanian S (2018) Transportation of fresh water fishes: a case study using multi-objective transportation problem. IJRAR 5(4):115–121

Kacher Y, Singh P (2021) A comprehensive literature review on transportation problems. Int J Appl Comput Math 7:1–49

Kumar BS, Nandhini R, Nanthini T (2017) A comparative study of ASM and NWCR method in transportation problem. Malaya J Mate 5(2):321–327

Maity G, Roy SK (2016) Solving multi-objective transportation problem with interval goal using utility function approach. Int J Oper Res 27(4):513–529

Maity G, Roy SK, Verdegay JL (2016) Multi-objective transportation problem with cost reliability under uncertain environment. Int J Comput Intell Syst 9(5):839

Marwan M (2022) Optimasi biaya distribusi material dengan metode NWC (North West Corner) dan metode vam (Vogel Approximation Method) pada pt. xyz. IESM J (Ind Eng Syst Manage j) 2(2):137–146

Mishra S (2017) Solving transportation problem by various methods and their comparison. Int J Math Trends Technol 44(4):270–275

Niluminda KPO, Ekanayake EMUSB (2022) Kruskal's algorithm for solving both balanced unbalanced acceptable and prohibited route transportation problems. North Am Acad Res 5(12):17–33

Niluminda O, Ekanayake U (2023) Graph colouring to solve both balanced and unbalanced transportation problems. Am J Traffic Transport Eng 8(6):135–144

Nomani MA, Ali I, Ahmed A (2017) A new approach for solving multi-objective transportation problems. Int J Manage Sci Eng Manage 12(3):165–173

Roy SK, Maity G, Weber GW, Gök SZA (2017) Conic scalarization approach to solve multi-choice multi-objective transportation problem with interval goal. Ann Oper Res 253:599–620

Roy SK, Ebrahimnejad A, Verdegay JL, Das S (2018) New approach for solving intuitionistic fuzzy multi-objective transportation problem. Sādhanā 43:1–12

Singh K, Rajan S (2019) Matrix maxima method to solve multi-objective transportation problem with a Pareto optimality criteria. Int J Innov Technol Expl Eng 8(11):1929–1932

Yano H (2015) Multiobjective programming problems involving fuzzy coefficients, random variable coefficients, and fuzzy

random variable coefficients. Int J Uncertain Fuzz Knowl-Based Syst 23(04):483–504

**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.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/387518262

# Aerodynamic Influences on Structures with Varied Plan Configurations

**Article** *in* Journal of Polymer & Composites · December 2024

CITATIONS
0

READS
11

**4 authors**, including:



Rahul Kumar Meena
National Institute of Technology Delhi
**46** PUBLICATIONS **157** CITATIONS
SEE PROFILE



Prateek Roshan
Delhi Technological University
**11** PUBLICATIONS **47** CITATIONS
SEE PROFILE



Manvendra Verma
GLA University
**57** PUBLICATIONS **1,020** CITATIONS
SEE PROFILE

All content following this page was uploaded by Manvendra Verma on 30 December 2024.

The user has requested enhancement of the downloaded file.

STM JOURNALS

Journal of
Polymer & Composites

ISSN: 2321-2810 (Online)
ISSN: 2321-8525 (Print)
Volume 13, Special Issue 1, 2025
DOI (Journal): 10.37591/JoPC

https://journals.stmjournals.com/jopc

Research JoPC

# Aerodynamic Influences on Structures with Varied Plan Configurations

Rahul Kumar Meena [1], Abhishek Prakash Paswan [2], Prateek Roshan [3], Manvendra Verma[4,*]

***Abstract***

*The wind-induced effects on tall buildings demonstrate a level of unpredictability, attributed to the intricate influence of the building's plan shape on the distribution of wind loads. The oscillations experienced by tall structures, both longitudinally and laterally, are influenced by factors such as wind direction, building configuration, height, and structural characteristics. This investigation is centred around examining the impact of wind on tall buildings featuring diverse plan cross-sectional shapes. The analysis employs Computational Fluid Dynamics (CFD) through ANSYS CFX version 2023 R2, scaled at 1:200. The study places particular emphasis on exploring aerodynamic modifications to building shapes, delving into their effectiveness and acceptance within the wind engineering community. Noteworthy findings reveal the inherent positivity of wind forces on the windward face and negativity on the leeward face. The graphical presentation of results, including pressure contours and streamlines in both plan and elevation, adds a visual dimension to the research. Significantly, the study establishes a direct correlation between the plan layout and size of the model and wind pressure distribution on the windward face. An intriguing observation is the similar pressure distribution patterns on both windward and leeward faces across all models. These discoveries contribute valuable insights for structural designers and architects, offering guidance on incorporating various plan shapes into tall building designs. Recognizing the nuanced influence of plan layouts on wind pressure distribution becomes crucial in crafting tall building designs that are not only structurally robust but also architecturally viable.*

**Keywords:** Effects of wind; computational fluid dynamics (CFD); pressure distribution; flow patterns; tall buildings

## INTRODUCTION

Tall structures have several uses, and there are a number of reasons why they are necessary. These fundamental factors—urbanization and population density, for example—indicate the need for tall structures. As cities expand and the population grows, towering buildings offer a useful way to fit more people and activities into constrained spaces. There could not be as much horizontal room for building due to land shortages in highly populated places. The increasing need for residential and commercial properties, coupled with advancements in construction methodologies, has led to a substantial surge in the construction of tall or high-rise structures globally. With the increasing height of buildings and advancements in construction methodologies leading to lighter and more flexible structures, there is a heightened vulnerability to substantial amplitude vibrations (1). The slender nature of a tall building is characterized by the ratio of its height to the smallest plan dimension. Tall

***Author for Correspondence**
Manvendra Verma

[1]Assistant Professor, Department of Civil Engineering, NIT, Delhi, India
[2]PhD Scholar, Department of Civil Engineering, Delhi Technological University, Delhi, India
[3]PhD Scholar, Department of Civil Engineering, Delhi Technological University, Delhi, India
[4]Assistant Professor, Department of Civil Engineering, GLA University, Mathura, Uttar Pradesh, India

Received Date: April 23, 2024
Accepted Date: August 08, 2024
Published Date: December 18, 2024

**Citation**: Rahul Kumar Meena, Abhishek Prakash Paswan, Prateek Roshan, Manvendra Verma. Aerodynamic Influences on Structures with Varied Plan Configurations. Journal of Polymer & Composites. 2025; 13(Special Issue 1): S377–S388p.

© STM Journals 2025. All Rights Reserved 

buildings with a slenderness exceeding five or possessing a fundamental frequency below 0.2 Hz are prone to experiencing significant amplitude vibrations (2). Substantial oscillations of considerable magnitude have the potential to cause harm to structural integrity, possibly resulting in structural damage or, in extreme cases, catastrophic failure (3). Wind-induced vibrations pose concerns for occupant comfort, with perceptions influenced by the level of horizontal acceleration. Various standards exist to gauge acceptable vibrations in tall buildings subjected to wind loads. These standards establish a permissible frequency range for building vibrations, typically ranging from 0.1 to 1.0 Hz (4). Factors like root mean square (RMS) or peak response values, recurrence intervals, and the building's structural characteristics further delineate acceptable frequency levels. Considering the significance of wind-induced vibration in ensuring both structural safety and occupant comfort, it becomes imperative to devise strategies to minimize vibrations in tall buildings subjected to wind forces. High-rise structures often experience cross-dynamic responses caused by vortex-induced vibration and along-wind dynamic response resulting from turbulence-induced buffeting and torsional vibrations originating from uneven loading (5). The predominant factor influencing strength and serviceability design criteria is the cross-wind response. Along-wind vibration is initiated by buffeting loads aligned with the wind direction among the various types of vibrations. Torsional vibration, on the other hand, results from twisting torque along the vertical axis of the building, arising from the uneven distribution of wind loads (6). Traditionally, three main strategies have been used for mitigating wind-induced vibration: adjusting the structural design, incorporating passive or active auxiliary dampers, and making aerodynamic modifications. The goal of modifying the structural design is to reduce wind-induced vibrations by altering properties like mass and stiffness. However, this approach is economically impractical, as it entails increasing the cross-sectional area of the framing system, leading to higher construction costs.

The active control system reduces wind-induced vibrations in tall buildings by modifying structural damping or stiffness through a hydraulic or electro-mechanical actuator system. This system can attenuate vibrations across a broad frequency spectrum, adapting to different excitation conditions. Furthermore, it enables precise, controlled application of force within a short timeframe (4,7). However, a drawback is the substantial energy consumption required to operate the actuators and the potential risk of destabilization if parameters are altered abruptly (8). To mitigate vibrations, a passive control system operates without sensors, actuators, or an external power supply. Damping devices control vibrations induced by wind loads or seismic activity across diverse structures like cables (9), traffic signals(10), and buildings. Among these devices, tuned mass dampers (TMDs) are the preferred choice for managing vibration energy in tall buildings. A TMD comprises a system with a spring and a mass that oscillates out of sync with the structural movement, effectively dissipating vibration energy (11).

An alternative strategy for mitigating wind loads involves the modification of the aerodynamic shape of tall buildings. The shape of an object significantly influences its aerodynamics, playing a crucial role in reducing wind loads on structures. This approach has gained substantial attention in the wind engineering community due to its widespread adoption in recent years. The effectiveness of aerodynamic modifications is typically assessed through wind tunnel experiments and Computational Fluid Dynamics (CFD) simulations. Optimizing aerodynamic shapes has proven to be a key factor in achieving optimal building designs while minimizing stress induced by wind. Current research indicates that the integration of these algorithms with advanced CFD modeling shows promise in overcoming challenges associated with Atmospheric Boundary Layer (ABL) wind tunnel testing. Meena et al.(12) discussed the effects of wind on the bracing system used in various types of multistoried steel buildings. Pal et al.(13) Conducted a comparative analysis of wind-induced mutual interference effects on tall buildings with square and fish plan shapes. Examined the maximum efficiency in terms of wind-induced mutual interference pressure and the exhibited base shear, finding that the square plan-shaped model demonstrated optimal performance under full blockage conditions(14–16). The configuration of a building plays a pivotal role in influencing the dynamic forces induced by wind and subsequent structural responses. Meticulous design of tall building components and shapes is paramount to mitigating wind excitation, ensuring cost-effectiveness. A strategic alteration of flow patterns around structures through passive aerodynamic modifications, often achieved through innovative building shapes, stands out as a potent and versatile construction approach. This technique proves highly

© STM Journals 2025. All Rights Reserved 

effective in substantially minimizing the impact of wind forces. Despite the abundance of wind effects data for regular building models in various international standards, exploring and implementing novel strategies for shaping buildings offer exciting avenues for further advancements in wind engineering (17–23) However, a notable limitation lies in the presentation of wind load data primarily for rectangular models. The dominance of wind effects is intricately linked to the plan shape of tall buildings. While various studies have explored different plan shapes, there remains a gap in addressing corner configurations or modifications in these areas. The current study aims to fill this gap, providing valuable insights for structural designers in considering comprehensive corner configurations and modifications for tall building designs.

This research delves into the examination of wind effects on a building with eight sides, exploring various configurations like the Z-plan shape and a more intricate complex shape. The investigation employs the Computational Fluid Dynamics (CFD) tool ANSYS CFX, specifically the 2023 R2 version, to analyze wind effects. Simulating the wind profile within a virtual wind tunnel using a power-law approach, the study visually represents the wind-generated effects through graphical elements like pressure contours and streamlines in both plan and elevation. Subsequent sections of this study will thoroughly discuss and interpret the findings derived from this investigation.

## NUMERICAL ANALYSIS BY ANSYS CFX

### Model

In the ANSYS CFX Design Module, the modeling process for studying wind effects on tall buildings involves the creation of a scale model with a length scale of 1:200, aligning with wind tunnel manual specifications. This meticulous adherence to the length scale ensures accurate representation and consistency with experimental standards. The model is configured with eight sides, a key parameter in capturing the complexity of the building's geometry. This modeling approach facilitates a detailed examination of wind effects on tall structures, providing valuable insights into their aerodynamic behavior and structural response. Fig. 1 shows the building model with the plan and elevation.

### Boundary Condition

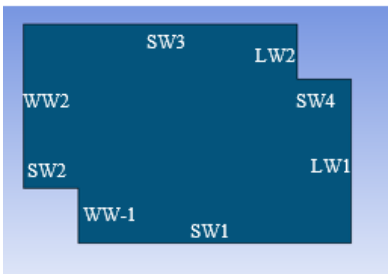
In the ANSYS CFX Design Module, articulating precise boundary conditions is crucial for establishing a well-defined problem that yields a unique solution. The accuracy of the solution is markedly influenced by the clarity and accuracy of the specified boundary conditions. This entails identifying key elements such as the solver, inlet, outlet, walls, and other parameters essential for effectively defining the problem at hand. The meticulous definition of these boundary conditions ensures a robust foundation for obtaining accurate and meaningful results within the simulation. zidan et al.(24) Proposed recommendations for the numerical simulation involve defining the domain and specifying boundary conditions. In this study, the domain pertains to a virtual wind tunnel, constituting a fluid domain. Within the ANSYS CFX software, the fluid properties are pre-established, while the building faces are designated as walls. Additionally, the ANSYS CFX incorporates pre-defined wall functions to streamline the simulation process. These measures collectively contribute to the effective execution of the numerical simulation. Fig 2 shows the computational domain dimensions and boundary conditions.

### Meshing

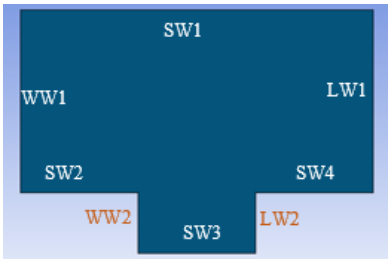
Crafting a high-quality mesh is not a guarantee of an optimal solution, but a poorly constructed mesh invariably leads to suboptimal outcomes. The meshing technique, illustrated in Figure. 3, includes inflation, as depicted in Figure. 3(b) for the Z-shaped building and Figure. 3(c) for the U-shaped model. This inflation is strategically implemented to account for the effects of air in the separation zone. The challenge lies in generating an appropriate mesh, as its intricacy is contingent upon the complexity of the problem at hand. Due to the limitations of governing equations in applying to arbitrary shapes, the computer simulation cannot directly operate on the actual geometry. Mesh elements play a critical role in enabling the predictable and mathematical solution of governing equations within defined volumes. Figure. 3(a) showcases full domain meshing, with (b) and (c) demonstrating building meshing, which is finer compared to other meshing methods. Additionally, Figure. 3(d) illustrates ground meshing, strategically introduced to mitigate unusual flow patterns and ensure the simulation's accuracy. In the

© STM Journals 2025. All Rights Reserved 

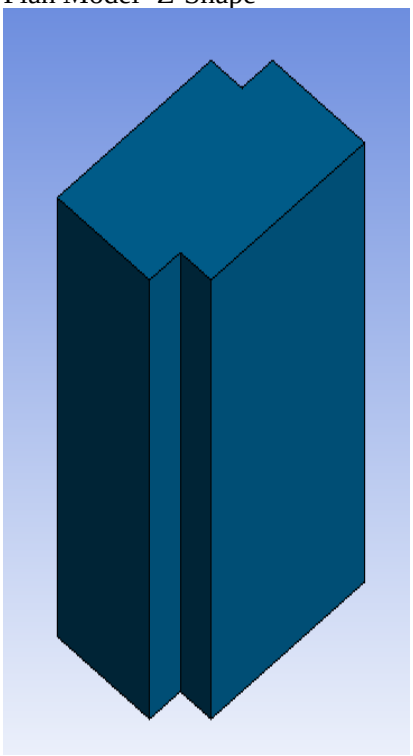
context of investigating wind effects using ANSYS CFX, the meshing strategy plays a pivotal role in achieving accurate and reliable simulation results. Meshing is the process of discretizing the geometric domain into smaller elements, allowing for the numerical solution of governing equations that describe fluid flow and other physical phenomena. In the ANSYS CFX simulation for wind effects on structures, a carefully crafted mesh is crucial for capturing the intricate details of the flow field and ensuring computational efficiency. To facilitate an effective representation of the complex geometry associated with tall buildings, a structured meshing approach is often employed. Structured meshing involves dividing the computational domain into logically arranged elements, providing a systematic grid that aligns with the geometry. This type of meshing is particularly advantageous for simulations involving architectural structures, as it enables a refined representation of surfaces and boundaries. Furthermore, the meshing process involves considerations such as element size, especially in areas of interest where the wind effects are significant, such as around corners, edges, and other aerodynamically sensitive regions. Fine meshing in these critical zones allows for a more accurate capture of pressure distribution and flow patterns. Inflation layers are another essential aspect of meshing, especially in the vicinity of building surfaces. These layers help resolve the boundary layer, where the flow experiences significant changes in velocity and pressure. By incorporating inflation layers, the simulation can more effectively capture the boundary layer effects, enhancing the overall accuracy of the wind effects prediction (25–30).



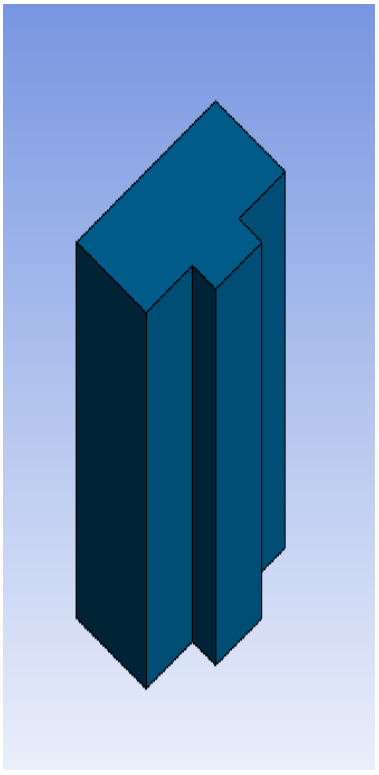

Plan Model- Z-Shape




Plan Model- T Shape



Elevation model - Z Shape



Elevation model- T Shape

**Figure 1.** Building Model with plan and elevation

© STM Journals 2025. All Rights Reserved 

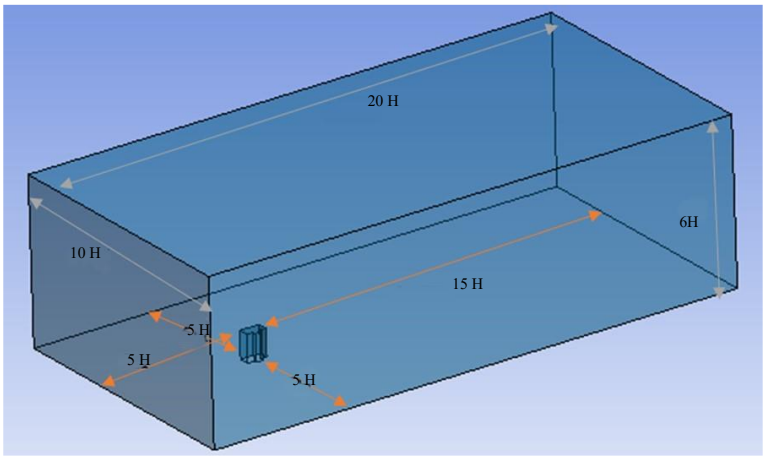
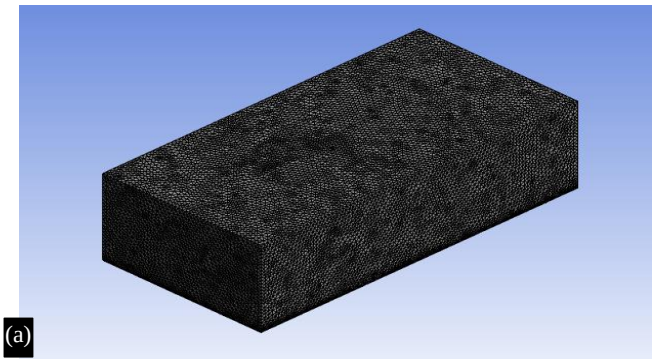
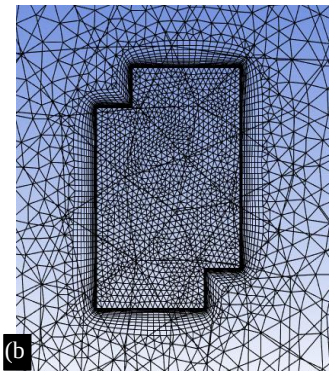


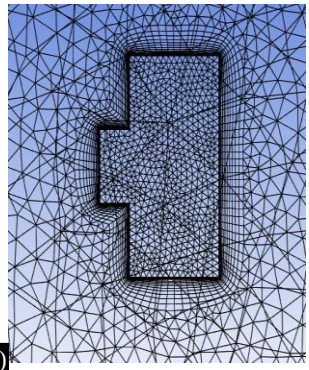

**Figure 2.** Computational domain dimensions and boundary condition.



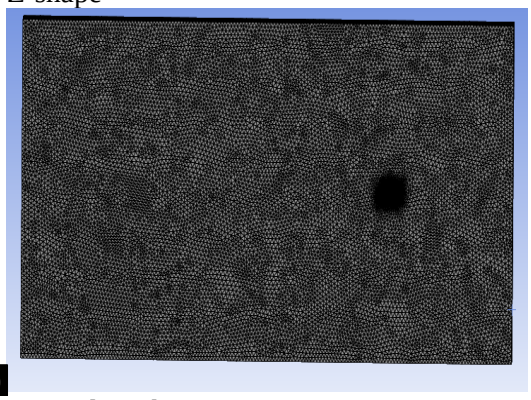
(a) Domain meshing



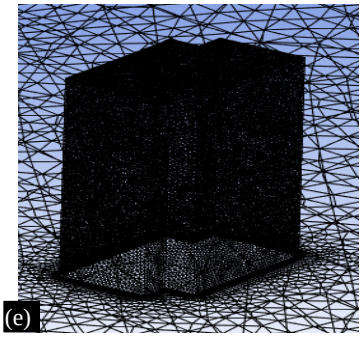
(b) Z-shape



(c) T-Shape model



(d) Ground meshing



(e) Building meshing

**Figure 3.** Meshing for different component of the building.

© STM Journals 2025. All Rights Reserved 

## RESULT AND DISCUSSION

### Velocity Profile and Turbulent Profile

In addition to the Coriolis force, pressure differences in various parts of the Earth cause wind movement. An atmospheric boundary layer forms as a result of the interaction between wind and the Earth's surface, especially close to the ground. Increased wind turbulence is caused by obstructions on the ground. The ground's roughness and the drag produced by nearby projections that impede the wind both affect the vertical profile of wind speed. The elevation at which drag effects stop is indicated by the gradient height, and the accompanying velocity is called the gradient velocity. The atmospheric boundary layer also denotes the height at which terrain affects wind velocity.

In accordance with the Power Law, depicted in Figure 4, the wind speed profile within the atmospheric boundary layer follows the equation (1). This law encapsulates the relationship between wind speed and height, providing a foundational understanding of how wind velocity evolves in the lower atmosphere, considering the impact of various factors on its vertical profile.

$$\frac{U}{U_H} = \left(\frac{Z}{Z_H}\right)^{\alpha} \tag{1}$$

Expressed as U, the horizontal wind speed at elevation Z is represented mathematically. $U_H$ corresponds to the speed at the reference elevation ZH, set at 10 m/s. The parameter α, which fluctuates based on ground roughness and holds a value of 0.147 for terrain category 2, plays a pivotal role in the equation. The reference elevation $Z_H$ is standardized at 1.0 m. This mathematical formulation encapsulates the nuanced relationship between wind speed, elevation, ground roughness, and reference parameters, offering a comprehensive framework for understanding wind behaviour in the specified terrain category.

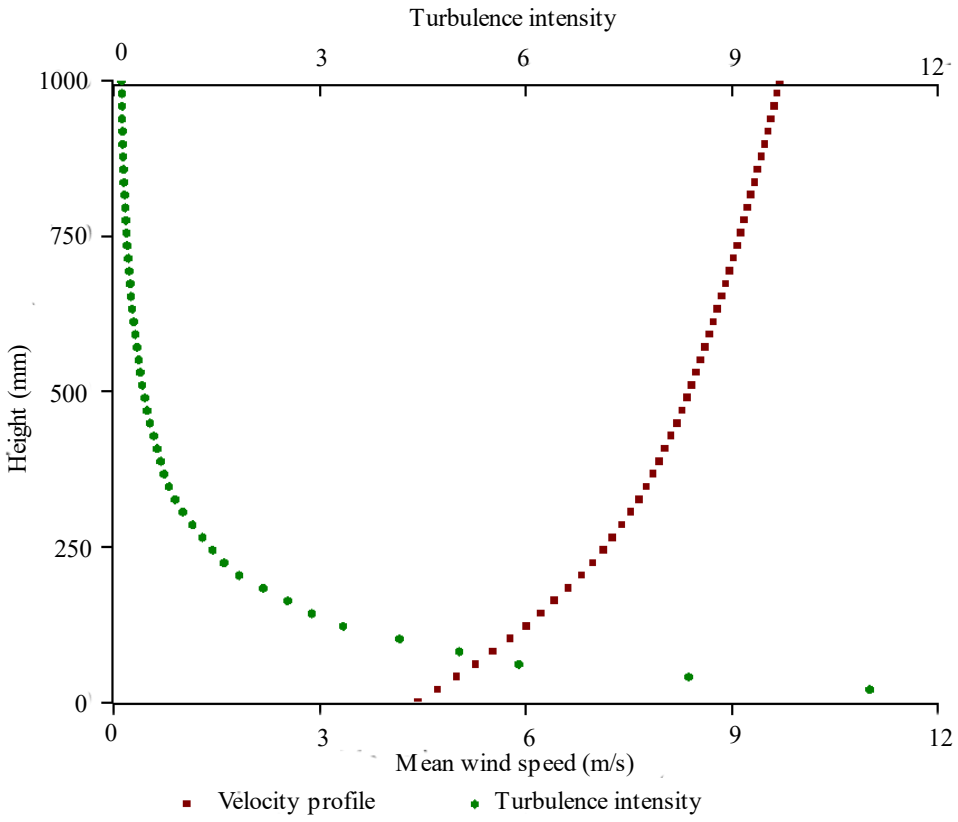



**Figure 4.** Velocity and turbulence intensity.

© STM Journals 2025. All Rights Reserved 

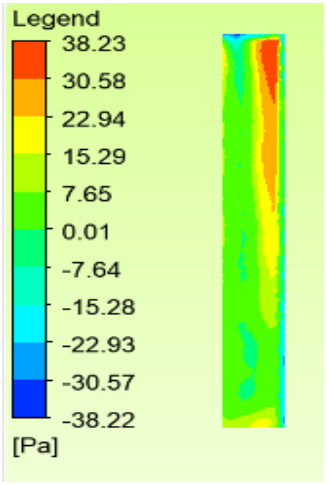
In Figure. 4, it is evident that the wind speed near the surface is notably reduced, primarily due to the friction between the surface and the moving air. The wind profile observed in this study closely resembles the wind profile outlined in ASCE 7 and aligns with findings from prior research in this domain.

**Pressure Contour on Side Face**

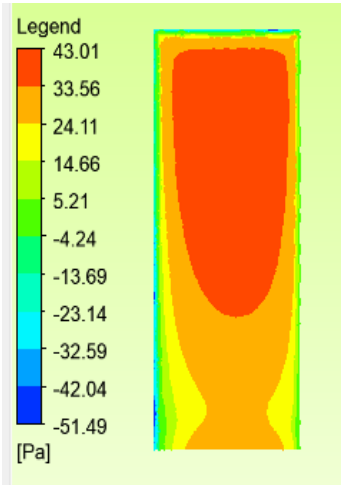
***Pressure contour on Z-shape***

The generation of pressure contours for the Z-plan-shaped building entails the representation of pressure distribution across the entire structure. This visual depiction serves to illustrate the diverse pressure magnitudes at different locations, providing valuable insights into the specific aerodynamic characteristics inherent to the Z-plan shape. The practice of visualizing pressure contours is considered standard in the examination of wind effects on architectural structures and has been widely incorporated in analogous research within the discipline.

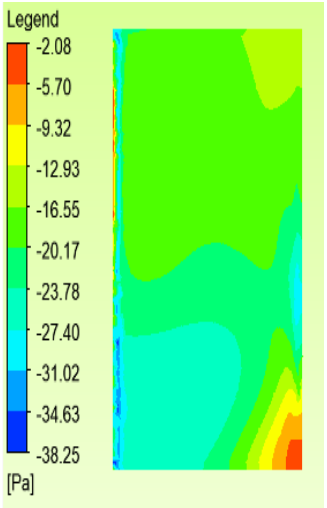
In Figure 5, the pressure on the windward face is denoted as "ww," on the leeward face as "lw," and on the side wall as "sw." Notably, the pressures on the windward faces exhibit a nearly identical nature, while on the leeward face, it is characterized by a negative magnitude. These pressure contours, showcasing the variation in pressure across the elevation of the building model, contribute to a comprehensive understanding of the intricate dynamics influenced by the Z-plan shape.



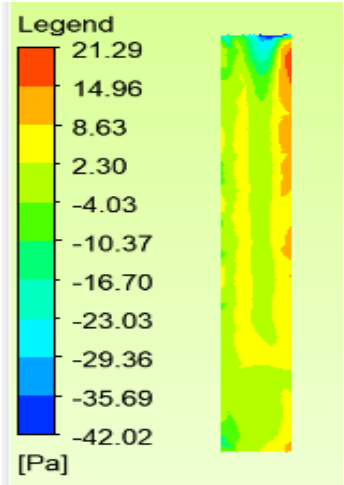

WW-1




WW-2




SW-1




SW-2

© STM Journals 2025. All Rights Reserved 

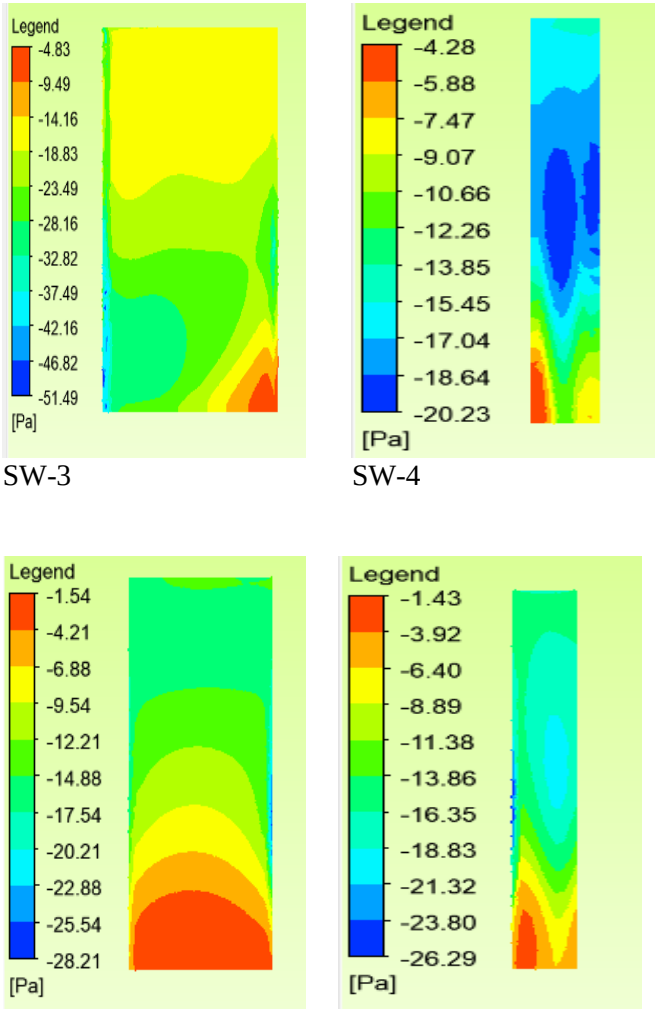



**Figure 5.** Pressure contour for 0 wind for Z plan shape model

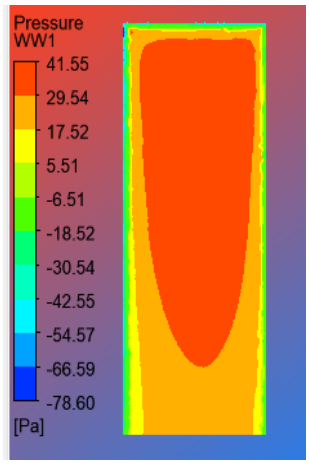
***Pressure contour on T-shape:***

Examining the wind pressure distribution on an irregular plan-shaped building with a T-type configuration involves analyzing how wind forces vary across the structure. This investigation aims to understand the nuanced aerodynamic effects specific to the irregular T-shaped plan. The exploration of wind pressure distribution on irregular plan shapes is a standard practice in the field of wind engineering and has been widely integrated into analogous studies. This study delves pressure contours on Figure 6,into the intricate details of how wind pressure interacts with the irregular T-shaped building, shedding light on the complex aerodynamic behavior associated with this particular architectural configuration.

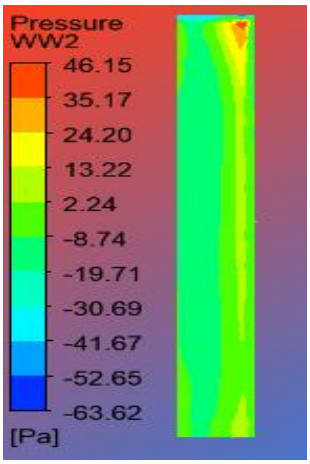
**Pressure Contour on Roof Face**

Investigating the pressure distribution on the roof of the building model entails a detailed analysis of how pressure varies across this specific architectural element. This examination aims to unravel the intricate dynamics of pressure interactions with the roof, providing valuable insights into the aerodynamic characteristics of the structure. The exploration of pressure distribution on building roofs is a standard practice in the realm of wind engineering and has been extensively utilized in analogous studies within the field. This study contributes to a nuanced understanding of how wind pressure is distributed on the roof, offering unique perspectives on the structural response to wind forces in this particular context. The pressure on the roof is increasing in suction as wind is moving from the wind ward side to leeward side and the contours are presented in figure 7 for both the type building model.

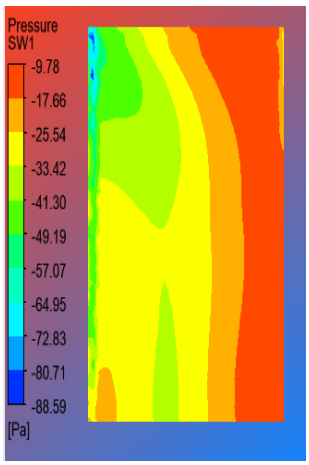
© STM Journals 2025. All Rights Reserved 



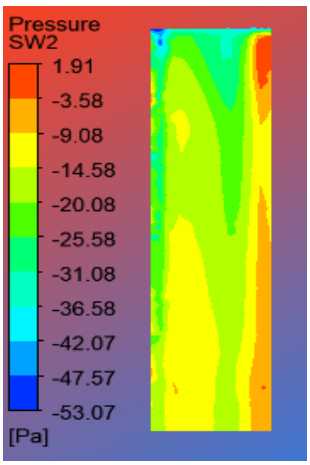

WW1



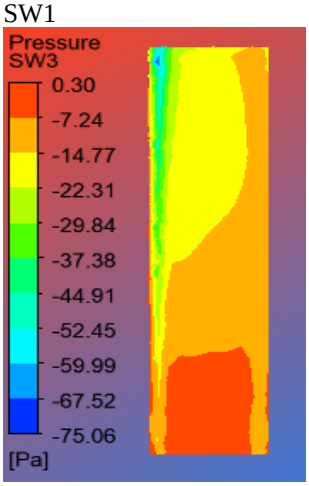

WW2



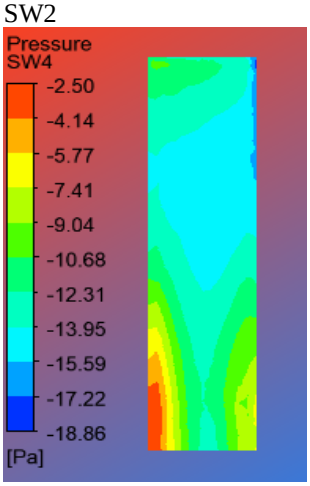

SW1




SW2

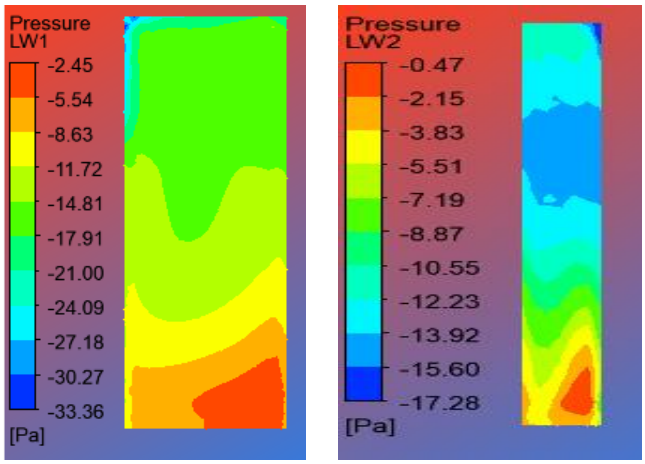



SW3



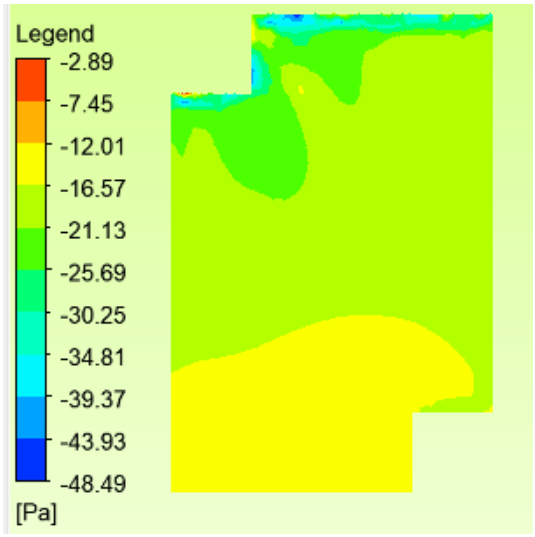

SW4

© STM Journals 2025. All Rights Reserved 

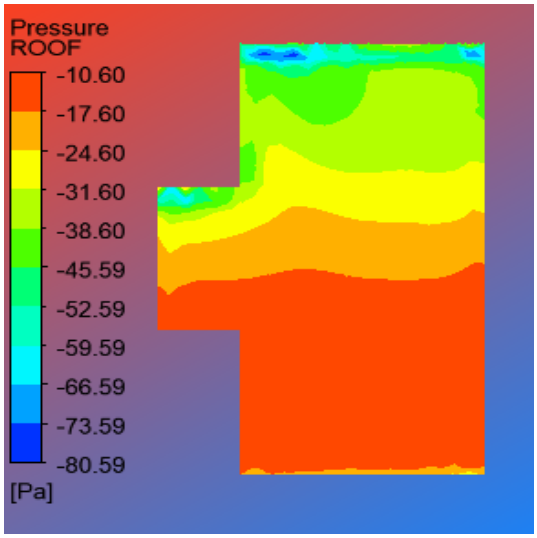



LW1 LW2

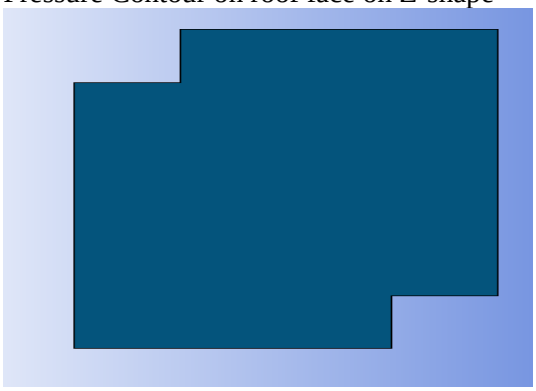
**Figure 6.** Pressure Contours for the T Shape Model



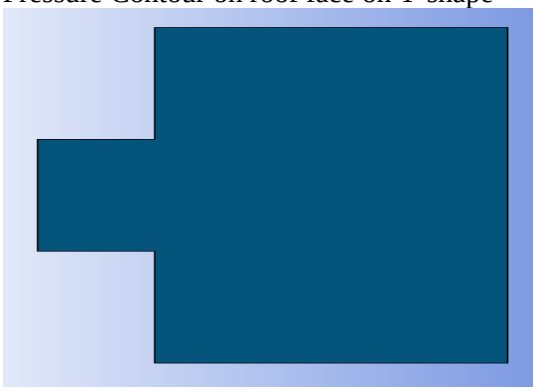

Pressure Contour on roof face on Z-shape




Pressure Contour on roof face on T-shape



Z shape building roof



T Shape building roof

**Figure 7.** Pressure contours on building roof.

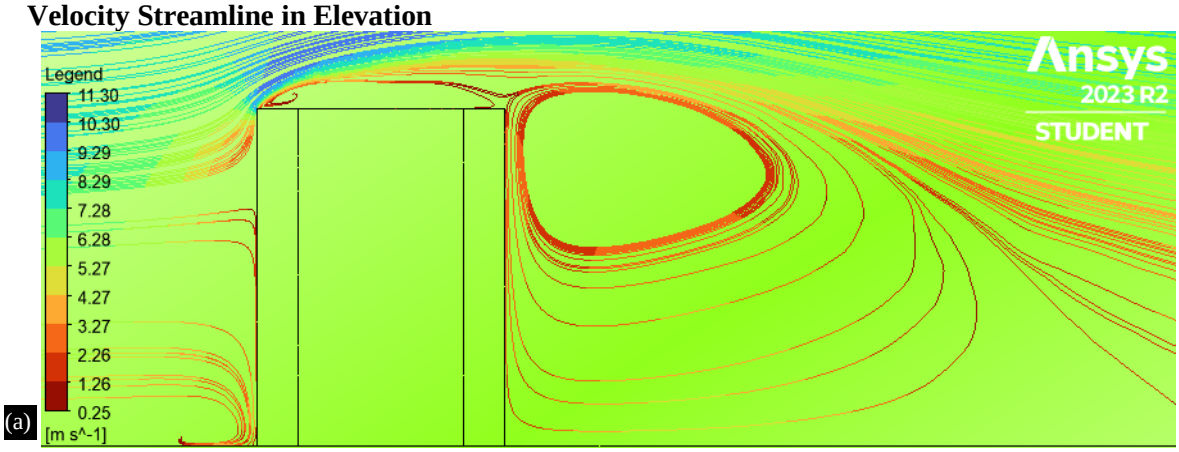
**Velocity Streamlines**

Exploring the velocity streamlines on both the elevation and plan of a high-rise building model serves as a crucial component in the investigation of wind effects. This analysis involves a detailed examination of how wind flows along the surfaces of the building, providing essential insights into the aerodynamic behaviour of the structure. The visualization of velocity streamlines is presented in

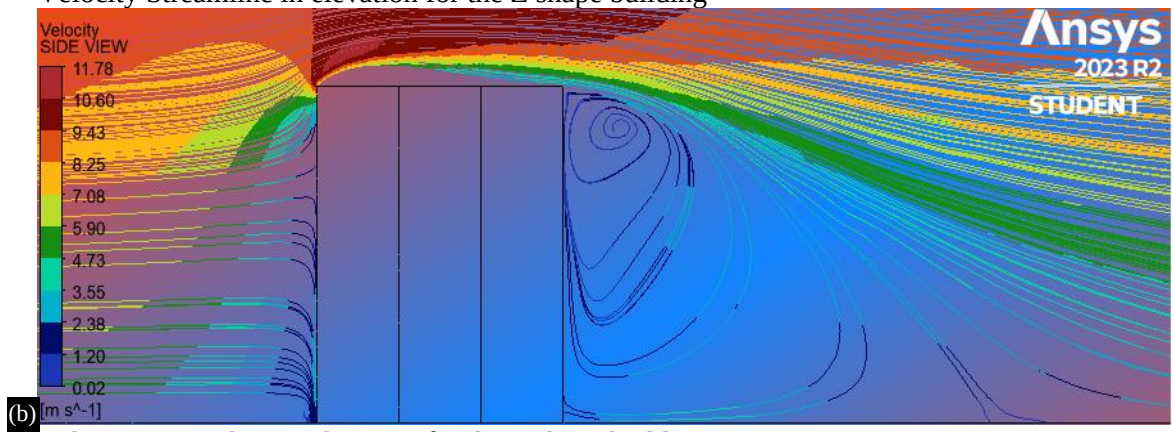
© STM Journals 2025. All Rights Reserved 

elevation in figure 8 and while in the plan in figure 9, it is a standard and widely accepted practice in wind engineering studies, contributing to a comprehensive understanding of how wind interacts with high-rise buildings. This study uniquely contributes to the field by offering detailed insights into the complex flow patterns and velocities on both the vertical and horizontal aspects of the building, enriching our understanding of the wind effects on high-rise structures.

**Velocity Streamline in Elevation**



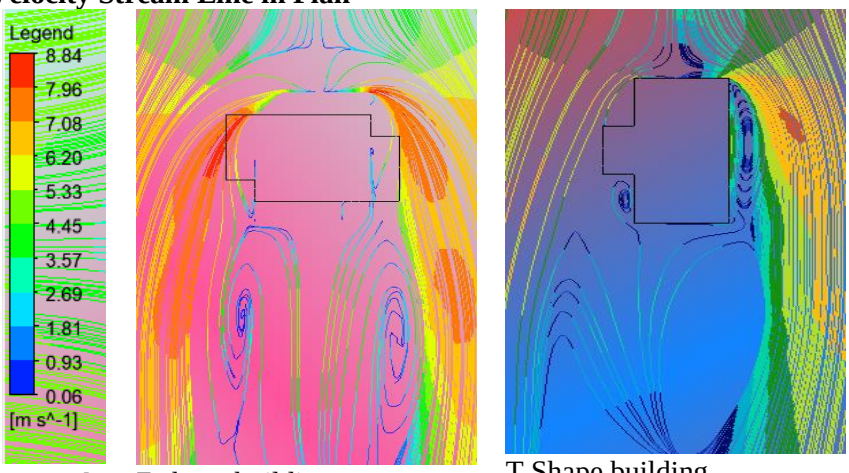

(a) Velocity Streamline in elevation for the Z shape building




(b) Velocity Streamline in elevation for the Z shape building

**Figure 8.** Velocity streamline in elevation for the high-rise building

**Velocity Stream Line in Plan**




Legend Z shape building T Shape building

**Figure 9.** Velocity Streamline in the plan for the high-rise building.

© STM Journals 2025. All Rights Reserved 

## CONCLUSIONS

In this paper, a comparative analysis is conducted on the pressure contour, mean pressure coefficients, and velocity streamlines for high-rise building models with Z and T Shapes at a wind incidence angle of zero degree. The simulation for this research is performed using the k-ε Model. The key discoveries derived from this study are outlined below.

- With the use of CFD, wind impacts can be virtually simulated, offering an economical and effective way to investigate different design options and evaluate how they might affect both low- and high-rise structures.
- Because of CFD's versatility, it is possible to optimise building forms and configurations to improve performance under various wind conditions through iterative design revisions.
- Changes to the building's layout have a major effect on the models' leeward and side faces' suction.
- For both Z-shaped and T-shaped building styles, the windward face is under the most strain, while the side face is under the least.
- The supplied graphics provide a visual representation of velocity streamlines in both plan and elevation.
- The correctness of the results depends on how well the geometry model is meshed and how precisely flow physics is defined, simulating boundary layer wind tunnel conditions.
- The leeward face wind pressure distribution in this research emphasises the formation of vorticity, suggesting higher degrees of turbulence.

## REFERENCES

1. Jafari M, Alipour A. Methodologies to mitigate wind-induced vibration of tall buildings: A state-of-the-art review. J Build Eng. 2021;33:1–60.
2. Kayvani K. Design of high-rise buildings : past , present and future. 23rd Australas Conf Mech Struct Mater. 2014;I:15–20.
3. Baskaran A. Wind engineering studies on tall buildings-transitions in research. Build Environ. 1993;28(1):1–19.
4. Ankireddi S, Yang HTY. Simple ATMD Control Methodology for Tall Buildings Subject to Wind Loads. J Struct Eng. 122(1).
5. Mendis P, Ngo T, Haritos N, Hira A, Samali B, Cheung J. Wind loading on tall buildings. Electron J Struct Eng. 2007;7:41–54.
6. Isyumov N, Case PC. Wind-Induced torsional loads and responses of buildings. Struct Congr 2000 Adv Technol Struct Eng. 2004;103(April 2000).
7. Tse KT, Kwok KCS, Hitchcock PA, Samali B, Huang MF. Vibration control of a wind-excited benchmark tall building with complex lateral-torsional modes of vibration. Adv Struct Eng. 2007;10(3):283–304.
8. Kim HS, Kang JW. Semi-active fuzzy control of a wind-excited tall building using multi-objective genetic algorithm. Eng Struct. 2012;41:242–57.
9. Ni YQ, Chen Y, Ko JM, Cao DQ. Neuro-control of cable vibration using semi-active magneto-rheological dampers. Eng Struct. 2002;24(3):295–307.
10. Jafari M, Sarkar PP, Alipour AA. A numerical simulation method in time domain to study wind-induced excitation of traffic signal structures and its mitigation. J Wind Eng Ind Aerodyn. 2019;193(July):103965.
11. Elias S, Matsagar V. Wind response control of tall buildings with a tuned mass damper. J Build Eng. 2018;15(February 2017):51–60.
12. MEENA RK, Awadhiya GP, Paswan AP, Jayant HK. Effects of Bracing System on Multistoryed Steel Building. 2018;
13. Pal S, Raj R, Anbukumar S. Comparative study of wind induced mutual interference effects on square and fish-plan shape tall buildings. Sādhanā. 2021;0123456789.

© STM Journals 2025. All Rights Reserved 

14. Meena RK, Raj R, Anbukumar S. Effect of wind load on irregular shape tall buildings having different corner configuration. Sadhana - Acad Proc Eng Sci [Internet]. 2022;47(3). Available from: https://doi.org/10.1007/s12046-022-01895-2
15. Meena RK, Raj R, Anbukumar S. Numerical Investigation of Wind Load on Side Ratio of High-Rise Buildings Numerical Investigation of Wind Load on Side Ratio of High-Rise Buildings. Springer Singapore; 2021.
16. Meena RK, Raj R, Anbukumar S. Wind Excited Action around Tall Building Having Different Corner Configurations. Adv Civ Eng. 2022;2022.
17. ASCE: 7-16(2017). Minimum Design Loads and Associated Criteria for Buildings and Other Structures. Structural Engineering Institute of the American Society of Civil Engineering, Reston. AMERICAN SOCIETY OF CIVIL ENGINEERS,Reston. 2017. 1–330 p.
18. ASCE: 49-12(2012). Wind Tunnel Testing for Buildings and Other Structures.Structural Engineering Institute of the American Society of Civil Engineering, Reston. American Society of Civil Engineers 1801 Alexander Bell Drive Reston, Virginia 20191 www.asce.org/pubs; 2012.
19. AS/NZS:1170.2(2011). Structural Design Actions - Part 2: Wind actions. Standards Australia/Standards New Zealand, Sydney. 2011. 101 p.
20. IS: 875 (2015). Indian Standard design loads (other than earthquake) for buildings and structures-code of practice,part 3(wind loads). BIS, New Delhi. 2015. 51 p.
21. ETHIOPIAN STANDARD. ES ISO 4354 (2012) (English): Wind actions on structures. Vol. 2012, Ethiopian Standards Agency( ESA) Ethiopia, Addis Ababa. 2012. 76 p.
22. MNBC. MYANMAR NATIONAL BUILDING CODE 2020. International relation and legal Section Department of Building Minis; 2020.
23. GB 50009-2001. NATIONAL STANDARD OF THE PEOPLE'S REPUBLIC OF CHINA. 2002.
24. Abu Zidan Y, Mendis P, Gunawardena T. Optimising the computational domain size in CFD simulations of tall buildings. Heliyon.
25. Sharma U, Gupta N, Verma M. Prediction of compressive strength of GGBFS and Flyash-based geopolymer composite by linear regression, lasso regression, and ridge regression. Asian J Civ Eng [Internet]. 2023;24(8):3399–3411. Available from: https://doi.org/10.1007/s42107-023-00721-2
26. Sharma U, Gupta N, Verma M. Prediction of Compressive Strength of Geopolymer Concrete using Artificial Neural Network. Asian J Civ Eng [Internet]. 2023;24(8):2837–2850. Available from: https://doi.org/10.1007/s42107-023-00678-2
27. Verma M. Prediction of compressive strength of geopolymer concrete using random forest machine and deep learning. Asian J Civ Eng [Internet]. 2023;24(7):2659–68. Available from: https://doi.org/10.1007/s42107-023-00670-w
28. Verma M. Prediction of compressive strength of geopolymer concrete by using ANN and GPR. Asian J Civ Eng [Internet]. 2023;24(8):2815–2823. Available from: https://doi.org/10.1007/s42107-023-00676-4
29. Verma M, Upreti K, Khan MR, Alam MS, Ghosh S, Singh P. Prediction of Compressive Strength of Geopolymer Concrete by Using Random Forest Algorithm. In: ICACIS 2022. 2023. p. 170–9.
30. Verma M, Upreti K, Dadhich P, Ghosh S, Khatri V, Singh P, et al. Prediction of Compressive Strength of Green Concrete by Artificial Neural Network. In: ICACIS 2022. 2023. p. 622–32.

© STM Journals 2025. All Rights Reserved 

View publication stats

Journal of the Indian Society of Remote Sensing
https://doi.org/10.1007/s12524-024-02073-z

RESEARCH ARTICLE

# Air Quality Analysis in Desert Region in the Northern State of India: GIS Based Approach

**Vignesh Mohan**[1] · **Rajeev Kumar Mishra**[1] · **Vijay Kumar Soni**[2]

Received: 22 April 2024 / Accepted: 19 November 2024
© Indian Society of Remote Sensing 2024

## Abstract

Four year (2018 – 2022) measurements of particulate matter and gaseous pollutants were taken to analyse the air quality in one of the major urbanizing state Rajasthan, India. Bhiwadi, one of the major city in the state, experienced the highest pollution among 8 cities with annual particulate matter concentration of $PM_{2.5}$ ($106.75 \pm 59.25$ μg/m$^3$), and $PM_{10}$ ($231.46 \pm 115.45$ μg/m$^3$) in 2021 with an average increase of 16% when compared to 2020 and ranked one of the most polluted city in Asia. Alwar city, with a concentration of about 44.23 μg/m$^3$ for $PM_{2.5}$, is the lowest received concentration among all other study areas. The particulate matter shows high correlation of about 0.5 – 0.9 in almost all the study regions throughout the study period and a good correlation was observed between CO and $NO_2$ ranges from 0.5 to 0.8 in Jaipur. In the entire study period, the concentration of $PM_{2.5}$ lies between 35 – 135, and 70 – 310 μg/m$^3$ for $PM_{10}$ in all eight different locations. In Rajasthan, out of 8 monitoring locations, three received higher concentrations. In other locations, the concentration was about to exceed the national standards. Growing urban regions are experiencing air quality issues and few Indian cities are ranked globally for poor air quality. It is high time to mitigate the measures for air quality and the study's results determine the spatiotemporal behaviour of air pollution in growing urban regions, which helps initiate the regulatory measures for improving air quality.

**Keywords** Anthropogenic activities · $PM_{2.5}$ · $PM_{10}$ · Rajasthan air quality · Urban air pollution

## Introduction

Rapid urbanization is one of the major causes of depleting air quality in developing and developed cities (Dadhich et al., 2018). World Health Organization (WHO) reports that 37 Indian cities face this issue in the list of 100 global cities. Delhi, Mumbai, Lucknow, Kanpur, and Ghaziabad were considered highly polluted cities in 2014 (Naveen et al., 2021). Air pollution occurs naturally and through anthropogenic activities (Saraswat & Singh, 2021; Sharma et al., 2023). Some recent studies found that the issue of air pollution is not only restricted to one geographical location but also transboundary (Naveen et al., 2021). Apart from various health issues, particulate matter ($PM_{2.5}$, $PM_{10}$) can reduce visibility and increase atmospheric turbidity.

Along with particulate matter, gaseous pollutants such as $SO_2$, CO, $NO_X$, and Ozone play a significant role in determining the concentration of pollutants in the atmosphere by the secondary formation process. For e.g., $SO_2$ interacts with other contaminants to form sulphur trioxide, sulphuric acid. It can stay in the atmosphere for two to four days on average and can also travel to an extent of 1000 km from its origin (Renuka et al., 2020). Ozone was one of the major pollutants that needed to be taken into account because of its vulnerability to crops, plants, and human health. It was estimated that India as a whole will lose 0.3 – 6.7 million tonnes (0.3 – 6.3%) and 4.0 – 14.2 million tonnes (4.2 – 15.0%) of wheat annually (Lal et al., 2017). Apart from particulate matter and gaseous pollutants, meteorological parameters such as relative humidity, wind speed, wind direction, and temperature also play a role in determining the air quality of a locality because they also interact with the primary pollutants and form a complex mixture of sources (Saraswat & Singh, 2021). WHO guidelines suggest that the regions having $PM_{2.5}$ concentration lesser than 10 μg/m$^3$ have a good quality of air for living but in the Indian scenario, the majority of the population lives in the area where $PM_{2.5}$

✉ Rajeev Kumar Mishra
rajeevkumarmishra@dtu.ac.in

[1] Department of Environmental Engineering, Delhi Technological University, Delhi 110042, India

[2] India Meteorological Department, Ministry of Earth Sciences, New Delhi, India

Published online: 29 December 2024

concentration is more than 35 µg/m$^3$ (Naveen et al., 2021). Due to air pollution, the average life span of people living in urban areas is decreasing, and many premature deaths are also occurring (Naveen et al., 2021; Pipal et al., 2014; Yadav et al., 2014a, 2014b). Rajasthan is one of the largest states geographically but in terms of economy, it is a low-income state compared to many other states in India. In Rajasthan, there are 33 districts and Jaipur is the state capital and is considered the largest city in Rajasthan (Dadhich et al., 2018; Yadav et al., 2014a, 2014b). The population of Rajasthan is around 68.6 million and in 8th place by population. The capital city, Jaipur, is listed in the 33rd position as the most polluted city globally, as stated by WHO in 2016 (Bran & Srivastava, 2017; Jain & Mandowara, 2019). Jodhpur, the second largest city in the state, ranked 12th among the most highly polluted cities globally in 2018, as reported by Air Visual and Green Peace (World Air Quality Report, 2018). AQI increased 1.38 times in 2018 compared to 2017 in Jodhpur (Saraswat & Singh, 2021). World Air Quality Report 2021 (World Air Quality Report, 2021) shows that Bhiwadi in Rajasthan ranked 1st in the highly polluted city in central and south Asia, i.e., 106 µg/m$^3$ annual average and it exceeds the annual average concentration of Delhi 96.4 µg/m$^3$. Several studies reported the deteoriated air quality in residenstial and industrial areas with a concentration of about 240 – 660 µg/m$^3$ (Saraswat & Singh, 2021). Yearlong studies also exhibits higher concentration gradually in other cities of Rajasthan (Yadav, et al., 2014a, 2014b). The concentration exceeds nearly 10% of Delhi's annual average concentration. The study focuses on spatial and temporal variations in air pollution concentration in eight major cities in Rajasthan. It investigates the role of local gaseous and meteorological parameters to analyse the monthly changes that occurred in air quality from 2018 to 2022. The outcomes of the study helps to estimate the concentrations of growing urban regions in the country to improve air quality.

## Methodology

### Selection of Study Area

Rajasthan is considered one of the largest state in the country and located in the northwestern part of the country. It has a latitude and longitude ranging between 23° 30' and 30° 11' North, and 69° 29' and 78° 17' East, respectively, with an area of about 3,42,239 km$^2$. Industrialization is occurring at a high rate in Bhiwadi, Udaipur, Kota, Jaipur, and Bhilwara regions. The state's major industries are textiles, handicrafts, cement, ceramic, chemicals, steel, tourism, IT, and marble. Apart from these, Rajasthan is one of the leading mineral-producing states in India. The geographical location of the state also plays a significant role in air quality because apart from anthropogenic sources, natural sources such as dust from deserts and dust storms also contribute to the air quality. During summer, the state temperature may increase to 49 °C, receiving around 31 cm of average yearly rainfall. The selected eight major cities for the study in Rajasthan are presented in Table 1 and Fig. 1. Ajmer city has a population of 542,321 in the city, 551,101 including its suburbs with a total industrial unit of about 19,221 and well-connected transport facilities. Alwar regions have a population of about 3.6 million within a area of 8720 sq km area. The Alwar region have a total 22,465 industrial units and located at a mean sea level (MSL) of 271 m. Similarly Udaipur also have 4201 industrial units is located in the aravali ranges with marble factories located around in the east and northern borders. The Jaipur city with a total 25,935 industrial units covers around 3.23% of the total area of the city. The city having area of around 14,068 sq km, east and north of Jaipur is occupied by hill range rising to over 200 m above the surrounding plain. Pali is located 70 kms away from jodhpur having a population of about 18,19,201 population with an area of 12,387 sq km and 13,834 industrial units. Kota city is having population of about 19,51,014 with an area of about 5211.3 sq km and 12,908 industrial units located in the city. The study regions such as Jodhpur are highly urbanizing with an estimated population of about 1.32 million with a vehicle population of about 1.1 million.

### Data Collection and Analysis

In Rajasthan, there are eight continuous ambient air quality monitoring locations: Ajmer, Alwar, Bhiwadi, Jaipur, Jodhpur, Pali, Kota, and Udaipur and all the 10 different stations in 8 different cities were selected for this study. Hourly average concentration data from the year 2018 to 2022 is used for this study and primary pollutants such as $PM_{10}$, $PM_{2.5}$, $SO_2$, $NO_2$, CO, $O_3$ and meteorological parameters (Wind speed, Wind direction, and Temperature) are taken from the CPCB

**Table 1** Details of selected locations in Rajasthan, India

| City name | Longitude | Latitude | No. of air quality monitoring (AQM) stations |
|---|---|---|---|
| Ajmer | 26.47 | 74.64 | 01 |
| Alwar | 27.55 | 76.61 | 01 |
| Bhiwadi | 28.21 | 76.87 | 01 |
| Jaipur | 26.92 | 75.80 | 03 |
| Jodhpur | 26.29 | 73.04 | 01 |
| Kota | 25.14 | 75.82 | 01 |
| Pali | 25.77 | 73.34 | 01 |
| Udaipur | 24.59 | 73.70 | 01 |

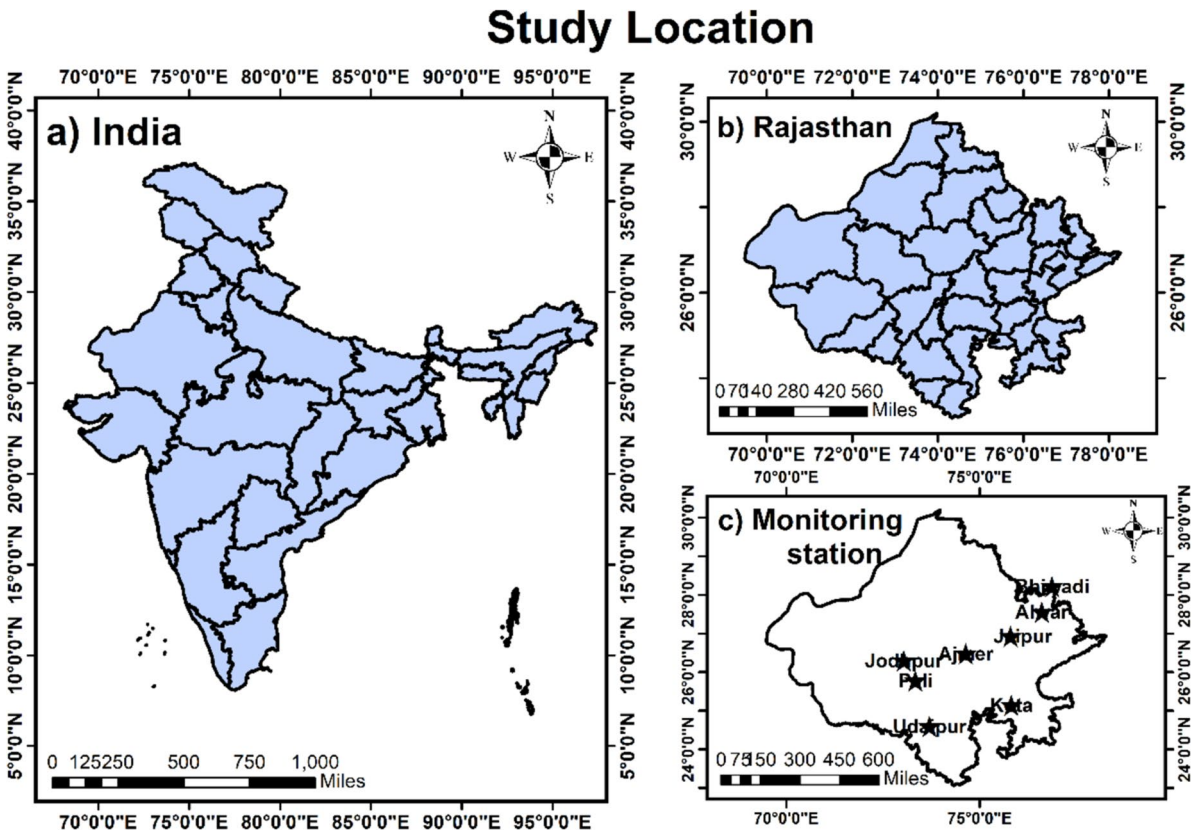



**Fig. 1** Map displaying the monitoring stations

(Central Pollution Control Board) website (https://app.cpcbccr.com/ccr/#/caaqm-dashboard-all/caaqm-landing/data). The collected dataset has been converted into an averaged daily mean concentration from the hourly time stamp. Then, the outliers from the data were removed using an outlier package from R Studio. The data visualization is done using ArcMap version 10.8 and R Studio version 2023.12.1 + 402.

## Result and Discussion

### Temporal Distribution of $PM_{2.5}$ and $PM_{10}$

Bhiwadi is recorded as the most polluted city in 2021 in central and south Asia as per the World Air Quality Report regarding $PM_{2.5}$ concentrations. It was around 180% more (Fig. 2d) than prescribed national ambient air quality standards (Annual NAAQs-40 μg/m$^3$), whereas the concentration of $PM_{10}$ in Bhiwadi was 230.67 μg/m$^3$ (Fig. 2d, Table S3). The Bhiwadi region is known for a mixture of industrial and residential activities. The temporal distribution of $PM_{10}$ concentration shows that the concentration of the pollutants is always higher than the annual standard (60 μg/m$^3$) in almost all the years at all the locations except the year 2020. The five-year average $PM_{2.5}$ concentration of Bhiwadi was around 106.75 μg/m$^3$ (Table S1), the highest recorded concentration compared to other cities. In Alwar, the concentration average is 44.23 μg/m$^3$ (Table S1), the lowest recorded average concentration of 44.23 μg/m$^3$ compared to other cities in the state of Rajasthan. The other major cities such as Pali, Jaipur, Ajmer, Kota and Udaipur are facing the issues of urban agglomeration. Still, the concentration of $PM_{2.5}$ is found in the range of ± 20% to the 24 h standard, i.e., 60 μg/m$^3$ as per the standards. The year-wise analysis of $PM_{2.5}$ concentration shows a reduction in the year 2020 of nearly 10 – 25% of concentration at all 8 locations due to the spread of COVID-19 based restrictions (Table S2) (Agarwal et al., 2020; A. B. Chelani & Gautam, 2022; A. Chelani & Gautam, 2021; Gautam, 2020; Kumar et al., 2023; Mallik et al., 2021; Mohan & Mishra, 2022). The concentration of $PM_{2.5}$ in Bhiwadi increased by 16% in 2021 compared to 2020. The five-year average concentration of $PM_{2.5}$ in Pali, Kota, and Jaipur was in the range of 50 – 60 μg/m$^3$ (Fig. 2) and in Ajmer, Alwar, and Udaipur it was from 40 to 50 μg/m$^3$ (Fig. S1, Table S3). The annual average $PM_{2.5}$ concentration in Jodhpur, increased from 50 to 165 μg/m$^3$ in terms of compared to the national ambient air quality standards from 2018 to 2022. In 2018, the annual average $PM_{10}$ concentration of

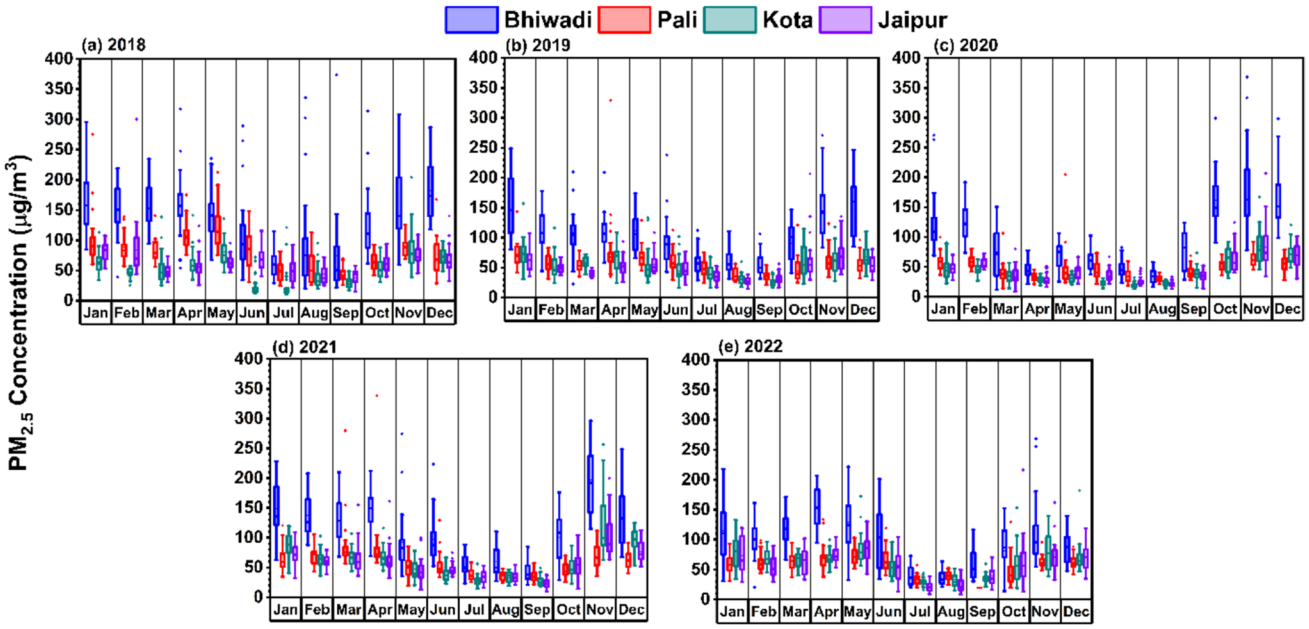



**Fig. 2** Temporal distribution of $PM_{2.5}$ concentrations during (**a**) 2018, (**b**) 2019, (**c**) 2020, (**d**) 2021, and (**e**) 2022, in Bhiwadi, Pali, Kota, and Jaipur

Bhiwadi was 310 µg/m$^3$, the highest recorded $PM_{10}$ and the lowest recorded in 2020 was 188.9 µg/m$^3$ (Table S3). The $PM_{10}$ concentration was higher than the standards in Pali, Jaipur, Kota (Fig. 3), Udaipur, and Jodhpur (Fig. S2). Still, Ajmer and Alwar's concentration was less than the annual average concentration (More than the limits). The increase in concentration was high in regions where urbanization and agglomeration are occurring at a larger rate and relatively stable in areas where urbanization and industrialization are pretty normal. Similar studies in the state estimated the increase in particulate matter and gaseous pollutants in the particular cities of Rajasthan (Kumar et al., 2023; Rai, 2016; Singh et al., 2021).

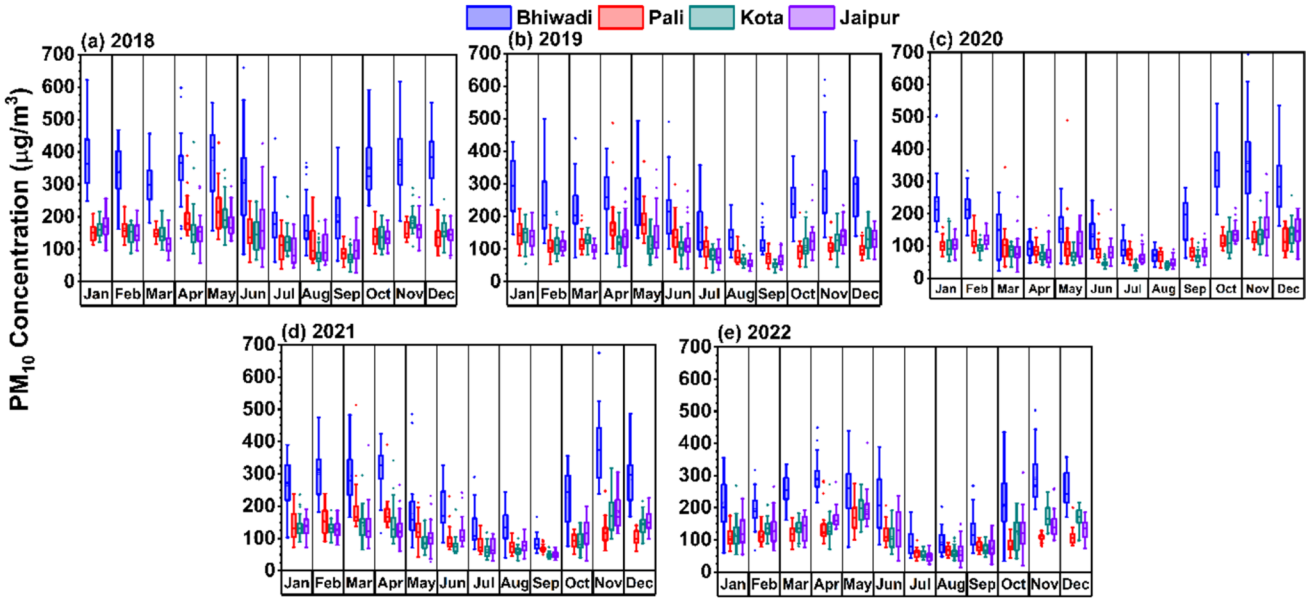



**Fig. 3** Temporal distribution of $PM_{10}$ concentration during (**a**) 2018, (**b**) 2019, (**c**) 2020, (**d**) 2021, and (**e**) 2022, in Bhiwadi, Pali, Kota, and Jaipur

## Correlation Analysis

The correlation analysis can be used to estimate the different roles of atmospheric processes and emissions responsible for air quality which are emitted primarily from different sources (Yadav et al., 2014a, 2014b). The correlation analysis helps to estimate the relation between different parameters and shows how they are related. The correlation coefficient matrices of different locations show a very good positive correlation between the particulate matter ($PM_{10}$ and $PM_{2.5}$).The correlation values range from 0.5 to 0.9 in all the locations from 2018 to 2022 and in the majority of the areas, the correlation value was found to be high, above 0.7. Apart from particulate matter, a good correlation was also observed between CO and $NO_2$ and the correlation values ranges from 0.5 to 0.8 in Jaipur location from 2018 to 2022. The correlation between $O_3$ and particulate matter shows moderate to negative correlation in different regions (Yadav et al., 2023). The regions like Bhiwadi and Kota show negative correlation and other regions such as pali and Jaipur shows less correlation (0.1 – 0.4). In all the locations of the study the wind flow is average throughout the year. The correlation plot exhibits an elliptical shape and values are also displayed high when the correlation is high. When the correlation is reduced, the plot shows a spherical shape, and the display values decrease accordingly. Based on this phenomenon, the correlation matrices were plotted against different parameters, and the values are exhibited in Fig. 4. The correlation analysis for other locations is given in the (Fig. S3a-d). It shows that the locations are dominated by local sources instead of long-range sources from different sectors and regions. In many urban regions, the contributions of local sources are found to be high (Bran & Srivastava, 2017; Pipal et al., 2014).

## Role of Winds on Pollutant Dispersion

The yearlong wind speed and direction measurements show the role of meteorology in dispersion of the particulates from source to the receptor. The wind speed and wind direction in the study region over the years shows that, based on the wind speed and direction the dispersion of particles occur. Higher wind speed paved for more dispersion of the particles and lesser wind speed causes accumulation of the particles

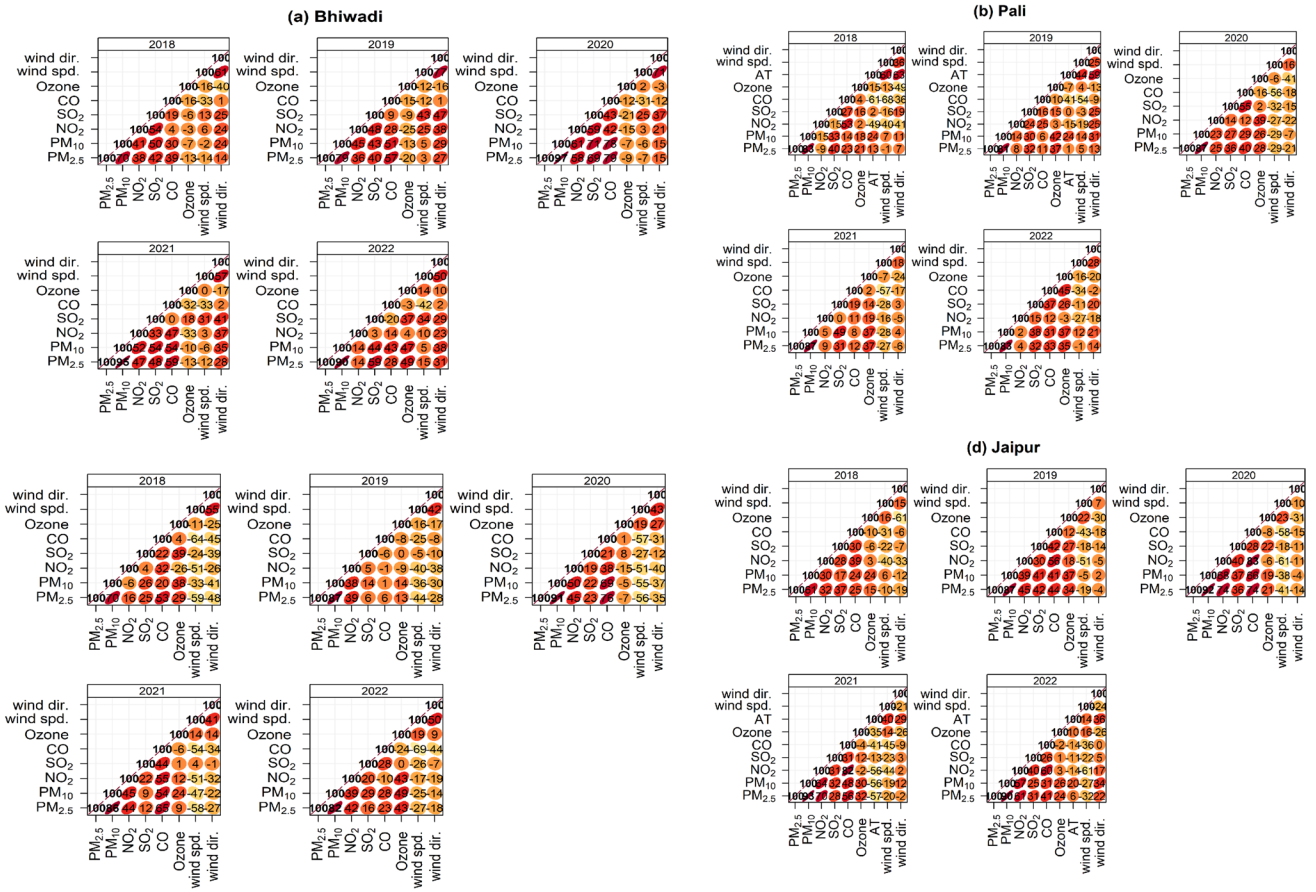



**Fig. 4** Correlation matrices analysis of pollutants concentration in (**a**) Bhiwadi, (**b**) Pali, (**c**) Kota, and (**d**) Jaipur from 2018 to 2022

at the study area. In Bhiwadi, during the study period, wind speed was moderate < 4 m/s in the west-to-east direction sources contributed significantly to $PM_{10}$ and $PM_{2.5}$ concentrations the direction of the wind carries windblown dust (Fig. 5a) from dessert to Bhiwadi and earlier studies in Bhiwadi also finds similar results in 2010, and 2011 in Udaipur City show similar results (Yadav et al., 2014a, 2014b). In kota the wind speed is lesser (i.e., < 3 m/s) than Bhiwadi (i.e., < 4 m/s), even though it is located a bit away from the desert regions, the major wind directions are from the southeast and southwest directions, which carries the desert sand to the study area. The Pali and Ajmer location received the highest average wind speed (3 – 5 m/s) compared to all other locations and the contribution of non-local sources is spread in all directions because it is located in the region where the desert regions covers the study area on one side and the human settlements on another. Windblown dust seems to have a stronger effect than the dilution of pollutants. In the regions adjustment to the desert, such as Pali, Jodhpur, Ajmer and Jaipur, the windblown dust from the desert is one of the prominent sources of the particulate matter and at higher wind speed above 4 m/s shows a good dispersion if the particles in the wind direction. The majority of the location receives wind speeds above 4 m/s. In Alwar location, east to west direction winds carries more concentration of pollutants to the city. Lower wind speed (< 2 m/s) shows a lower concentration at all the locations. The concentration of $PM_{2.5}$ for other locations is given in the Fig. S4a-d.

The concentration of $PM_{10}$ varies based on the increase or decrease of wind speed due to its higher mass, whereas for $PM_{2.5}$, lower wind speed can influence the particle dispersion. When the wind speed is greater than 4 m/s, $PM_{10}$ is observed to increase due to the influence of non-local sources, and also, locations that are located east of the Great Indian Desert receive a contribution from windblown dust. So, the wind blowing from the east, including the southwest and northwest direction, shows more concentration. In Alwar, the wind from the east to the west towards the location shows the concentration of particles at higher concentrations during the year 2019. In 2018, a similar pattern was also seen in the southwest wind direction since this location is near the state boundaries of the neighbouring states (Fig. 6). The concentration of $PM_{10}$ for other locations is given in the Fig. S5a-d.

## Spatial Distribution of Particulate Matter Over the Years

The $PM_{2.5}$ and $PM_{10}$ concentrations lie between 35 – 135 µg/$m^3$ and 70 – 310 µg/$m^3$ respectively. The spatial distribution

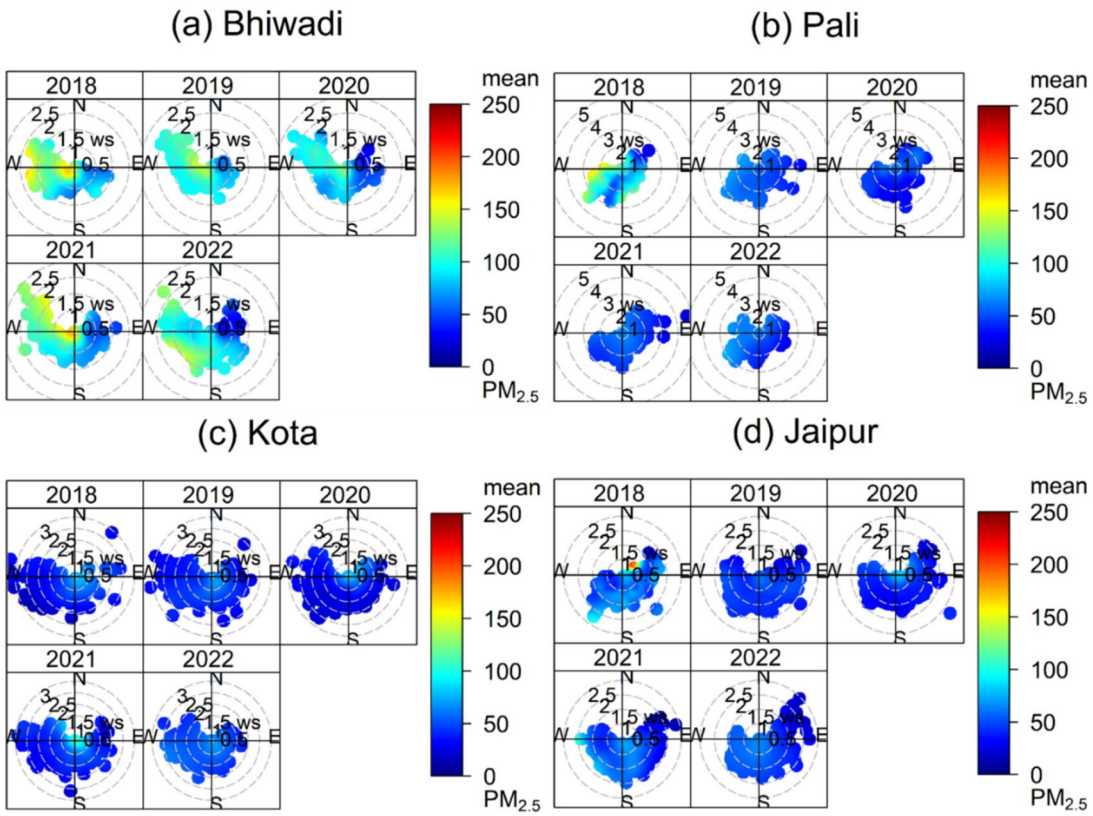



**Fig. 5** Polar cluster analysis of $PM_{2.5}$ concentrations in (**a**) Bhiwadi, (**b**) Pali, (**c**) Kota, and (**d**) Jaipur from 2018 to 2022

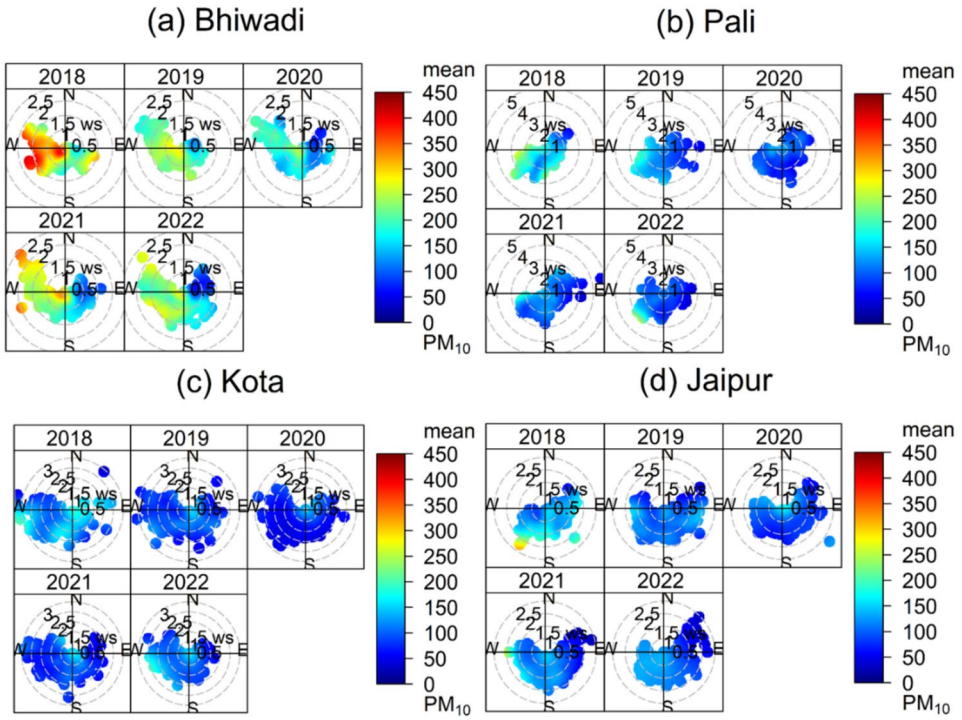



**Fig. 6** Polar cluster analysis of $PM_{10}$ concentration in (**a**) Bhiwadi, (**b**) Pali, (**c**) Kota, and (**d**) Jaipur from 2018 to 2022

of concentration during different years at different locations is analysed using Arc map to distribute the concentration to a particular domain (Figs. 7 and 8). The year 2018 received more concentration of pollutants in almost all the monitoring locations compared to the other years. The reduced concentration in 2020 due to the pandemic is also clearly visible in

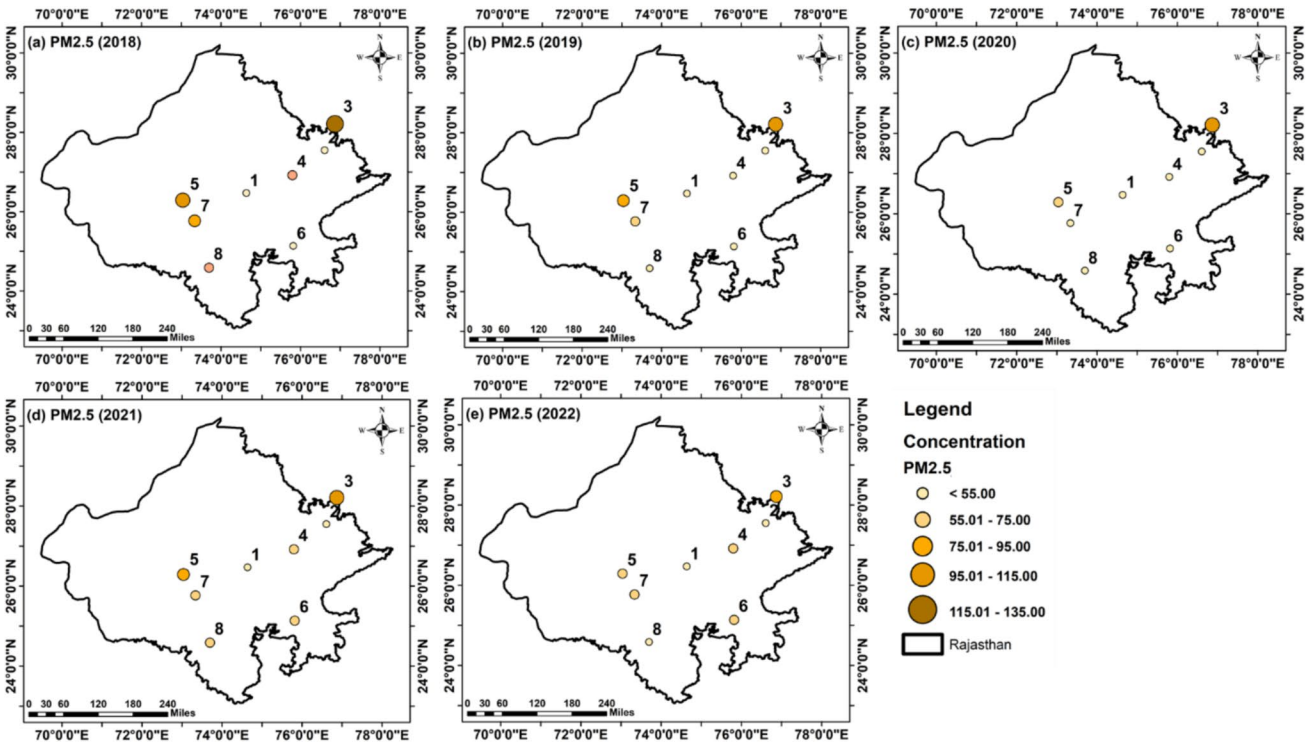



**Fig. 7** Spatial Variation of $PM_{2.5}$ concentration in Ajmer, Alwar, Bhiwadi, Jaipur, Jodhpur, Kota, Pali, and Udaipur from 2018 to 2022

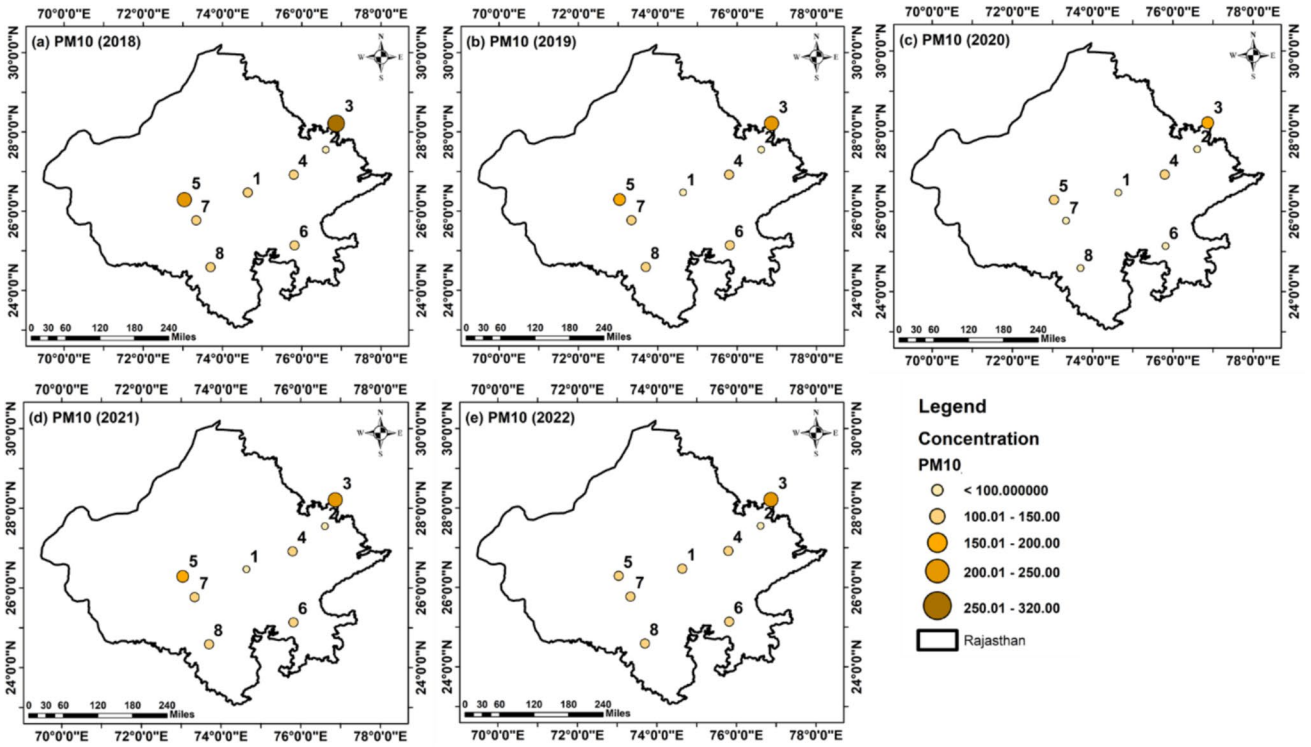



**Fig. 8** Spatial Variation of $PM_{10}$ concentration in Ajmer, Alwar, Bhiwadi, Jaipur, Jodhpur, Kota, Pali, and Udaipur from 2018 to 2022

the study. In the year 2021, the concentration at all the locations increased compared with the national standards like Ajmer (13%), Alwar (7%), Bhiwadi (177%), Jaipur (42%), Jodhpur (88%), Kota (51%), Pali (44%) and Udaipur (38%). The increased anthropogenic activities after the pandemic restrictions caused Bhiwadi to become a highly polluted Asian City. Along with $PM_{2.5}$, the $PM_{10}$ concentration also showed an increase of around 284% in Bhiwadi compared to national standards. Bhiwadi is considered the hub of many industries, and studies show that about 48% of the pollution in Bhiwadi is due to road dust emissions. Among all the other locations in Rajasthan, Ajmer recorded the lowest concentration and the concentration was slightly more (10 – 30%) than the standards. The GIS method is used to analyse the assumptions values throughout the domain in the absence of the data points through mathematical assumptions of the existing data.

## Conclusion

Rajasthan is one of the major state in the western part of the country and a highly growing urban city. Air quality-related studies are very limited to this domain. This study collected and analysed data on air pollutants such as particulate matter and gaseous pollutants along with the meteorology in different cities of Rajasthan from 2018 to 2022. Bhiwadi, a major city in Rajasthan, recorded a higher concentration of particulate matter in South Asia as per the World Air Quality Report 2021.

The annual variations from 2018 to 2022 show a gradual increase in concentration, with the lowest concentration in 2020 due to the COVID-19 pandemic restrictions. The overall average concentration of $PM_{10}$ and $PM_{2.5}$ was high in Bhiwadi, i.e., $231.46 \pm 115.45$ μg/m$^3$ and $106.75 \pm 59.25$ μg/m$^3$ respectively. In contrast, the lowest concentration of $PM_{2.5}$ was recorded in Alwar ($44.23 \pm 21.49$ μg/m$^3$) and Ajmer ($46.70 \pm 18.94$ μg/m$^3$). A very strong correlation was found between $PM_{2.5}$ and $PM_{10}$ at all the locations and a moderate correlation was observed between CO and $NO_2$. Especially from October to December and January to April, the concentration of $PM_{10}$ and $PM_{2.5}$ was extremely high because of the seasonal factor (winter) and reduced planetary boundary layer height. From June to September, the concentration was decreased due to the wet removal of pollutants during monsoon. The state receives a combination of anthropogenic and natural sources (Dust from the desert). Additionally, a few sources from long-range transportation also play a role in determining the concentration.

The role of meteorology, especially wind speed and wind direction, is another major factor in determining pollution concentration in Rajasthan. The locations are covered by desert in the east and the wind flow from the east shows the transport of pollutants (windblown dust) from the desert to

the location under higher wind speed (> 4 m/s). The lower wind speed (< 2 m/s) shows less dispersion. Out of 8 locations, around four locations in the state have increased pollution levels compared to the national ambient air quality standards (NAAQS). In other locations, the concentration is more or less equal to the standards. Appropriate actions in the growing urban regions and policy formulations will help to curb air quality issues at the later stages of urbanization.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s12524-024-02073-z.

**Acknowledgements** The authors are thankful to the Central Pollution Control Board (CPCB) for providing a platform for accessing data from various locations and also to other agencies such as the Rajasthan State Pollution Control Board (RSPCB) for providing data facility and members of the Advance Air and Acoustics Research Laboratory, Department of Environmental Engineering, Delhi Technological University (DTU) for their support to conduct this study.

**Author's Contribution** Vignesh Mohan – Conceptualization, Visualization, formal analysis, reviewing, and editing. Rajeev Kumar Mishra—Investigation, discussion, reviewing, and editing. Vijay Kumar Soni—Discussion, reviewing, and editing.

**Funding** Not applicable.

## Declarations

**Conflict of interest** The authors declared that they have no conflict of interest.

## References

Agarwal, A., Kaushik, A., Kumar, S., & Mishra, R. K. (2020). Comparative study on air quality status in Indian and Chinese cities before and during the COVID-19 lockdown period. *Air Quality, Atmosphere and Health, 13*(10), 1167–1178. https://doi.org/10.1007/s11869-020-00881-z

Bran, S. H., & Srivastava, R. (2017). Investigation of PM2.5 mass concentration over India using a regional climate model. *Environmental Pollution, 224*, 484–493. https://doi.org/10.1016/j.envpol.2017.02.030

Chelani, A., & Gautam, S. (2021). Lockdown during COVID-19 pandemic: A case study from Indian cities shows insignificant effects on persistent property of urban air quality. *Geoscience Frontiers*. https://doi.org/10.1016/j.gsf.2021.101284

Chelani, A. B., & Gautam, S. (2022). The influence of meteorological variables and lockdowns on COVID-19 cases in urban agglomerations of Indian cities. *Stochastic Environmental Research and Risk Assessment*. https://doi.org/10.1007/s00477-021-02160-4

Dadhich, A. P., Goyal, R., & Dadhich, P. N. (2018). Assessment of spatio-temporal variations in air quality of Jaipur city, Rajasthan, India. *Egyptian Journal of Remote Sensing and Space Science, 21*(2), 173–181. https://doi.org/10.1016/j.ejrs.2017.04.002

Gautam, S. (2020). COVID-19: Air pollution remains low as people stay at home. *Air Quality, Atmosphere and Health, 13*(7), 853–857. https://doi.org/10.1007/s11869-020-00842-6

Jain, S., & Mandowara, V. L. (2019). Study on particulate matter pollution in Jaipur City. *In International Journal of Applied Engineering Research., 14*(3), 637–645.

Kumar, A., Sarma, K., Pandey, A., Mishra, R. K., & Devara, P. C. S. (2023). COVID-19 lock-down in delhi: Understanding trends of particulate matter in context of land-use patterns, GIS mapping, and meteorological traits. *Environmental Engineering Science, 40*(1), 1–12. https://doi.org/10.1089/ees.2022.0020

Lal, S., Venkataramani, S., Naja, M., Kuniyal, J. C., Mandal, T. K., Bhuyan, P. K., Kumari, K. M., Tripathi, S. N., Sarkar, U., Das, T., Swamy, Y. V., Gopal, K. R., Gadhavi, H., & Kumar, M. K. S. (2017). Loss of crop yields in India due to surface ozone: An estimation based on a network of observations. *Environmental Science and Pollution Research, 24*(26), 20972–20981. https://doi.org/10.1007/s11356-017-9729-3

Mallik, C., Gadhavi, H., Lal, S., Yadav, R. K., Boopathy, R., & Das, T. (2021). Effect of lockdown on pollutant levels in the delhi megacity: Role of local emission sources and chemical lifetimes. *Frontiers in Environmental Science, 9*(October), 1–20. https://doi.org/10.3389/fenvs.2021.743894

Mohan, V., & Mishra, R. K. (2022). A picture of Delhi's regional air quality during diminished anthropogenic activities in the COVID-19 era. *Arabian Journal of Geosciences, 15*(15), 1–14. https://doi.org/10.1007/s12517-022-10567-8

Naveen, R., Kalyan, S. S. S., Umakanth, N., Madhav, B. T. P., & Rao, M. C. (2021). Influence of air pollutants over India during 2015. *AIP Conference Proceedings. Doi, 10*(1063/5), 0052441.

Pipal, A. S., Jan, R., Satsangi, P. G., Tiwari, S., & Taneja, A. (2014). Study of surface morphology, elemental composition and origin of atmospheric aerosols ($PM_{2.5}$ and $PM_{10}$) over Agra. *India. Aerosol and Air Quality Research, 14*(6), 1685–1700. https://doi.org/10.4209/aaqr.2014.01.0017

Rai, N. K. (2016). Comparative study of ambient air quality using air quality index in residential areas of Jodhpur City. *IOSR Journal of Environmental Science, 10*, 93–98. https://doi.org/10.9790/2402-1012029398

Renuka, K., Gadhavi, H., Jayaraman, A., Rao, S. V. B., & Lal, S. (2020). Study of mixing ratios of SO2 in a tropical rural environment in south India. *Journal of Earth System Science*. https://doi.org/10.1007/s12040-020-1366-4

Saraswat, P., & Singh, S. (2021). Study of particulate matter ($PM_{10}$ and Spm) distribution in ambient air of Jodhpur, Rajasthan. *International Journal of Researches in Biosciences, Agriculture and Technology, 17*, 181–191.

Sharma, M., Khare, M., & Mishra, R. K. (2023). Air quality changes in Delhi due to open waste burning: an accidental fire in Bhalswa landfill. *International Journal of Environmental Science and Technology*. https://doi.org/10.1007/s13762-023-04921-w

Singh, G., Kumar, A., Vaid, D., & Sharma, P. (2021). Stubble burning and its impact on air quality in Delhi NCT: A Case Study. *Journal of University of Shanghai for Science and Technology, 23*(06), 58–68. https://doi.org/10.51201/jusst/21/05223

World Air Quality Report. (2018). *IQAir Air Visual*, IQAir.

World Air Quality Report. (2021). *IQAir Air Visual*, IQAir.

Yadav, R., Beig, G., & Jaaffrey, S. N. A. (2014a). The linkages of anthropogenic emissions and meteorology in the rapid increase of particulate matter at a foothill city in the Arawali range of India. *Atmospheric Environment, 85*, 147–151. https://doi.org/10.1016/j.atmosenv.2013.09.007

Yadav, R., Sahu, L. K., Jaaffrey, S. N. A., & Beig, G. (2014b). Temporal variation of Particulate Matter (PM) and potential sources at an urban site of Udaipur in Western India. *Aerosol and Air Quality Research, 14*(6), 1613–1629. https://doi.org/10.4209/aaqr.2013.10.0310

Yadav, R. K., Gadhavi, H., Arora, A., Mohbey, K. K., Kumar, S., Lal, S., & Mallik, C. (2023). Relation between $PM_{2.5}$ and $O_3$ over different urban environmental regimes in India. *Urban Science*. https://doi.org/10.3390/urbansci7010009

**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.

AIP Conference Proceedings

RESEARCH ARTICLE | JANUARY 16 2025

# Analysing the performance of heptagonal photonic crystal fibre in terahertz regime

Gurmeet Singh; Shubham Sharma; Ajeet Kumar

*AIP Conf. Proc.* 3233, 020010 (2025)
https://doi.org/10.1063/5.0235051

View Online Export Citation

Articles You May Be Interested In

Modeling of the pernicious chemical sensor in the terahertz regime
*AIP Conference Proceedings* (May 2022)

Nonlinear photonic disclination states
*APL Photonics* (January 2023)

Generalized hendecagonal fuzzy number
*AIP Conf. Proc.* (May 2024)

27 January 2025 06:06:02

AIP Publishing

# Analysing the Performance of Heptagonal Photonic Crystal Fibre in Terahertz Regime

Gurmeet Singh[1], Shubham Sharma[1], Ajeet Kumar[1, a)]

[1]Advanced Photonics Simulation Research Laboratory, Department of Applied Physics, Delhi Technological University, Bawana Road, New Delhi, Delhi 110042, India

[a)]Corresponding Author Email: ajeetdph@gmail.com

Abstract: We designed a heptagonal cladding (HC) solid-core (SC) photonic crystal fibre (PCF) to improve the performance in the terahertz (THz) regime. The confinement loss, effective mode area, effective material loss, and nonlinear coefficient have culminated using Chalcogenide glass (ChGs) material as a core because it provides high optical transparency in the infrared spectrum. The cladding region is filled with air holes, providing stability to the PCF, which can operate on the principle of total internal reflection (TIR). The terahertz frequency range from 0.3 to 4 THz has been used to understand the confinement of infrared waves in the fibre. From existing results, confinement loss is reported as $10^{-9}$ dB/cm, which is reported as low as $10^{-15}$ dB/cm in the proposed design while varying air filling fraction (AFF) and THz frequency individually. Using COMSOL Multiphysics software, all the results are computed using computational electromagnetics, whose finite element method (FEM) is performed. With an appropriate choice of parameters, the losses can be minimised.

Index Terms: *Heptagonal cladding photonic crystal fibre, Terahertz, Confinement loss, and Effective material loss.*

## Introduction

Electromagnetic waves ranging from 0.1 to 10.0 THz are referred to as terahertz radiation, sometimes also known as submillimeter-wave radiation [1]. These frequencies are lower than infrared lights but higher than radio waves and microwaves. The terahertz (THz) frequency finds applications in various fields, including imaging [2], astronomy [3], sensing [4], spectroscopy [5], DNA hybridization [6], communications [7, 8], and so on. Terahertz (THz) radiation has also been utilized in medical diagnostics for conditions such as skin cancer, breast cancer, and other critical diseases [2].

Photonic crystal fibre (PCF), a micro-structured fibre, has been introduced in THz wave propagation. These unique fibres have hollow, porous, and solid cores that can confine light and light captivity features that are impossible with conventional optical fibre. Some of these PCFs' special features are low material loss and confinement loss (CL), high mode area, sensitivity, and birefringence. The hollow core fibre that uses photonic band gap characteristics for waveguiding is based on photonic crystals. Significant amounts of the THz field propagate in the air in this fiber. The solid core of a photonic crystal fiber (PCF) operates based on the principle of total internal reflection (TIR). The centre of these fibres transmits the most significant amount of wave energy. Terahertz communication devices use Topas and Zeonex materials as their applications, too [9].'

Many studies have used terahertz (THz) waveguides to describe and briefly analyse the design and performance of PCF structures. A chemical molecule known as a chalcogenide comprises an anion of chalcogen and one additional electropositive element. Materials based on chalcogenides, such as those made of telluride, selenide, and sulphide, are widely available. In various fields, including photovoltaic, photocatalyst, sensor, fuel cell, and battery, binary, ternary, and quaternary chalcogenide materials are helpful. For mid-wave-infrared (MWIR) fibres made of non-oxide components, chalcogenide glasses (ChGs) are an appealing option since they have optical transparency in both the mid- and long-wave infrared ranges [10].

HC PCF results are compared to the existing results from Ref. [11] comprising EML, CL, and effective mode area (EMA), which are $8.0 \times 10^{-2}$ cm$^{-1}$, $3.20 \times 10^{-13}$ cm$^{-1}$, and $1.0 \times 10^{-7}$ m$^2$, Ref. [12] talks about the same values, i.e., EML, CL, and EMA which are $5.0 \times 10^{-2}$ cm$^{-1}$, $7.79 \times 10^{-12}$ cm$^{-1}$, and $2.0 \times 10^{-5}$ m$^2$. Reference [13] is mentioned a slightly different parameter; in place of effective material loss, it talks about the nonlinearity coefficient, confinement loss, and effective mode area, which are $3.0 \times 10^{-8}$ rad/Wm, $10^{-9}$ dB/cm, and $5.62 \times 10^{-6}$ m$^2$.

In this paper, we have simulated the PCF using the finite element method. Four results are evaluated: EML, confinement loss, effective mode area, and nonlinearity coefficient.

*Recent Advancements in Artificial Intelligence and its Applications in Engineering*
AIP Conf. Proc. 3233, 020010-1–020010-6; https://doi.org/10.1063/5.0235051
Published under an exclusive license by AIP Publishing. 978-0-7354-5098-1/$30.00



27 January 2025 06:06:02

## Proposed heptagonal PCF design

The design of the proposed HC SC PCF is illustrated in Fig. 1, depicting the transverse cross-sectional view of the fiber. Five layers of heptagonal shapes have been utilized to construct this fiber, with the 1st, 2nd, 3rd, 4th, and 5th layers having 7, 14, 21, 28, and 35 air holes, respectively. They are introduced in the cladding region with different material.

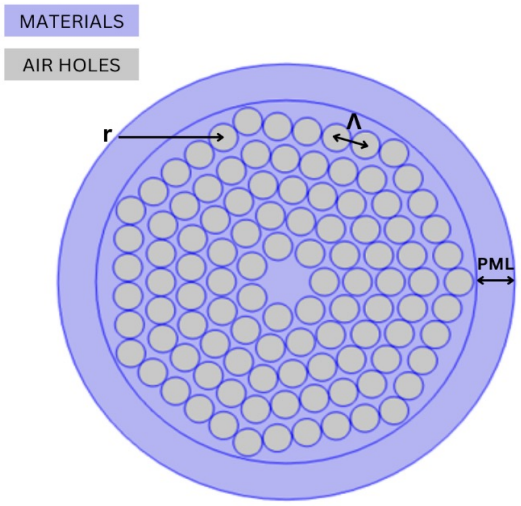



**Fig. 1.** Proposed HC SC PCF

In this design, chalcogenide glass which has various applications in the infrared region has been used. The separation between two air holes is termed as the pitch; using it, the air filling fraction (AFF) remains the same when the frequency varies ranging from 0.3 to 4.0 THz. At 1 THz of frequency, AFF goes between 0.4 to 0.6.

The CL is determined through the perfectly matched layer (PML).

## Formulations

The EML can be calculated as [14]

$$\alpha_{eff} = \sqrt{\frac{\epsilon_0}{u_0}} \frac{\int_{mat} n_{mat} \alpha_{mat} |E|^2 \, dA}{2\left|\int_{all} s_z \, dA\right|} \left(\frac{dB}{cm}\right) \tag{1}$$

Where $\alpha_{mat}$ is bulk material loss, $\vec{s}_z$ is poynting vector in z-direction.

CL which is the loss in a perfectly matched layer (PML), can be evaluated as [14]

$$L\left[\frac{dB}{cm}\right] = \frac{40\,\pi}{\ln(10)\lambda} Im\left(n_{eff}\right) \times 10^{-2}$$

$$= 8.686 k_0 \, Im\left(n_{eff}\right) \times 10^{-2} \tag{2}$$

Where $Im\left(n_{eff}\right)$ is complex portion of the fundamental mode.

A more EMA will make a better PCF which can be calculated using [11],

$$A_{eff} = \frac{\left[\int I(r) dr\right]^2}{\left[\int I^2(r) dr\right]} \tag{3}$$

The nonlinear coefficient, which must be less for the highly effective mode area, can be simulated as

$$\gamma = \left(\frac{2\pi}{\lambda}\right)\left(\frac{n_2}{A_{eff}}\right) \tag{4}$$

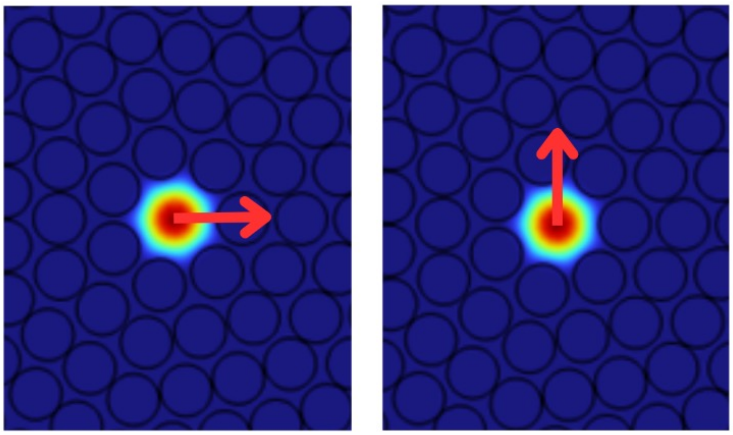
Where $n_2$ is the nonlinear refractive index [12].

27 January 2025 06:06:02

## Numerical results

Using the FEM, we have culminated the EML, EMA, CL, and nonlinear coefficient for the given structure. Chalcogenide glass with a refractive index of 2.65 is used to get desired results.

Initially, frequency is varied from 0.3 to 4 THz, by keeping some parameters fixed, such as radius of air holes, $r = 0.57\,mm,\ \Lambda = 1.9\,mm,\ cladding = 7.5\,mm,\ and\ 20\%$ of clad ding as PML. Then, the diameter/pitch ratio, also known as air filling fraction (AFF), ranging from 0.4 to 0.6, is used to see the effect on the confinement loss. Figure 2 shows the contour plot of the field intensity for the *x-pol and y-pol* at *1 THz* (fundamental mode = *2.60*).



(a) (b)

**Fig. 2.** Fundamental Mode at 1 THz for (a) x-pol., and (b) y-pol.

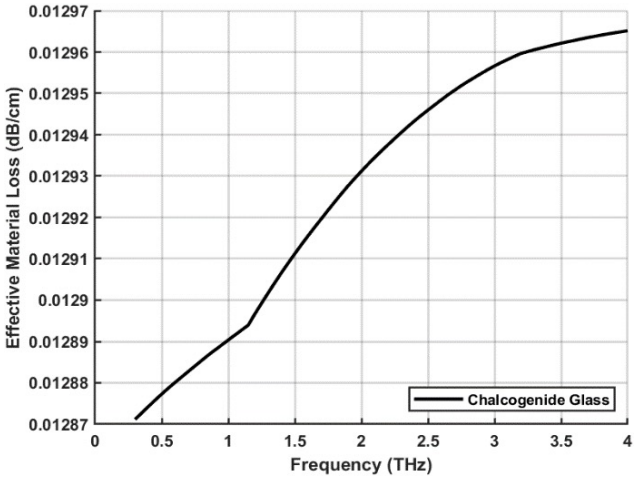



**Fig. 3.** Variation of EML with frequency.

Figure 3 depicts effective material loss (EML) variation against the varying frequency. The EML reported as low as $1.28 \times 10^{-2}$ dB/cm at 0.3 THz and $1.29 \times 10^{-2}$ at 4.0 THz. The EML reaches absorption loss at a higher frequency but remains under the value for the lower one.

The negligible confinement loss is obtained with 20 % extra cladding used as PML. It is reported $3.91 \times 10^{-15}$ dB/cm at 0.3 THz and $4.06 \times 10^{-13}$ dB/cm at 4.0 THz. This is due to better radiation confinement in the infrared region by Chalcogenide Glass and high AFF near the core. Figure 4 shows its variation with frequency. The effective mode

27 January 2025 06:06:02

area is simulated with a very high value of $2.04 \times 10^{-6}$ m$^2$ at the lowest frequency of 0.3 THz. According to the variation shown in Fig. 5, it goes down to $1.85 \times 10^{-6}$ m$^2$ at 4.0 THz because wave energy increases with frequency.

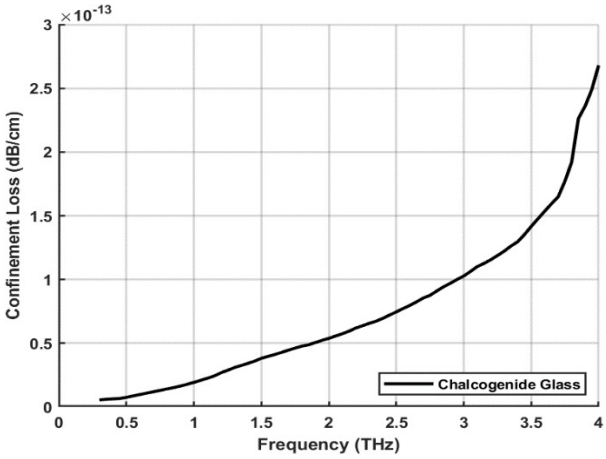



**Fig. 4.** Variation of confinement loss with frequency.

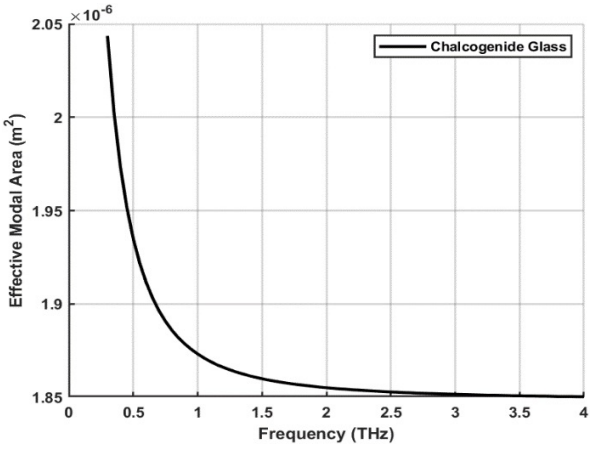



**Fig. 5.** Variation of effective mode area with frequency.

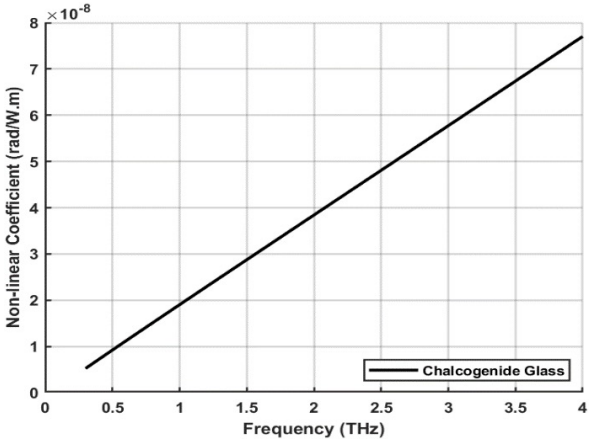



**Fig. 6.** Variation of γ with frequency.

27 January 2025 06:06:02

Nonlinearity losses become low for high effective mode area, which can be seen by comparing Figures 5 and 6. The nonlinearity coefficient is $5.23 \times 10^{-9}$ rad/Wm at 0.3 THz and becomes maximum when the effective mode area is low. This is due to the low energy entering the fibre's core at a lower frequency.

Figure 7 depicts the variation of confinement loss with diameter/pitch ratio, also known as AFF, ranging from 0.4 to 0.6, making the diameter of air holes from $7.6 \times 10^{-4}$ m to $1.14 \times 10^{-3}$ m. The confinement loss is high for lower AFF, i.e., $1.68 \times 10^{-13}$ dB/cm for 0.4 and decreases to $5.36 \times 10^{-15}$ dB/cm for 0.6.

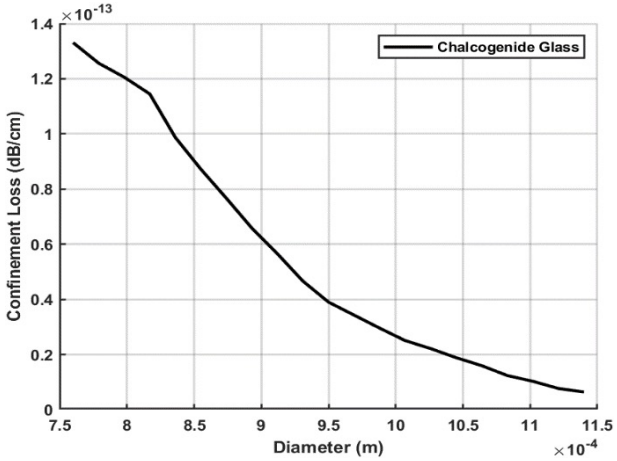



**Fig. 7.** Variation of Confinement Loss with Diameter of air holes.

Table – 1 relates existing results [12 – 14] and simulated results of HC SC PCF in this paper.

**Table 1.** Proposed HC SC PCF with the existing designs.

| Ref. | EML (dB/cm) | L (dB/cm) | $A_{eff}$ ($m^2$) | γ (rad/Wm) |
|---|---|---|---|---|
| [11] | $8.0 \times 10^{-2}$ | $3.20 \times 10^{-13}$ | $1.0 \times 10^{-7}$ | - |
| [12] | $5.0 \times 10^{-2}$ | $7.79 \times 10^{-12}$ | $2.0 \times 10^{-5}$ | - |
| [13] | - | $10^{-9}$ | $5.62 \times 10^{-6}$ | $3.0 \times 10^{-8}$ |
| Proposed design | $1.28 \times 10^{-2}$ | $3.91 \times 10^{-15}$ | $2.04 \times 10^{-6}$ | $5.23 \times 10^{-9}$ |

## Conclusion

In conclusion, the solid core PCF with very low CL and high EMA is proposed for THz wave propagation. The HC SC PCF performance has improved using the chalcogenide glass as the core material. In the broad range of THz from 0.3 to 4.0 THz, the EML of $1.28 \times 10^{-2}$ dB/cm and CL of $3.91 \times 10^{-15}$ dB/cm can be achieved. High EMA can help make this PCF reliable; hence, it can be used for various applications such as tele-communications, spectroscopy, etc.

## References

1. Islam, Md Saiful, Jakeya Sultana, Alex Dinovitser, Mohammad Faisal, Mohammad Rakibul Islam, Brian W-H. Ng, and Derek Abbott. "Zeonex-based asymmetrical terahertz photonic crystal fiber for multichannel communication and polarization maintaining applications." Applied Optics, 57(4), 666-672 (2018).

2. Chen, Q., Jiang, Z., Xu, G. X., & Zhang, X.-C. "Near-field terahertz imaging with a dynamic aperture." Optics Letters, 25(15), 1122-1124 (2000).

3. Ho, Louise, Pepper, M., & Taday, P. "Signatures and fingerprints." Nature Photonics, 2(9), 541-543 (2008).

27 January 2025 06:06:02

4. Strachan, Clare J., Taday, P. F., Newnham, D. A., Gordon, K. C., Zeitler, J. A., Pepper, M., & Rades, T. "Using terahertz pulsed spectroscopy to quantify pharmaceutical polymorphism and crystallinity." Journal of Pharmaceutical Sciences, 94(4), 837-846 (2005).
5. Zhang, Jiangquan, & Grischkowsky, D. "Waveguide terahertz time-domain spectroscopy of nanometer water layers." Optics Letters, 29(14), 1617-1619 (2004).
6. Nagel, M., Haring Bolivar, P., Brucherseifer, M., Kurz, H., Bosserhoff, A., & Büttner, R. "Integrated THz technology for label-free genetic diagnostics." Applied Physics Letters, 80(1), 154-156 (2002).
7. Jamison, S. P., McGowan, R. W., & Grischkowsky, D. "Single-mode waveguide propagation and reshaping of sub-ps terahertz pulses in sapphire fibers." Applied Physics Letters, 76(15), 1987-1989 (2000).
8. Jin, Yun-Sik, Kim, G.-J., & Jeon, S.-G. "Terahertz dielectric properties of polymers." Journal of the Korean Physical Society, 49(2), 513 (2006).
9. Abdullah-Al-Shafi, Md, Akter, N., Sen, S., & Hossain, M. S. "Design and performance analysis of background material of zeonex based high core power fraction and extremely low effective material loss of photonic crystal fiber in the terahertz (THz) wave pulse for many types of communication areas." Optik, 243, 167519 (2021).
10. Yakasai, Izaddeen K., Abas, P. E., Ali, S., & Begum, F. "Modelling and simulation of a porous core photonic crystal fibre for terahertz wave propagation." Optical and Quantum Electronics, 51, 1-16 (2019).
11. Zhang, Yani, Xue, L., Qiao, D., & Guang, Z. "Porous photonic-crystal fiber with near-zero ultra-flattened dispersion and high birefringence for polarization-maintaining terahertz transmission." Optik, 207, 163817 (2020).
12. Sultana, Jakeya, Islam, M. S., Ahmed, K., Dinovitser, A., Ng, B. W.-H., & Abbott, D. "Terahertz detection of alcohol using a photonic crystal fiber sensor." Applied Optics, 57(10), 2426-2433 (2018).
13. Dash, Shikhar, Mathur, V., Pandey, N., & Sinha, R. K. "Terahertz Wave Propagation Characteristics in Graded Teflon Based Solid-Core Photonic Crystal Fibre." Journal of Physics: Conference Series, 2426(1), 012021 (2023).
14. Singh, Gurmeet, Sharma, S., & Kumar, A. "Design and analysis of spider-web photonic crystal fiber in the terahertz regime." Indian Journal of Physics (2023), 1-6. [DOI: 10.1007/s12648-023-03038-7]

27 January 2025 06:06:02

Arabian Journal for Science and Engineering
https://doi.org/10.1007/s13369-024-09950-2

RESEARCH ARTICLE-ELECTRICAL ENGINEERING

# ANFIS-Based Resonant Controller for Mitigating Torque Ripples and Addressing Parametric Variation in PMSM-Driven Electric Vehicle

**Shilpa Ranjan**[1] · **Madhusudan Singh**[1] · **Mini Sreejeth**[1]

Received: 15 July 2024 / Accepted: 29 December 2024
© King Fahd University of Petroleum & Minerals 2025

**Abstract**
The permanent magnet synchronous motor (PMSM) offers very competing innovations for traction drive in electric vehicles (EVs) due to its ameliorated efficiency and power density. Sparse damping of torsional vibration and torque ripple in PMSM can impact passenger comfort. This research introduces an enhanced control method for PMSMs that integrates an Adaptive Neuro-Fuzzy Inference System (ANFIS) with a resonant (RES) control technique to tackle these challenges. The linear PI component has issues with nonlinearities and parameter fluctuations under high-dynamic situations; the fuzzy controller mainly handles these issues. However, the RES component is more appropriate for maintaining control accuracy over particular frequencies, where a single fuzzy might not be as reliable without a lot of tuning and extra rules. ANFIS is a highly appreciated method for dealing with uncertainties. The proposed ANFIS-based RES controller suppresses the undesirable torque ripples and associated noise in EVs and achieves precise reference tracking with disturbance repudiation. The execution of the proposed method for PMSM is analyzed for parametric variation and found to be robust. Simulation performed in MATLAB, additionally experimental validation using dSPACE 1104-based laboratory setup, confirms the robustness and effectiveness of the proposed method.

**Keywords** Permanent magnet synchronous motor · Electric vehicle · Resonant control · Adaptive neuro-fuzzy inference · Torque ripple

## 1 Introduction

The increased awareness to improve air quality, reduce greenhouse gas emissions, and the need for an alternative to fossil fuels has intrigued the thoughts of researchers, vehicle manufacturers, governments, and consumers to look for an alternative to conventional internal combustion engines for modes of transportation. EVs are gaining popularity as a personal and public transportation mode due to inherent advantages like zero-emission, high efficiency, smooth-low-noise operation, etc. [1]. Several governments support EVs as a crucial component of their ideas for mechanisms required to substantially decrease greenhouse gas release and enhance energy capability in the transportation sector. However, EVs have a few disadvantages, such as the cost of battery and range anxiety due to limited range and more charging time. Therefore, it is essential to select the proper drive motor and its control technologies [2, 3]. DC motors have historically been favored in electric vehicle applications because of their straightforward control mechanisms. Induction motors (IMs) are among the suitable options for powering EVs and are broadly used in the latest electric vehicles due to the rapid growth in control technology [4, 5]. Compared to IMs, the PMSM has more power density, better torque-to-current ratios, better efficiency, and durability. Hence, PMSMs are frequently utilized in EV applications these days. However, PMSM drives exhibit torque ripples during low-speed operation, imposed by dead time effects, errors in current measurement, imperfections in the designing of the machine, and other unpredictability. These ripples can further result in speed ripples, which may be tolerable for many industrial purposes but not for EV applications. Since an electric vehicle's engine and gearbox are connected directly to the wheels

✉ Shilpa Ranjan
Shilparanjan_2k21phdee28@dtu.ac.in

Madhusudan Singh
madhusudan@dce.ac.in

Mini Sreejeth
minisreejeth@dce.ac.in

[1] Electrical Engineering Department, Delhi Technological University, New Delhi, India

Published online: 11 January 2025

without a flywheel, torque ripples cause undesirable vibrations and associated discomfort to the passengers. Torque ripples in such applications should be minimal to ensure smooth and efficient operation [6–8]. Thus, there is a scope to ameliorate the performance of PMSM drives by minimizing torque ripples to ensure better driving comfort and efficiency for EV drive applications.

The techniques used for controlling PMSM in EV applications include the proportional–integral (PI) controller, the proportional–integral–derivative (PID) controller, the neural network controller (NNC), the fuzzy logic controller (FLC), the model predictive control (MPC), the sliding mode controller (SMC), the hybrid controllers, and many more [9]. The compound sliding mode control (CSMC) method offers a technique to measure the speed of PMSM. The CSMC consists of a SMC and an extended sliding mode disturbance observer; where the SMC uses a hybrid reaching law, making it very complex to implement [10]. The traditional PI control has an easy design as compared to advanced control systems. However, it is difficult to properly suppress the torque ripples because of their sensitivity to parameter errors and outside disturbances. A frequency-varying resonant controller can be combined with traditional PI controllers to address these issues [11]. The resonant controller's compensatory torque current is merged with the reference current produced by the PI controller, effectively reducing speed fluctuations. The hysteresis current controller (HCC) regulates the current in a voltage source inverter (VSI) by maintaining the actual stator current within a specified hysteresis band around the reference motor current. This approach ensures the generation of a nearly sinusoidal current waveform [12, 13].

Since intelligent controllers can work in highly uncertain circumstances and deal with immense quantities of data, their concept is often used to describe anything that cannot be classified as conventional control. PMSM drives can be controlled using artificial neural networks (ANNs), fuzzy logic controllers (FLCs), and neuro-fuzzy controllers (NFCs). Fuzzy logic improves system nonlinearities by leveraging human experience. However, it needs to enhance its membership functions, as it currently relies heavily on the designer's knowledge and experience [14, 15]. The ANN can handle the system's complex nonlinear features; however, there is a problem with the extended training and convergence times. An Adaptive Neuro-Fuzzy Inference System (ANFIS) combines the features of fuzzy inference and neural networks. An ANFIS-based PI controller provides faster dynamics and improved disturbance rejection, but the computing load increases due to the PI term [16, 17]. Incorporating the RES controller into the speed loop minimizes the torque ripple, and precise reference tracking with disturbance rejection is achieved [18].

The suggested ANFIS-based resonant controller for the PMSM drive consists of two control loops: the outer loop, which uses ANFIS-RES for speed regulation, and the inner loop, which employs HCC to control the output current and ensures the reference signal is tracked. The outer loop of the ANFIS-RES controller generates a compensation torque current, which is combined with the primary reference current in the inner loop to enhance speed response and minimize torque ripples.

The key innovation of the proposed controller and the authors' contributions are summarized as follows:

(i) A novel ANFIS-RES-based speed controller has been proposed for EV drive application to exploit the advantage of addressing nonlinearities, faster dynamics, and improved disturbance rejection with less computational time. RES possesses high controlling and tracking capabilities of the sinusoidal signal.
(ii) The consequence of parametric variation on the performance of speed, torque, and stator current has been analyzed when there is a change in stator resistance of PMSM.
(iii) The PMSM drive has been checked in different scenarios, such as speed change, reversal of speed, and load change.
(iv) Simulation and experimental analysis are performed to validate the ripples caused by uncertainties in motor parameters.

The structure of this paper includes six parts, beginning with the introduction part. Section 2 addresses the modeling of PMSM, Sect. 3 introduces the control approach, and Sects. 4 and 5 discuss the simulation and experimental results. The paper concludes with Sect. 6.

## 2 PMSM Modelling

When modeling a PMSM, the focus is primarily on the stator since the rotor contains permanent magnets. This approach defines the stator currents within the $d-q$ reference frame.

$$\frac{\mathrm{d}}{\mathrm{d}t}i_q = \frac{1}{L_q}u_q - \frac{R_s}{L_q}i_q - \frac{L_d}{L_q}N_p\omega_r i_d - \frac{\lambda N_p\omega_r}{L_q} \tag{1}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}i_d = \frac{1}{L_d}u_d - \frac{R_s}{L_d}i_d + \frac{L_q}{L_d}N_p\omega_r i_q \tag{2}$$

where $i_d$ and $i_q$ represent currents along $d-q$-axis, $L_d$ and $L_q$ denotes respective inductances of $d-q$-axis, $R_s$ refers to stator resistance, $N_P$ signifies pole pairs, $\lambda$ indicates flux linkage, and $\omega_r$ signifies the rotor speed.

The electrical torque is expressed as

$$T_E = 1.5N_P\left[\lambda i_q + \left(L_d - L_q\right)i_d i_q\right] \tag{3}$$

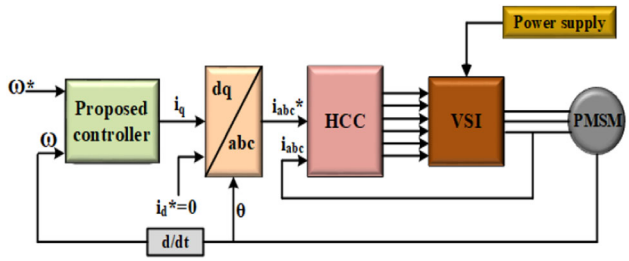



**Fig. 1** Schematic block diagram employing FOC PMSM drive

The motor's speed is obtained as

$$\frac{N_P}{2}(T_E - T_L) = J\frac{\mathrm{d}\omega_r}{\mathrm{d}t} \tag{4}$$

where,$T_E$ indicates electromagnetic torque, $T_L$,the load torque, $J$ is the inertia of the motor.

### 2.1 Field-Oriented Control of PMSM

Figure 1 illustrates the block diagram representing the Field-Oriented Control (FOC) method for a PMSM. The proposed speed controller takes the speed error, the distinction between the reference and actual motor speeds, to get the required reference torque. The reference value for the $q$-axis current is obtained by dividing the torque by the torque constant. To achieve optimal torque vector control at a 90-degree angle, the reference value $i_d$ is set to zero. The $d-q$ voltages of the motor are determined using Park transformation [19]. The hysteresis current controller (HCC) uses the reference 3-phase motor currents and the measured actual 3-phase motor currents to produce the switching signals. This method allows for accurate control of the motor's torque and flux by converting the stator currents into a rotating reference frame, which boosts efficiency and dynamic performance. The integration of HCC further improves the system's ability to precisely track and maintain the desired motor parameters with greater reliability [20, 21].

## 3 Control Strategy

The proposed control strategy incorporates two closed loops: an inner current loop and an outer speed loop. In the traditional control approach, both loops are managed by PI controllers. In contrast, the proposed ANFIS-RES strategy integrates an intelligent ANFIS-RES controller into the outer speed loop, while the inner current loop utilizes an HCC. In this paper, there is a comparative analysis of the PI-RES and ANFIS-RES control methods. For the PI-RES control, the inner loop employs an HCC, and the outer loop uses a PI-RES controller [22].

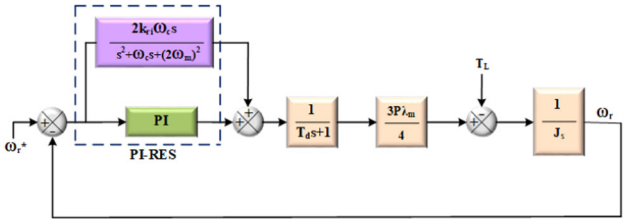



**Fig. 2** Outer speed loop of PMSM drive using PI-RES

## 4 PI-RES Control Strategy

In a PI-RES controller, a proportional resonant controller and a traditional PI controller are combined, where the term resonance relates to the second harmonic rotor frequency. In the proportional resonant controller, the expression in the s-domain for tracking the AC signal is expressed as

$$G_{\mathrm{PR}}(S) = K_p + K_{\mathrm{ri}}\frac{s}{s^2 + \omega o^2} \tag{5}$$

where $K_{\mathrm{ri}}$ and $K_p$ are the resonant coefficient and proportional gain, respectively, and $\omega_o$ is the resonant frequency [23]. Figure 2 illustrates the outer loop incorporating the PI-RES controller. The s-domain open loop transfer function for $G_{\mathrm{PI-RES}}$ is

$$G_{\mathrm{PI-RES}}(S) = K_p + \frac{K_i}{s} + K_{\mathrm{ri}}\frac{s}{s^2 + \omega o^2} \tag{6}$$

where $K_i$ is the integral gain. In the speed control loop, $K_p$ governs the system's dynamic response, whereas $K_i$ and $K_{\mathrm{ri}}$ are responsible for removing any errors at a steady state. A larger gain value can help eradicate the steady state error and lessen the settling time, but it produces a large overshoot. In PI-RES, the controller gains, $K_p$ and $K_i$ are required to be tuned, and the resonant coefficient is calculated using

$$\begin{bmatrix} K_{\mathrm{ps}} \\ K_{\mathrm{is}} \\ K_{\mathrm{ri}} \end{bmatrix} = \begin{bmatrix} \frac{L_q}{4T_d} \\ K_{\mathrm{ps}}\frac{R}{L_q} \\ K_{\mathrm{is}} \end{bmatrix} \tag{7}$$

where, $K_{\mathrm{ps}}$, proportional gain & $K_{\mathrm{is}}$, integral gain of speed loop and $T_d = 1.5T_s$, $T_s$, represents the sampling time.

### 4.1 Proposed ANFIS-RES Control Strategy

In this proposed approach, the PI controller in PI-RES is substituted with an ANFIS controller. Conventional PI controllers frequently provide a satisfactory response for applications requiring minimal power; however, in intense power applications, tuning of PI is not often satisfactory for increased performance of the drive. In order to balance the

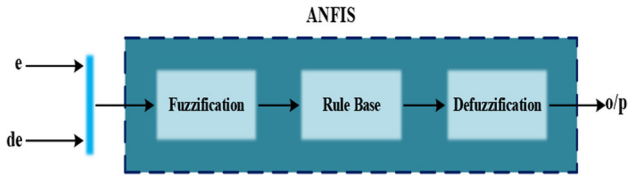



**Fig. 3** Block diagram of ANFIS with two inputs and one output

benefits of fuzzy logic with the particular frequency compensation offered by the RES component, it has been decided to keep the resonant (RES) portion of the PI-RES controller while just changing the linear PI portion. The RES component of the controller is very useful in applications where harmonic rejection is essential, as is often the case with PMSM control, because it effectively attenuates harmonics and disturbances at specific frequencies. Fully replacing the PI-RES controller with a nonlinear fuzzy PI controller introduces design complexity and computational effort challenges, especially for applications with real-time constraints. The fuzzy PI controller may not effectively replicate the frequency selective properties of the RES component, potentially compromising the system's performance in periodic disturbance rejection. Our aim in this work was to propose a practical and incremental enhancement to the standard PI-RES methodology by introducing nonlinearity only where it is most impactful within PI control action. Therefore, retaining the RES part ensures that these benefits are maintained. ANFIS model is more evident to the user and has minimum memorization errors [24].

## 4.2 ANFIS

The Adaptive Neuro-Fuzzy Inference System (ANFIS) is a class of artificial neural networks that relies on the Takagi—Sugeno fuzzy inference system. This advanced and precise model uses input–output data pairs from the analyzed system. ANFIS is particularly useful for modeling and controlling systems that exhibit uncertainty [25–27]. Figure 3 illustrates a simplified block diagram of the ANFIS framework.

The ANFIS algorithm relies on input–output data pairs from the target system for its training process. The ANFIS framework is built on the Ist-order Sugeno model and is defined using two fuzzy IF–THEN rules, as shown in Fig. 4.

- Rule.1: if $x_1$ is $A_1$ and $x_2$ is $B_1$, then $y_1 = p_1x_1 + q_1x_2 + r_1$
- Rule.2: if $x_1$ is $A_2$ and $x_2$ is $B_2$, then $y_2 = p_2x_1 + q_2x_2 + r_2$

To execute the two rules, the ANFIS structure is organized into five layers: Input, Fuzzification, Rule, Normalization and Output, as illustrated in Fig. 5. In this setup, $A_i$ & $B_i$ $(i = 1,2)$ represent the fuzzy sets, while $p_i$, $q_i$,$r_i$ $(i = 1,2)$ correspond

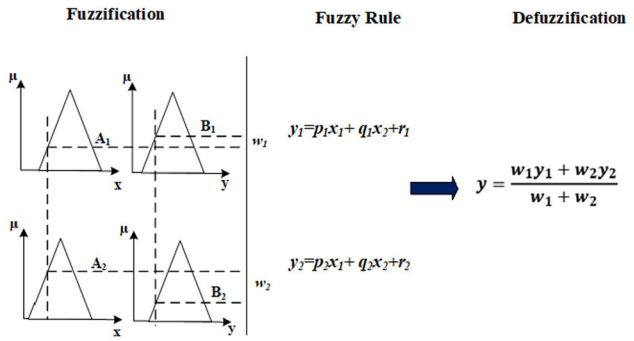



**Fig. 4** Takagi–Sugeno-based fuzzy inference system with triangular-shaped membership function (For inputs $x_1$ and $x_2$, $A_1$,$A_2$ & $B_1$,$B_2$ are the fuzzy sets whose outputs are denoted by $y_1$ & $y_2$ in which $p_1$,$q_1$,$r_1$ & $p_2$,$q_2$,$r_2$ are the constant values decided by the least square method(LSM), and w is the membership degree of the fuzzy rule)

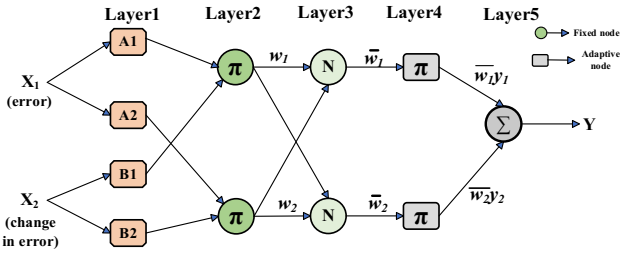



**Fig. 5** Basic structure of ANFIS

to the design parameters, which are adjusted throughout the training process.

1. Input Layer: PMSMs often operate under dynamic conditions influenced by speed, torque, and current parameters. This layer serves as a sensory processor, where crisp inputs (e.g., error in speed and rate of change of error in speed) are transformed into fuzzy linguistic terms (e.g., "low," "medium," "high") using membership functions. At every $i$-th position, the node is adaptive, and its functionality is represented by the node function as

$$O_{1,i} = \begin{cases} \mu_{Ai}(x); & \text{for } i = 1, 2 \\ \mu_{Bi-2}(y); & \text{for } i = 3, 4 \end{cases} \tag{8}$$

2. Fuzzification Layer: This layer evaluates the relevance of each rule by calculating the "firing strength" based on the degree of membership of input variables. For example, if a speed error is "high" and its rate of change is "increasing," a particular rule may dominate the decision-making process. Each node, indexed by *"i"*, represents the strength of the rule's activation by multiplying the input signals and generating the corresponding output as

$$O_{2,i} = w_i = \mu_{A_i}(x)\mu_{B_i}(y);\ i = 1, 2 \tag{9}$$

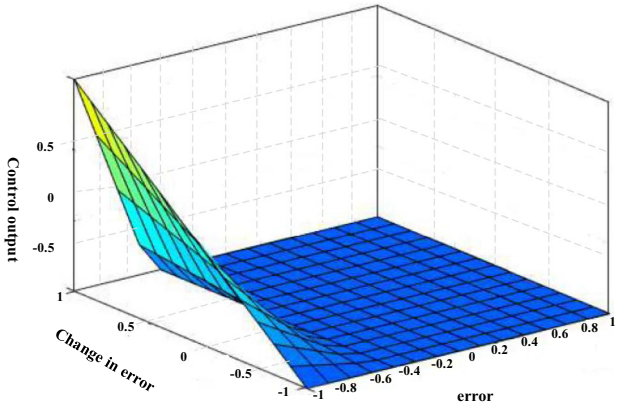



**Fig. 6** Surface view of error vs change in error for PMSM control

3. Rule Layer: In this layer, the weighted outputs from different rules are normalized to ensure balanced contributions. This harmonization resembles a team of experts aligning on the most reliable control strategy. The $i$-th node calculates normalized firing strength as

$$O_{3,i} = \overline{w_i} = \frac{w_i}{w_1 + w_2}; \quad i = 1, 2 \tag{10}$$

4. Normalization Layer: The normalized outputs are used to compute crisp control actions. This layer combines fuzzy reasoning with adaptive learning to generate precise decisions. For PMSM, this could mean determining optimal voltage and current commands to maintain efficient operation. The node function in this layer is calculated by following the expression

$$O_{4,i} = \overline{w_i} y_i = \overline{w_i}(p_i x + q_i y + r_i) \tag{11}$$

5. Output Layer: Using neural network learning techniques, it adjusts the parameters of membership functions and rule weights. This adaptability ensures that the system continually refines itself, accommodating changes in PMSM dynamics or operating conditions. The final output is calculated as

$$0_{5,i} = \sum_i \overline{w_i} yi = \frac{\sum_i w_i y_i}{\sum_i w_i} \tag{12}$$

In Fig. 6, the graphical illustration of the surface of the fuzzy controller for its operating space is shown. This figure demonstrates that the $X$-axis (Change in Error) and the $Y$-axis (Error) represent the input variables to the fuzzy inference system. The $Z$-axis (Output) represents the system's output based on the inputs.

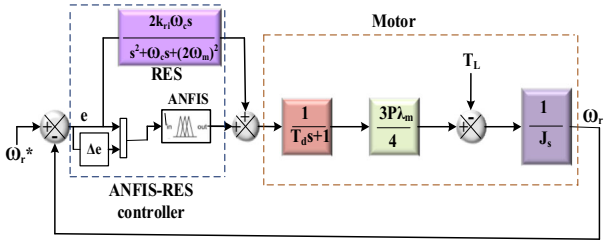



**Fig. 7** Outer speed loop of PMSM drive employing ANFIS-RES controller

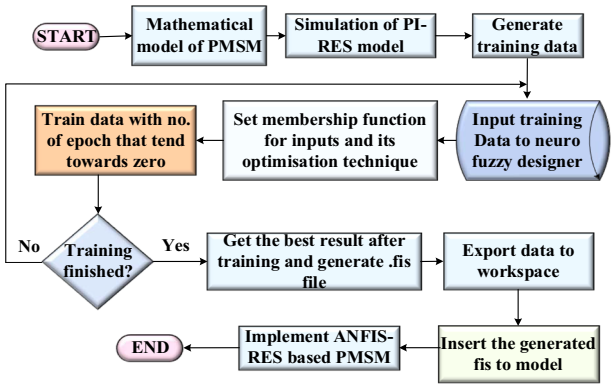



**Fig. 8** Flowchart of proposed ANFIS-RES control Strategy

### 4.3 ANFIS-RES Algorithm

ANFIS-RES has been developed by integrating the fundamental ANFIS architecture with the RES controller. It provides the benefit of significant gain at its resonant frequency, which shortens steady-state error and overshoots or undershoots while tracking or rejecting a sinusoidal input. Figure 7 illustrates the block diagram of the proposed controller within the speed loop. ANFIS-RES utilizes the reference and actual speeds to determine the $q$-axis current in the outer speed loop. The reference 3-phase current ($i^*_{abc}$) obtained through the $dq - abc$ transformation along with the measured motor current ($i_{abc}$) is sent to the HCC to produce the firing signals for the VSI to drive the PMSM.

The proposed ANFIS-RES control strategy is depicted in the flowchart shown in Fig. 8. Speed data are considered as input training data, and each input is defined employing a membership function. The process involves:

*Step 1* The dataset used for training is created using the reference model employing the PI-RES controller, which is normalized for the proposed controller and loaded in the neuro-fuzzy designer. In this study, the input–output variables are scaled to the [0,1] range for ANFIS training. The dataset is then split into two subsets: one for training and the other for testing. The training subset is utilized to optimize the parameters of the membership functions to achieve the best fit. In contrast, the testing subset helps prevent overfitting by

**Table 1** PMSM parameters

| Parameters | Ratings |
|---|---|
| Rated speed, $\omega_r$ | 3000 rpm |
| Rated torque, $T_e$ | 11 Nm |
| Rated voltage, V | 380 V |
| Rated current, I | 6.9 A |
| No. of poles pairs, $N_P$ | 3 |
| Flux linkage | 0.265 wb |
| Inertia | 0.0011 kg $m^2$ |
| Stator resistance, $R_s$ | 1.93 Ω |
| Inductance | 11.4 mH |

evaluating the model's generalization ability. The input variables include speed error ($e$) and the change in speed error ($\Delta e$), with current being the output. 80% of the data is used for training the ANFIS model, while the remaining 20% is reserved for testing.

*Step 2* The grid partition technique is considered to develop the fuzzy inference system (FIS) with the desired number and type of membership functions. The proposed controller uses five triangular membership functions.

*Step 3* The hybrid optimization technique trains the generated FIS with an error tolerance of 0.0116 subject to 50 iterations and backpropagation for better system accuracy.

*Step 4* The above steps are repeated until the root mean square error (RMSE) value is acceptable.

*Step 5* Upon completing the training process, the output from the final layer is processed to determine the production of the ANFIS-RES controller, which involves calculating the resonant coefficient, $K_{\mathrm{ri}}$. The drive system then uses the combined output from the ANFIS-RES controller to generate the necessary pulses for the inverter.

## 5 Simulation Studies and Discussion

A 3-phase, 3.6 kW surface-mounted permanent magnet machine (SPMSM) drive with specifications provided in Table 1 is used for simulation studies using MATLAB/Simulink and experimental validation. The sampling time taken for simulation is 10 µs. The performance of the PMSM drive is evaluated under varying operating scenarios, and the PI-RES method is compared with the ANFIS-RES approach. In the simulation results, the actual torque and speed of the motor are shown in blue and green, respectively, while the reference torque and speed are shown in red.

### 5.1 PMSM Drives Performance During the Starting

Figure 9a and b shows the starting characteristics of SPMSM

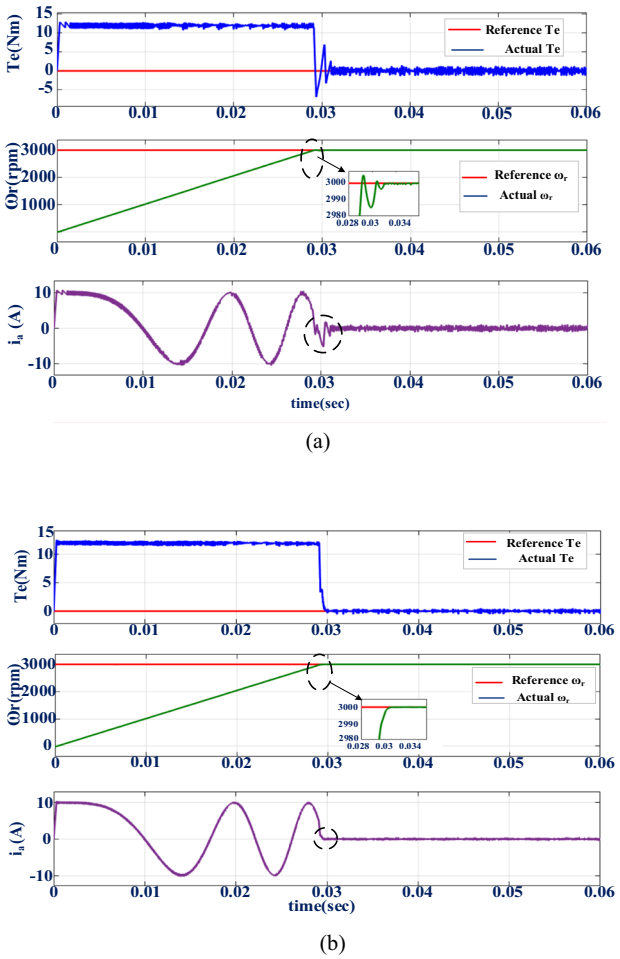



**Fig. 9** Starting characteristics of PMSM drive at rated speed, no-load using **a** PI-RES **b** ANFIS-RES

**Table 2** Comparative analysis of starting performance

| Control technique | Rise time, *tr* (ms) | Settling time, *ts* (ms) | %Overshoot in speed |
|---|---|---|---|
| PI-RES | 23.175 | 13.968 | 1.003 |
| ANFIS-RES | 23.025 | 13.811 | 0 |

when the motor is started at rated speed with no load. When a PI-RES controller is utilized as a speed controller, some overshoot and undershoot in speed and distortion in stator current are seen; however, an ANFIS-RES controller regulates motor speed without any overshoot and provides a quick and smooth response. With PI-RES, oscillations in torque have been observed, while ANFIS-RES does not show such oscillations. At the beginning of operation, the stator current increases as anticipated to generate the necessary dynamic torque to counteract the motor's inertia. Table 2 shows a comparative analysis of the starting performance of the PMSM,

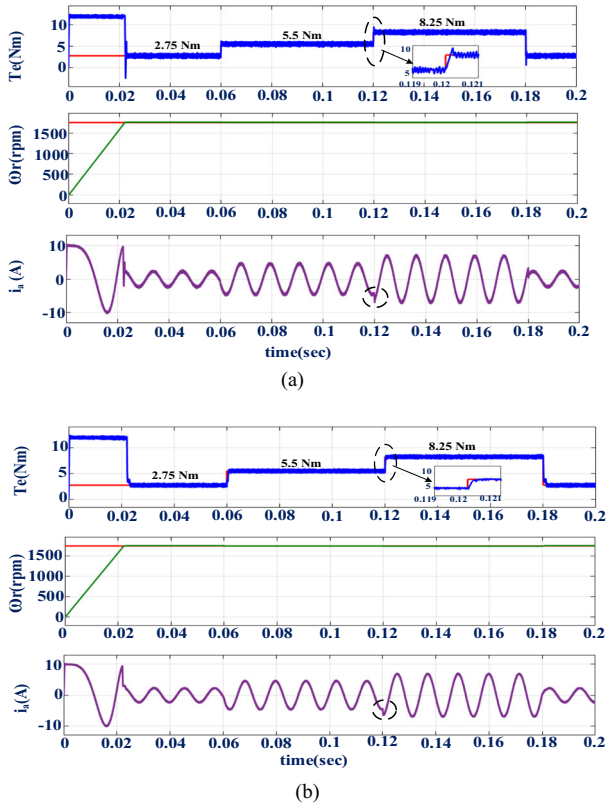



**Fig. 10** Dynamic characteristics of PMSM drive during a step change in load at rated speed using **a** PI-RES **b** ANFIS-RES

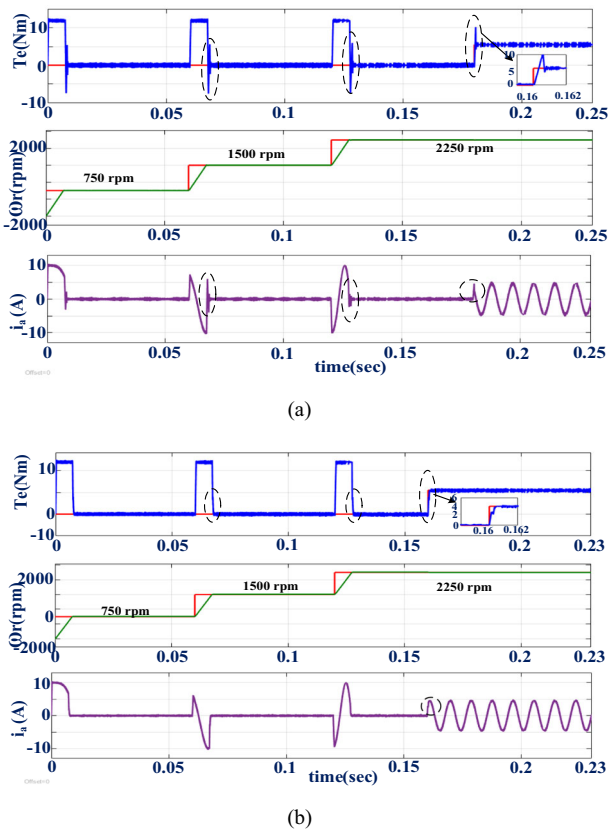



**Fig. 11** Dynamic response of PMSM drive during a step change in speed and load using **a** PI-RES **b** ANFIS-RES

**Table 3** Comparison of torque ripple (TR)

| Scenarios | TR using PI-RES | TR using ANFIS-RES |
|---|---|---|
| 25% of rated load | 0.483 | 0.316 |
| 50% of rated load | 0.252 | 0.203 |
| 75% of rated load | 0.129 | 0.1 |

considering the aspects of settling time, rise time, and overshooting for both control techniques.

## 5.2 PMSM Drive Performance for a Step Change in Load at 1750 rpm

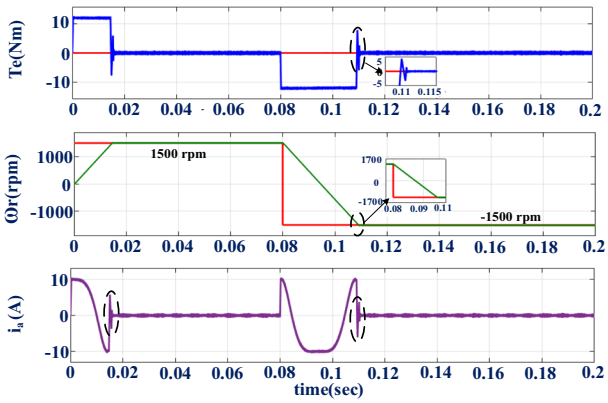
Figure 10a and b shows how the drive performs during a step change in load. At startup, the motor operates at no-load condition and achieves a speed of 1750 rpm. At $t = 0.02$ s, 25% of the rated load torque is exerted on the rotor, and this torque is increased to 50% at 0.06 s and 75% at 0.12 s. During this load transition, some overshoots are observed with PI-RES. However, no such overshoot is observed with the proposed method. Table 3 shows the comparative analysis of torque ripple at different load conditions. Eventually, the motor's speed is recuperated, and it tracks the reference speed.

## 5.3 PMSM Drive Performance for a Step Change in Speed with Sudden Application of Load

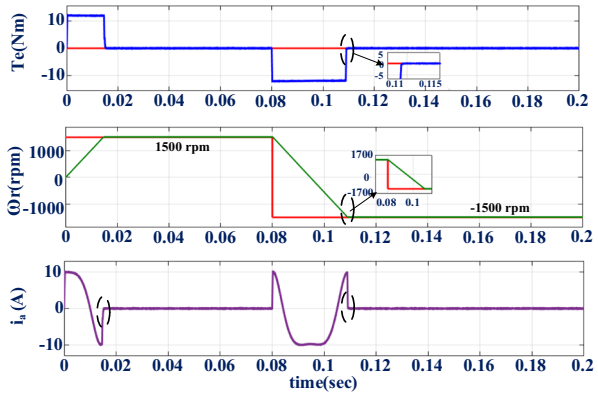
The transient behavior of the PMSM drive is analyzed during step changes in rotor speed from 750 to 1500 rpm and 1500 to 2250 rpm. The motor proceeded with no load at 25% of the rated speed. At 0.06 s, the motor's speed increased to 50% of the rated speed; at 0.16 s, the speed surged from 50 to 75% of the rated speed. At 0.16 s, the motor exhibited half of the rated load torque while maintaining a smooth trajectory of its reference speed. Due to the sudden load application, a drop of 14 rpm is observed in PI-RES, whereas no such drop is observed in ANFIS-RES. Figure 11a and b illustrates a step change in drive speeds under no-load conditions and a step change in load at 0.16 s.

## 5.4 PMSM Drive Performance During Speed Reversal at No-Load Condition

At no load, the motor initiates at 1500 rpm. Figure 12a and b illustrates the dynamic behavior of the PMSM drive during speed reversal from 1500 to − 1500 rpm when no load is applied. At 0.08 s, the motor speed is reversed. In the case of




(a)




(b)

**Fig. 12** Dynamic characteristics of PMSM drive during speed reversal at no load using **a** PI-RES **b** ANFIS-RES

PI-RES, some oscillations are noticed in torque during the reversal of speed, and no such oscillations are observed with the proposed method.

## 6 Experimental Results

The laboratory prototype developed for experimental validation of simulation studies is shown in Fig. 13.

A 3.6 kW, 380 V, 6Pole SPMSM motor with a permanent magnet synchronous generator (PMSG) functioning as the load has been chosen for experimental validation. The PMSG is attached to the shaft of PMSM for electric loading with resistor banks. A dual-level insulated gate bipolar transistor (IGBT) system operating in a three-phase configuration-based inverter drives the SPMSM. The interface between the host PC and the experimental setup is a DSP-based dSPACE1104 controller. A rotatory encoder is fitted on the motor shaft that measures the actual motor speed to confirm the accuracy of the estimated speed. The rotor speed, electromagnetic torque, and stator current responses are examined under various running conditions. The experimental results

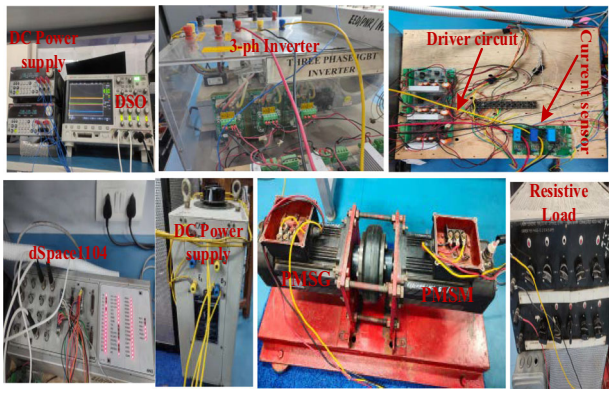
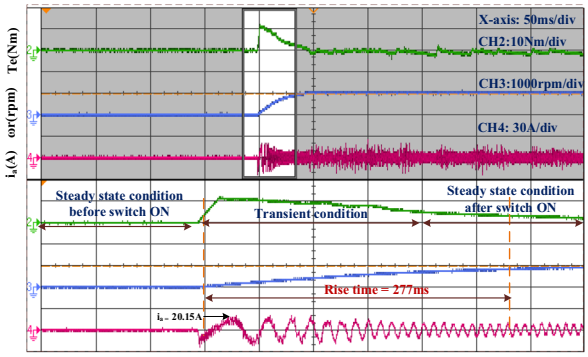
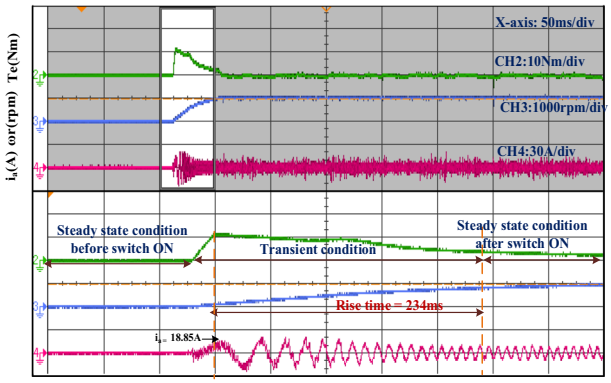



**Fig. 13** Laboratory prototype of PMSM drive




(a)




(b)

**Fig. 14** Dynamic response of laboratory prototype during starting under no load at rated speed using **a** PI-RES **b** ANFIS-RES

obtained from the same simulation setup are presented. These results provide insights into the consistency and reliability of the simulation process.

### 6.1 PMSM Drives Performance During the Starting

Figure 14a and b shows the reaction of the PMSM drive to dynamic changes during starting under no-load condition. The proposed controller exhibits a rise time that is 16%

**Table 4** Comparative analysis of starting performance of laboratory prototype

| Control technique | Rise time, *tr* (ms) | Settling time, *ts* (ms) |
|---|---|---|
| PI-RES | 277 | 463 |
| ANFIS-RES | 234 | 407 |

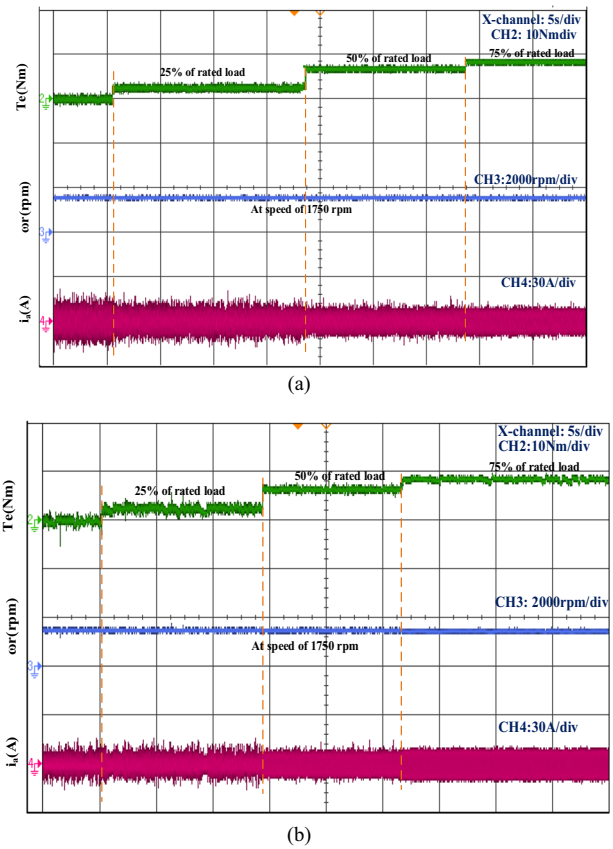



**Fig. 15** Dynamic performance of laboratory prototype under load variation using **a** PI-RES **b** ANFIS-RES

shorter than that of the PI-RES controller. When a motor starts, it demands an immense amount of current to overcome its inertia, but this current slowly drops until it settles to zero. Both the controllers possess comparable no-load dynamics, with the ANFIS-RES controller reaching a steady state a little faster which is given in Table 4.

## 6.2 PMSM Drive Performance for a Step Change in Load at 1750 rpm

Figure 15a and b depicts the dynamic behaviors of the PMSM drive during loading variation. The motor has been started at no load, and different loads are enforced at regular intervals to confirm the efficacy of the proposed method. The ripples induced in torque are given in Table 5, which shows

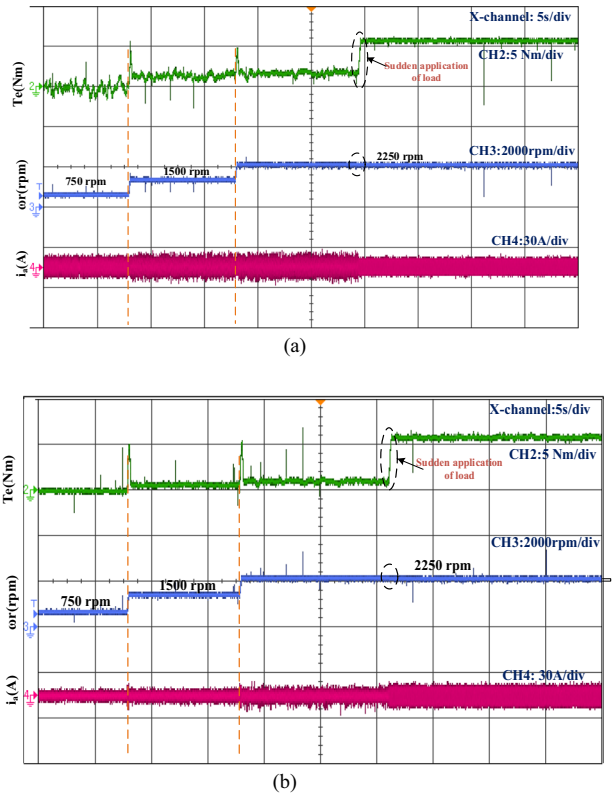



**Fig. 16** Dynamic response of laboratory prototype during step change in speed and load using **a** PI-RES **b** ANFIS-RES

that ANFIS-RES possesses appreciably fewer ripples in each scenario compared to PI-RES.

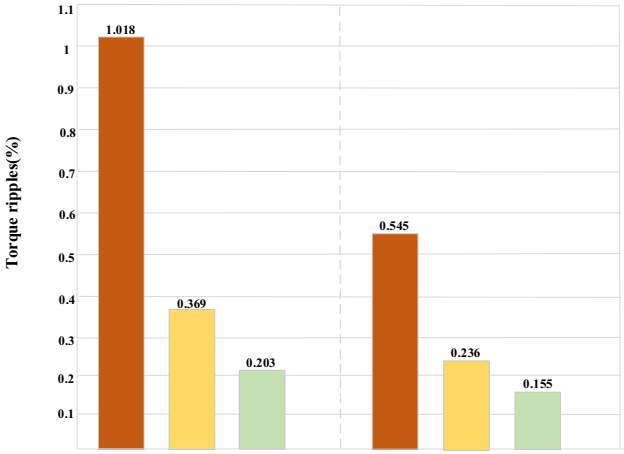
## 6.3 PMSM Drive Performance for a Step Change in Speed and Sudden Application of Load

Figure 16a and b illustrates the dynamic behavior of PMSM drive under varying speeds (25%, 50%, and 75% of rated speed). The motor is powered on at no load at a 25% rated speed. The speed is subsequently increased to 50% and then 75% of the rated speed. After the motor achieves steady-state speed, a step increase in load (50% of rated load) is applied. It is observed that at 25% and 50% of the rated speed, the ripples induced in torque are greater, and the current drawn is also greater while using the PI-RES controller as compared to ANFIS-RES. At 75% of the rated speed, the response for both controllers is almost the same. ANFIS-RES responds better even at low speeds, which is favorable for EV operations.

## 6.4 PMSM Drive Performance During Speed Reversal at No-Load Condition

Figure 17a and b illustrates the dynamic behavior of the

**Table 5** Comparison of torque ripple (TR) of laboratory prototype




| At 1750 rpm | PI-RES | ANFIS-RES |
|---|---|---|
| 25% of rated load | 1.018 | 0.545 |
| 50% of rated load | 0.369 | 0.236 |
| 75% of rated load | 0.203 | 0.155 |

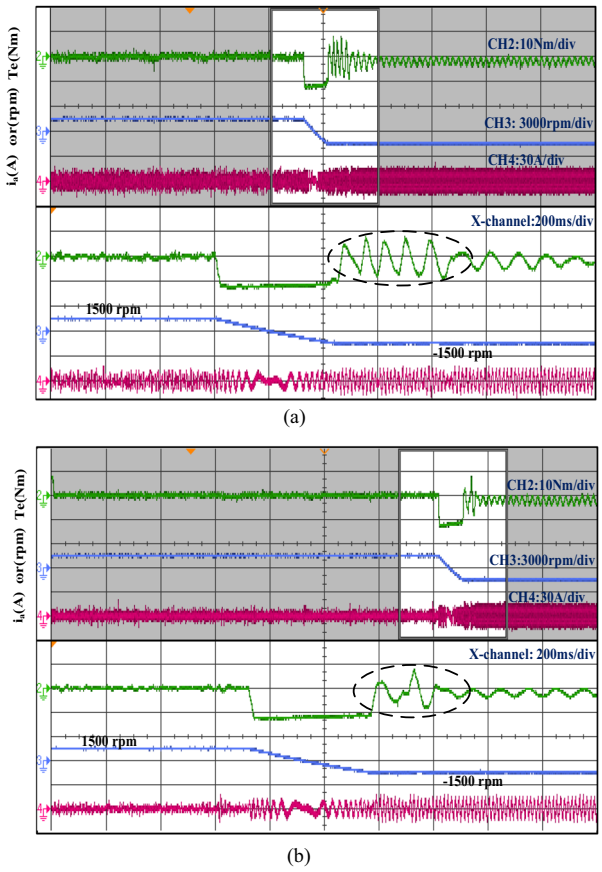



**Fig. 17** Dynamic response of laboratory prototype during speed reversal using **a** PI-RES **b** ANFIS-RES

PMSM drive during a speed reversal. The results demonstrate that ANFIS-RES produces significantly lower torque ripples compared to PI-RES.

From all the characteristics response of the PMSM drive, it is perceived that the proposed ANFIS-RES controller provides faster responses with fewer ripples in torque and current as compared to the PI-RES controller.

## 6.5 Effect of Parameter Variation on Speed and Torque

For stable motor drive operation, the controller should be robust enough to cater to variations in parameters of the motor, such as permanent magnet flux, d and q-axes inductances, and armature resistance, which fluctuate during motor operation [28, 29]. As the temperature changes, a change in stator resistance value occurs. The resistance variation caused by machine losses can be represented as

$$R_s = R_{s0}[1 + \alpha(T - T_0)] \tag{13}$$

where, $R_s$ is the actual value of resistance, $R_{s0}$ is the initial resistance value, $\alpha$ is the temperature coefficient, $T$ and $T_0$ are the actual and initial temperature of the stator winding.

To analyze the effect of an increase in temperature, the motor is continuously run for a couple of hours, and the change in resistance and performance of the laboratory prototype is observed. The stator resistance increased from 3.3 to 4Ω after the said continuous operation of the motor. The key parameters for evaluating the performance of the motor under increased stator resistance are tabulated in Table 6.

**Table 6** Effect of parametric variation on the speed of PMSM drive

| Scenario | PI-RES controller | | ANFIS-RES controller | |
|---|---|---|---|---|
| | Rise time, *tr* (ms) | Settling time, *ts* (ms) | Rise time, *tr* (ms) | Settling time, *ts* (ms) |
| Without change in *Rs* | 277 | 463 | 234 | 407 |
| With change in *Rs* | 279 | 471 | 234 | 407 |

It is observed that for up to a 20% increase in stator resistance, the designed controller effectively follows the reference speed without any deterioration in performance. A slight change in the rise time of speed is seen while using the PI-RES. The robustness of the proposed ANFIS-RES control strategy is thus established.

## 7 Conclusion

An ANFIS-RES controller has been designed and developed to regulate the speed of a field-oriented control (FOC)-based PMSM drive. The controller was evaluated on a laboratory prototype to validate its performance under various operating conditions. The experimental and simulation results confirm that the proposed ANFIS-RES controller outperforms the conventional PI-RES approach, delivering enhanced steady-state and dynamic performance. By replacing the linear PI component with a fuzzy logic-based ANFIS design, the methodology achieves an optimal balance between the flexibility of fuzzy control and reduced implementation complexity, making it adaptable and efficient.

Key findings from this research include:

The ANFIS-RES controller exhibits a faster speed response with zero overshoot under diverse conditions, such as startup, load changes, acceleration, and speed reversal.
It provides improved starting performance, meeting the demands of electric vehicle (EV) applications, particularly for frequent start-stop scenarios common in urban driving.
The controller demonstrates robustness against parameter variations, making it suitable for the stringent requirements of vehicular systems.

The proposed strategy offers a faster dynamic response with reduced torque ripple and minimal mechanical vibration, addressing critical factors for passenger comfort and drive efficiency. Additionally, the controller maintains consistent performance despite variations in stator resistance, further establishing its suitability for EV applications compared to conventional control schemes.

**Acknowledgements** This work is supported by the Centre of Excellence for Electric Vehicle and Related Technologies, Delhi Technological University.

## Declarations

**Conflict of interest** The authors declare that the content in this manuscript is original and does not come from any other source. There are no financial or scientific conflicts of interest pertaining to the research reported in this publication for any of the authors.

## References

1. Zhu, Z.Q.; Howe, D.: Electrical machines and drives for EV electric, hybrid and fuel cell vehicles. Proc. IEEE **95**(4), 746–765 (2007)
2. U. S. Environmental Protection Agency, "Sources of greenhouse gas emissions," Accessed: Aug. 17, 2021. [Online]. Available: https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions
3. Chan, C.C.: The state of the art of electric, hybrid and fuel cell vehicles. Proc. IEEE **95**(4), 704–718 (2007)
4. Buyukdegirmenci, V.T.; Bazzi, A.M.; Krein, P.T.: Evaluation of induction and permanent magnet synchronous machine using drive-cycle energy and loss minimization in traction applications. IEEE Trans. Ind. Appl. **50**(1), 395–403 (2014)
5. Gebregergis, A.; Chowdhury, M.H.; Islam, M.S.; Sebastian, T.: Modeling of the permanent-magnet synchronous machine including torque ripple effects. IEEE Trans. Ind. Appl. **51**(1), 232–239 (2015). https://doi.org/10.1109/TIA.2014.2334733
6. Chau, K.T.; Chan, C.C.; Liu, C.: Overview of permanent-magnet brushless drives for electric and hybrid electric vehicles. IEEE Trans. Ind. Electron. **55**(6), 2246–2257 (2008)
7. Uddin, M.N.; Radwan, T.S.; Rahman, M.A.: Performance of interior permanent magnet motor drive over wide speed range. IEEE Trans. Energy Convers. **17**(1), 79–84 (2002)
8. Inoue, Y.; Morimoto, S.; Sanada, M.: Comparative study of PMSM drive systems based on current control and direct torque control in flux-weakening control region. IEEE Trans. Ind. Appl. **48**(6), 2382–2389 (2012)
9. Hussain, S.; Bazaz, M. A.: Comparative analysis of speed control strategies for vector controlled PMSM drive. In: International Conference on Computing, Communication and Automation (ICCCA), Noida, India, 2016, pp. 1314–1319
10. Sun, X.; Cao, J.; Lei, G.; Guo, Y.; Zhu, J.: A composite sliding mode control for SPMSM drives based on a new hybrid reaching law with disturbance compensation. IEEE Trans. Trans. Electrif. **7**(3), 1427–1436 (2021)
11. Gao, J.; Wu, X.; Huang, S.; Zhang, W.; Xiao, L.: Torque ripple minimization of permanent magnet synchronous motor using a new proportional resonant controller. IET Power Electron **10**(2), 208–214 (2017)
12. Xia, C.; Ji, B.; Yan, Y.: Smooth speed control for low-speed high-torque permanent-magnet synchronous motor using proportional–integral–resonant controller. IEEE Trans. Ind. Electron. **62**(4), 2123–2134 (2015)

13. Suryakant, S.; Sreejeth, M.; Singh, M.: Performance analysis of PMSM drive using hysteresis current controller and PWM current controller. In: 2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), Bhopal, (2018)
14. Uddin, M.N.; Abido, M.A.; Rahman, M.A.: Development and implementation of a hybrid intelligent controller for interior permanent magnet synchronous motor drive. IEEE Trans. Ind. Appl. **40**(1), 68–76 (2004)
15. Ranjan, S.; Sreejeth, M.; Singh, M.: Analysis of the performance of an HCC-based PMSM drive using adaptive neuro-fuzzy inference. In: 2022 2nd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2022, pp. 1–4. https://doi.org/10.1109/CONIT55038.2022.9848167
16. Yang, Y.; Vilathgamuwa, D. M.; Rahman, M. A.: Implementation of an artificial-neural-network-based real-time adaptive controller for an interior permanent-magnet motor drive. In: IEEE Transactions on Industry Applications, vol. 39, No. 1, pp. 96–104 (2003)
17. Li, S.; Won, H.; Fu, X.; Fairbank, M.; Wunsch, D.C.; Alonso, E.: Neural-network vector controller for permanent-magnet synchronous motor drives: simulated and hardware-validated results. IEEE Trans. Cybern. **50**(7), 3218–3230 (2020)
18. Zhang, Q.; Guo, H.; Guo, C.; Liu, Y.; Wang, D.; Lu, K.; Zhang, Z.; Zhuang, X.; Chen, D.: An adaptive proportional-integral-resonant controller for speed ripple suppression of PMSM drive due to current measurement error. Int. J. Electr. Power Energy Syst. **129**, 106866 (2021)
19. Pillay, P.; Krishnan, R.: Modeling, simulation, and analysis of permanent-magnet motor drives. I. The permanent-magnet synchronous motor drive. IEEE Trans. Ind. Appl. **25**(2), 265–273 (1989)
20. Sankararao, E.; Gandhi, R.; Chakraborti, A.; Roy, R.: Comparative analysis of model predictive and hysteresis current control for PMSM drive application. In: 2022 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES), Jaipur, India, 2022, pp. 1–6. https://doi.org/10.1109/PEDES56012.2022.10080791
21. Abassi, M.; Khlaief, A.; Saadaoui, O.; Chaari, A.; Boussak, M.: Performance analysis of FOC and DTC for PMSM drives using SVPWM technique. In: 2015 16th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA), 2015, pp. 228–233
22. Ranjan, S.; Sreejeth, M.: Minimization of ripples in torque of hysteresis current controlled PMSM using PI-RES controller. In: 2021 2nd International Conference for Emerging Technology (INCET), Belagavi, India, 2021, pp. 1–6. https://doi.org/10.1109/INCET51464.2021.9456269
23. Suryakant; Sreejeth, M.; Singh, M.; Seth, A. K.: Minimization of torque ripples in PMSM drive using PI-resonant controller-based model predictive control. Electr. Eng. **105**(1), 207–219 (2023). https://doi.org/10.1007/s00202-022-01660-y
24. Rahman, A.; Hoque, M.A.: On-line adaptive artificial neural network-based vector control of permanent magnet synchronous motors. IEEE Trans. Energy Convers. **13**(4), 311–318 (1998)
25. Gaur, P.; Singh, B.; Mittal, A. P.: Artificial neural network based controller and speed estimation of permanent magnet synchronous motor. In: 2008 Joint International Conference on Power System Technology and IEEE Power India Conference, New Delhi, India, 2008, pp. 1–6. https://doi.org/10.1109/ICPST.2008.4745351
26. Alasvandi, M.; Moussavi, S.; Morad, E.; Rasouli, E.: ANFIS based IMC PID controller for permanent magnet DC motor. In: Proceedings of the 16th International Conference on Informatics in Control, Automation and Robotics Prague, 2019, pp. 235–242
27. Salem, W. A. A.; Osman, G. F.; Arfa, S. H.: Adaptive neuro-fuzzy inference system based field oriented control of PMSM & speed estimation. In: Twentieth International Middle East Power Systems Conference (MEPCON), Cairo, Egypt, 2018, pp. 626–631
28. Sreejeth, M.; Singh, M.; Kumar, P.: Sensitivity analysis of induction motor using fourth order slip estimator. Int. J. Power Energy Syst. **33**(2), 2013
29. Sreejeth, M.; Singh, M.; Kumar, P.: Particle swarm Optimisation in efficiency improvement of vector-controlled surface mounted permanent magnet synchronous motor drive. IET Power Electron. **8**(5), 760–769 (2015)

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
APPLIED PACKAGING RESEARCH

RESEARCH ARTICLE

# Biodegradable Polymeric Film for Food Packaging

**Raminder Kaur**
Delhi Technological University

**Neelam Yadav***
Delhi Technological University

## ABSTRACT

*Food packaging films that show post-consumer biodegradability are rarely explored by researchers. The present study was carried out to investigate the physicomechanical characteristics of ethyl cellulose (EC) based films to be used in food packaging. Ethyl cellulose was plasticized with different percentage of polyethylene glycol (PEG). The samples of standard dimensions were subjected to different testing such as soluble matter content, moisture content, oil permeability, surface morphology, mechanical testing etc. The data obtained was analysed to decide the moderate percentage of plasticizer that can be used to provide a rational explanation of a perfect quality specimen. It has been revealed that too high or too low percentage of plasticizer was not appropriate for a good film. Tensile stress analysis was used to estimate the mechanical properties of the films. The results have shown an increase in tensile properties with increase in coalescence temperature. Temperature-dependent plastic deformation was observed for coalescence temperatures above 50°C.*

## KEY WORDS

*Ethyl cellulose, Plasticizer, Biopolymers, Food packaging.*

***Raminder Kaur**
*Corresponding Author*
rkwaliadtu@gmail.com

## INTRODUCTION

Packaging protects food from the external environment and provides information to the consumers about the food inside the package. Foods are packaged to deliver safe, wholesome and nutritious food to the consumers and it gives a high significance of link between packaging and consumer protection [1]. Plastics commonly used in food packaging include polyethylene (both HDPE and LDPE), polypropylene, polyvinyl chloride, polyethylene terephthalate and polystyrene [2].These plastics used in food packaging, are derived from petroleum and generate greenhouse gases, thus cause a huge impact on environment. Moreover, they also interact with food material and may release toxic chemicals into food [3]. In addition, the quality of the packaged food is likely to be deteriorate during its storage due to change in its natural ingredients, which may include the loss in the moisture content or volatile odour. However, this deterioration can be reduced by selecting passable food packaging or coatings. Many researchers have reported the enrichment of these packaging films by the addition of other active agents such as $O_2$ and $CO_2$ absorbers, antimicrobial agents and antioxidants etc. in order to enhance their utility [4-7]. However, the food packaging films which show post-consumer biodegradability, without causing much burden on the environment were not much explored [8]. Like other polymeric products [9-15], researches are now focused on the development of the bio-based performance end products in the food packaging, with an attention to make them environment friendly. Most widely used bio-based polymers are the cellulose derivatives, such as methyl cellulose, cellulose acetate, hydroxypropyl cellulose and ethyl cellulose, that, have a backbone of cellulose with repeating structure of anhydrous glucose units, each having three reactive -OH (hydroxyl) sites [16]. Ample of work has been done in the past for the preparation and investigation of modified atmosphere packaging films based on protein, lipid and polysaccharide, which provide barrier to the volatile component of food as well as its moisture content [17,18]. However, the main concern is that, these bio-degradable packaging films lack many desirable characteristics of packaging materials to be used at a commercial scale. Various approaches may be used to overcome these problems, including the modification of raw material or the incorporation of plasticizers to improve their film forming characteristics. The type of plasticizer and its concentration can modify the different physical properties of the films [19]. The present study has been conducted to synthesize ethyl cellulose based bio-degradable films with improved physical properties by the incorporation of polyethylene glycol as plasticizer. The ethyl cellulose films were synthesized by taking different ratios of ethyl cellulose and polyethylene glycol and their various properties are studied.

## MATERIALS AND METHOD

### 2.1 Materials

Ethyl cellulose containing ethoxy content of about 48-49.5% was used for film formation. Ethanol (99.9%) was used as solvent for dissolving ethyl cellulose to form a 2.5% w/v solution with viscosity about 18-24 cp. Polyethylene glycol (MW 400) was used as plasticizer. All the chemicals used were of AR grade and were supplied by CDH (P) Ltd., New Delhi.

### 2.2 Method

#### 2.2.1 Preparation of Ethyl Cellulose Film

All films were prepared by simple casting method. A 2.5% (w/v) solution of ethyl cellulose in ethanol was prepared by constant stirring at 100 rpm for 20-25 minutes using a magnetic stirrer at room temperature. After the formation of a homogenized solution, polyethylene glycol (PEG); as a plasticizer was added to it. The ethyl cellulose and polyethylene

glycol used were in the ratio 4:1, 2:1 and 4:3 for different samples i.e. Sample 1, 2 and 3 respectively. The cast solution was then gently poured into a petridish (diameter = 4 cm) avoiding any bubble formation. The content of the petridish was then dried overnight at room temperature on a levelled surface, thus, resulting in a spotless film that was manually peeled unbroken from the casting surface.

### 2.3 Assessment of Ethyl Cellulose Films for their Suitability as Food Packaging Material

The samples of ethyl cellulose films containing different ratio of PEG were assessed for the physical and chemical properties so as to be used in a food packaging. Five replicates of each sample was tested for assuring the reproducibility of the results.

#### 2.3.1 Film Appearance

The appearance of the ethyl cellulose films having different ratio of polyethylene glycol was assessed by its texture and uniformity.

#### 2.3.2 Film Thickness

All ethyl cellulose films i.e. Sample1, Sample 2 and Sample 3 were acclimatised at 24°C ± 2°C and 75% RH for 24 hours and the thickness was determined using micro-meter, which recorded upto exactness of ± 0.001mm. The thickness measurements was done at 6 diverse points of the ethyl cellulose films. Mean value was used to estimate the thickness of the resulted ethyl cellulose films [20, 21].

#### 2.3.3 Soluble Matter (SM)

The soluble matter percentage of the resulted films were estimated by immersing the samples (2.0 cm in diameter each) in 50 ml of distilled water, which was slowly stirred for a duration of 24 hrs. The samples were removed and dried in oven for 24 hrs. The following equation was used to calculate the amount of soluble matter in the resulted films [22].

$$\textbf{\textit{Soluble matter \%}} = \left(\frac{\textit{Initial dry weight} - \textit{Final dry weight}}{\textit{Initial dry weight}}\right) \boldsymbol{x}\ \mathbf{100} \quad (1)$$

#### 2.3.4 Moisture Content (MC)

The moisture content of all samples was calculated by measuring the weight loss of films by drying it in a furnace at 105°C ± 3°C for approximately 12 hrs, till a constant weight had been achieved (dry sample weight). Moisture content was determined as percentage loss in mass during drying of the films using following equation [23, 24].

$$\textbf{\textit{Moisture content}}\ (\%) = \left(\frac{\textit{Intial mass} - \textit{Final mass}}{\textit{Initial mass}}\right) * \mathbf{100} \quad (2)$$

#### 2.3.5 Scanning Electron Microscopy (SEM)

The microstructural examination of ethyl cellulose films was done using a scanning electron microscope (SEM). Film samples to be tested for SEM were confirmed for no water present in them. For the purpose, they were kept in desiccators with P2O5 for 2 weeks. Further, ethyl cellulose films were frozen using liquid nitrogen and then a cryo-fractured cross section of these films were gold sputtered (*placing gold layer of about 100 A° so as to escape charging up of the films under electron beam*) and then observed under scanning electron microscope (Hitachi S3700 SEM) at 10 kV as accelerating voltage [21].

#### 2.3.6 Mechanical Properties

The wet films were dried in a levelled- forced air drying oven, pre-set at a desired coalescence temperature (The temperature at which the components of miscible solution attract each other) for 24hrs [25]. Mechanical properties like tensile strength ($\sigma_{ut}$), elongation at break (EAB %), elastic modulus (E) and work of failure for these ethyl cellulose films were determined by ASTM method, using an Universal Testing Machine (Micro UTM KLA Tencor T150 UTM with a gauge length of 2 mm to 10 mm and a load variable of 50mN to 20 kN.).The samples of each ethyl cellulose film was

cut into a rectangular shape having a measurement of 60 mm x 60 mm and with an average thickness of 1 mm. The samples was griped on the device with a gap of 1 mm and pulled away at a speed of 0.1 mm/s and load of 1 N. The curves plotted between stress (Tension force (F) / cross sectional area(S)) and strain (change in length (**ΔL**) /original length (**$L_0$**) were used to determine the tensile properties as per equation 3 and 4, given below. [24, 26].

$$\sigma_{ut} = F/S \quad (3)$$

$$EAB = \Delta L / L_0 \times 100\% \quad (4)$$

Equation 3 was used to calculate the elastic modulus of the ethyl cellulose films, which was calculated by taking the ratio of the maximum force exerted on the film by the cross sectional area and is represented in MPa. Equation 4 represents the percentage change in strain.

Equation 5 is used to calculate the elastic modulus which is the ratio of stress and strain at elastic limit

$$E = \sigma/\epsilon \quad (5)$$

Where, σ denotes stress and ϵ denotes strain at elastic limit

(6)

**Work of failure = (Area under curve x cross head speed)/ (film thickness x film width)**

Work of failure represented by Equation 6 is a function of work done in breaking the film specimen and is representative of the film thickness.

Applied load is an important parameter to be considered, especially when comparing the results of two studies, since the strain or percent elongation of a polymeric film is normally inversely proportional to the rate of strain and the parameter changes due to various reasons namely sample preparation method, coalescence temperature and also on coating process variables.

### 2.3.7 Anti-permeability for Oils

To test the anti-permeability of ethyl cellulose films (incorporated with different percentage of PEG) for liquid oils and oily foods, the samples of resulted films of size 2 cm x 2 cm was firmly pressed on the open end of the test tube (inner diameter- 25 mm, outer diameter-27 mm) containing liquid oil (soya bean oil). Further, the test tube was placed upside down on the filter paper, and was observed for the permeability of the oil through the ethyl cellulose films as time went. Observations were repeated on 5 samples of similar composition, for a duration of 3 days storage under normal conditions of temperature, pressure and relative humidity [27].

### 2.3.8 Reliability, Collapsibility and Water-solubility of Ethyl Cellulose films

The samples of the ethyl cellulose films were considered reliable, as they maintained the integrity of appearance, when peeled from petri dish. Also, the samples of resulted ethyl cellulose films were creased for 20 times at the same line to test its plasticity or collapsibility. Water solubility of the samples were examined by immersing a square piece of each film in 100 ml of distilled water, without stirring, at a temperature of 25°C. Water solubility of ethyl cellulose film was estimated by observing the time taken for the film to totally dissolve in water [21].

## RESULTS AND DISCUSSIONS

### 3.1 Film appearance

The samples of ethyl cellulose films formed on a flattened surface was peeled off unspoiled from the petri-dish. The appearance of the sample 1 was hazy and that of sample 3 was translucent while the appearance of sample 2 was opaque, as seen from Figure 1. However the texture of all the samples was smooth and homogeneous and showed no breakage while peeling from the surface. These same samples were used for further testing.



**(a)**



**(b)**



**(c)**

*Fig. 1: Appearance of surface of Ethyl cellulose film containing different ratio of ethyl cellulose and polyethylene glycol in ethanol (a) EC: PEG-4:1 (b) EC: PEG-2:1 (c) EC: PEG-4:3*

*Table 1: Mean values of weights (mg) and soluble matter (%) of various samples*

| SAMPLE | EC : PEG | INITIAL WEIGHT(sd) | FINAL WEIGHT (sd) | SOLUBLE MATTER (sd) |
|---|---|---|---|---|
| 1 | 4:1 | 0.041 (0.00018) | 0.016 (0.00028) | 60.97 (0.00023) |
| 2 | 2:1 | 0.032 (0.00029) | 0.014 (0.00035) | 56.25 (0.00032) |
| 3 | 4:3 | 0.052 (0.000167) | 0.013 (0.00061) | 75 (0.00039) |

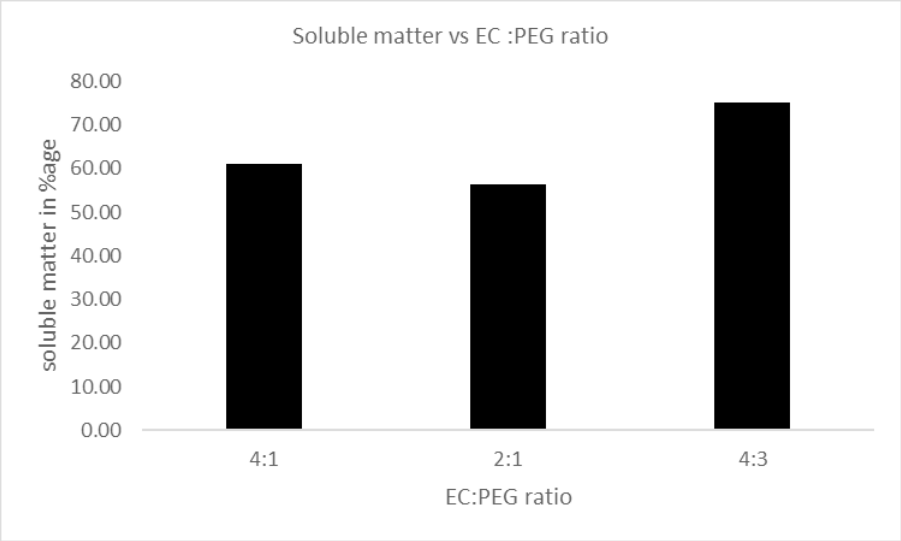



*Fig. 2: Soluble Matter Content for Ethyl Cellulose Films of Different EC: PEG Compositions*

### 3.2 Films Thickness

The samples of resulted films were measured for the film thickness at 6 different points (centre of each film, at centre of each quarter of the film and edges of the film). Average value of all the points was calculated to determine the thickness of each ethyl cellulose film. The mean thickness of various samples was measured as 1.010 mm with a standard deviation of 0.01, 1.017 mm with standard deviation of 0.04 and 1.014 mm with a standard deviation of 0.02 for sample 1 (EC:PEG-4:1), sample 2 (EC:PEG-2:1)and sample 3 (EC:PEG-4:3) respectively.

### 3.3 Soluble Matter (SM)

The samples of the resulted ethyl cellulose films immersed in distilled water for 24 hrs, showed no swelling, as well as, they maintained their structure. The result of soluble matter among different categories of ethyl cellulose film was maximum for the ethyl cellulose film containing EC: PEG ratio 4:3, which showed approximately 75.00 % total soluble matter and minimum for the ethyl cellulose film containing EC: PEG ratio 2:1, having approximately 56.25% total soluble matter. However, the soluble matter for ethyl cellulose film containing EC: PEG ratio 4:1 showed the total soluble matter of about 60.97 %. The mean values of initial and final weights (with standard deviation) for each sample their Percentage of soluable matter is as given in Table 1. The soluble matter content for ethyl cellulose films of different EC: PEG compositions is as shown in Figure 2.

Sample of ethyl cellulose film with EC: PEG 2:1 ratio showed a slight lower moisture percentage (about 10% decrease) as compared with the sample having EC: PEG 4:1 ratio, while the ethyl cellulose film having EC: PEG 4:3 ratio showed a great increase in moisture content. Figure 3 gives the moisture content for ethyl cellulose films of different EC: PEG

*Table 2: Mean values of weights (mg) and moisture content (%) of different EC: PEG Compositions*

| SAMPLE | EC:PEG | INITIAL WEIGHT (mg) | FINAL WEIGHT (mg) | MOISTURE CONTENT (%) |
|---|---|---|---|---|
| 1 | 4:1 | 0.043 | 0.016 | 62.79 |
| 2 | 2:1 | 0.022 | 0.011 | 50.00 |
| 3 | 4:3 | 0.044 | 0.014 | 68.18 |

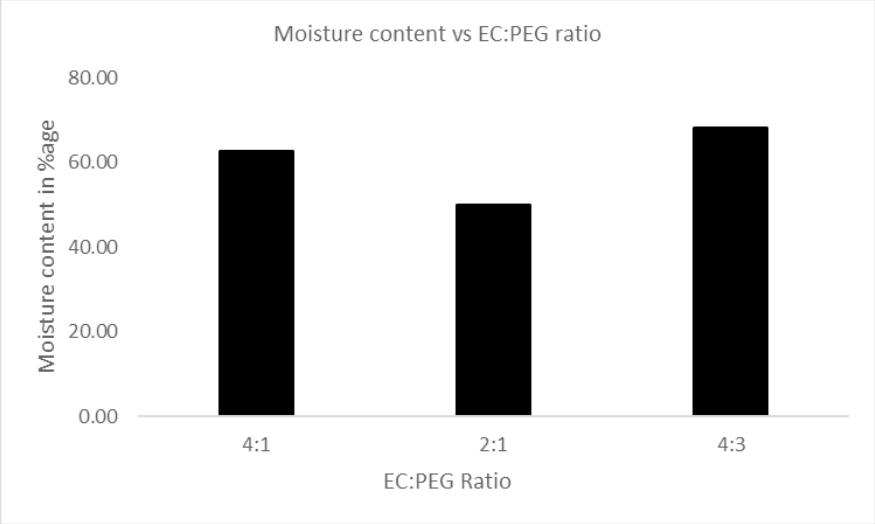


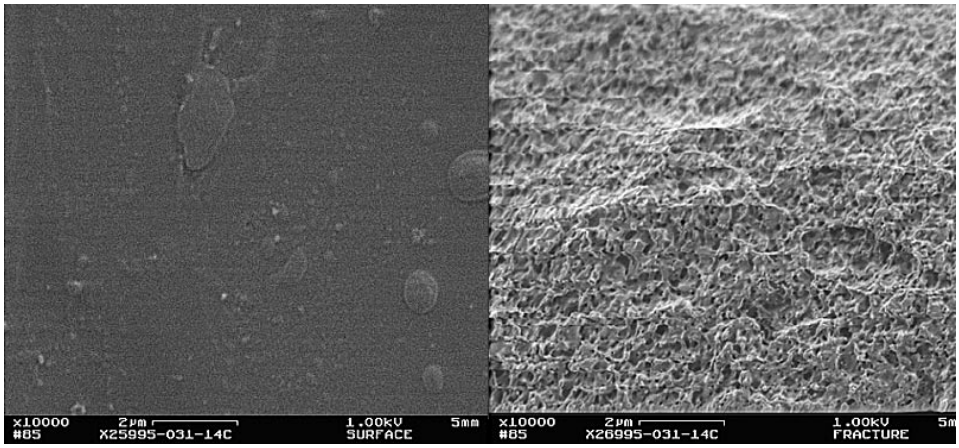

*Fig. 3: Moisture Content for Ethyl Cellulose Films of Different EC: PEG Compositions*

compositions.The low moisture percentage of EC: PEG 2:1 can be explained as a result of interference of water structure in the binding condition. The mean values of weights (mg) and moisture content (%) of different EC: PEG compositions are given in Table 2.

### 3.5 Scanning Electron Microscopy (SEM)

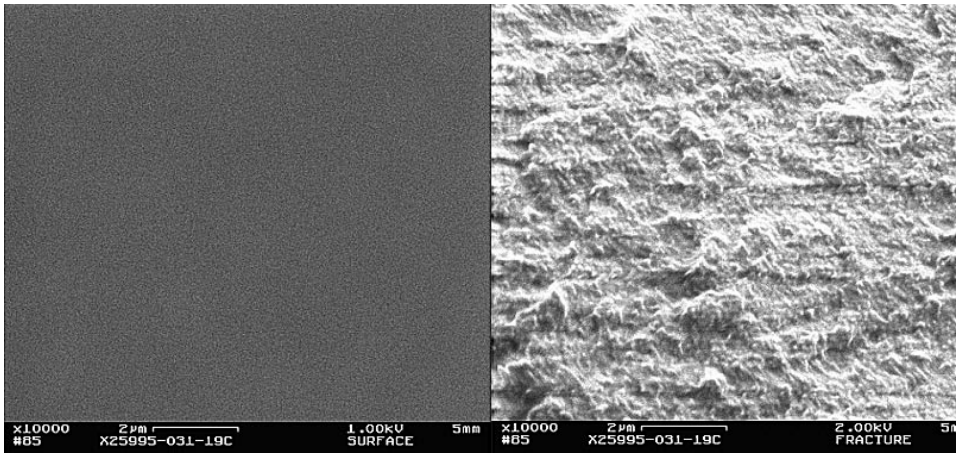
The microstructural examination of the ethyl cellulose films was qualitatively studied using scanning electron microscopy (SEM). The images of the surface and cross section of EC films revealed the effects of addition of different percentage of PEG. Figure 4 show the SEM images of PEG plasticized ethyl cellulose films with EC: PEG ratio 4:1, EC: PEG ratio 2:1 and EC: PEG ratio 4:3 respectively. The surface image of both ethyl cellulose films with EC:PEG ratio 4:1 and 2:1 showed unbroken, compact and smooth traits without any micro-cracks, which meant polyethylene glycol has been easily incorporated into ethyl cellulose as plasticizer. However the EC film with ratio 4:3 showed few cracks and the surface was not apparently smooth, which might be due to decreased polymer-polymer cohesiveness, with increases content of plasticizer, leading to the deterioration of the quality of film. After observing all the three films, it is inferred that the EC: PEG ratio of 2:1 and EC: PEG ratio of 4:1 both are considered suitable for uniform and good quality film.

The structure of the film is affected by the organization of various constituents in the film making solution and the interaction taking place between the constituents during the drying of the ethyl cellulose film. The microstructural analysis of the film is also helpful to know the mechanism of water vapor permeability across the films and the arrangement of components in the film and their effect on mechanical properties of the film [28].



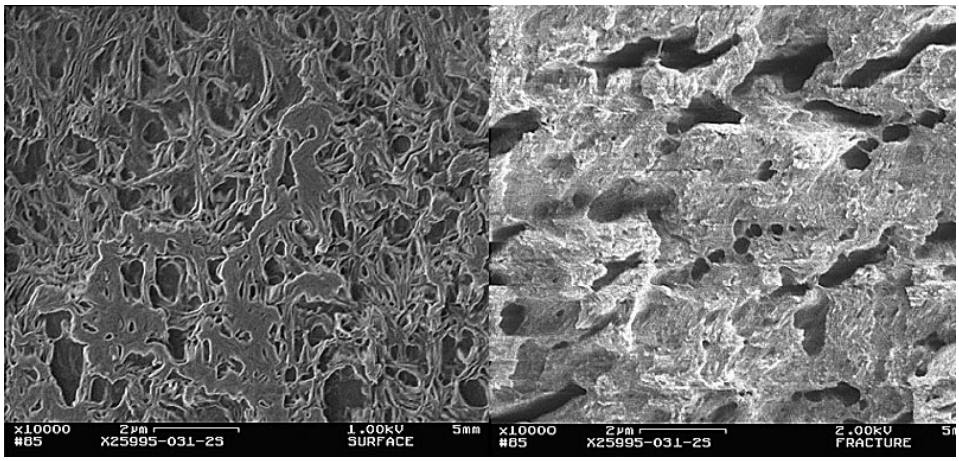

(i) (ii)

(a) EC: PEG - 4:1




(i) (ii)

(a) EC: PEG - 2:1




(i) (ii)

(a) EC: PEG - 4:3

*Fig. 4: SEM Images of PEG 400 Plasticized Ethyl Cellulose Films (a) EC: PEG ratio: 4:1; (i) surface of the film, (ii) cross section of the film (b) EC: PEG ratio: 2:1; (i) surface of the film, (ii) cross section of the film. (c) EC: PEG ratio: 4:3; (i) surface of the film, (ii) cross section of the film.*

*Table 3: Mean values of tensile strength of various samples at different coalescence temperature*

| Coalescence temperature in (°C) | Tensile strength in (MPa) | | |
|---|---|---|---|
| | SAMPLE 1 (sd) | SAMPLE 2 (sd) | SAMPLE 3 (sd) |
| 30 | 0.007 (0.00015) | 0.014 (0.001) | 0.013 (0.00125) |
| 40 | 0.013 (0.0017) | 0.015 (0.00146) | 0.014 (0.00115) |
| 50 | 0.016 (0.00114) | 0.017 (0.00126) | 0.016 (0.00321) |
| 60 | 0.016 (0.00125) | 0.017 (0.00111) | 0.015 (0.00167) |
| 70 | 0.0149 (0.00143) | 0.0156 (0.00129) | 0.0149 (0.00125) |

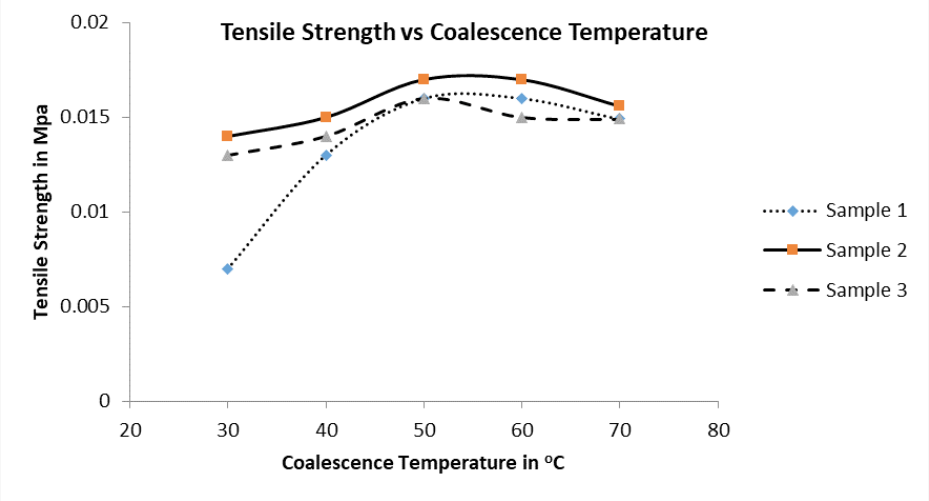



*Fig. 5: Plot of Tensile Strength Vs Coalescence Temperature (Sample1-EC: PEG - 4:1; Sample 2- EC: PEG- 2:1; Sample 3 – EC: PEG- 4:3).*

### 3.6 Mechanical Properties

Characteristic features of the load-time or typical stress-strain profiles from tensile tests are routinely used to characterize the polymer properties. Toughness is directly proportional to the area under a load time or load-displacement curve which is quantitated as energy. The higher the amount of energy a polymeric film can absorb prior to break under load, the higher is its toughness. Elastic modulus is an indicator of the elasticity of the film, with lower values corresponding to higher elasticity. On the other hand, tensile strength is an indicator of the film strength, with larger values corresponding to the stronger films. Ideally a film coating must be both strong and elastic and, therefore, must yield a higher tensile strength-to-elastic modulus ratio. This ratio has been used as an indicator of the overall mechanical quality of the films.

Figure 5 depicts the plot of tensile strength as a function of coalescence temperature for the samples prepared. The tensile strength of the films increased with a corresponding increase in coalescence temperature and did not reach a plateau within the temperature range used in this investigation. It was observed that sample 2 showed best characteristics in comparison to other samples. At temperatures above 60°C, no appreciable differences were observed in the tensile strengths of the films cast from all formulations. This could possibly be due to evaporation of the plasticizers from the films at elevated temperatures. Mean values of tensile strength of various samples at different coalescence temperatures is as given in Table 3.

Mean values of work of failure of various samples at different coalescence temperatures are given in Table 4. The values of work of failure

*Table 4: Mean values of work of failure of various samples at different coalescence temperature*

| Coalescence temperature in °C | Work of failure (Nm/sec) | | |
|---|---|---|---|
| | SAMPLE 1 (sd) | SAMPLE 2 (sd) | SAMPLE 3 (sd) |
| 30 | 5 (0.67) | 5.4 (0.78) | 0.1 (0.0124) |
| 40 | 4 (0.45) | 5.9 (0.64) | 3 (0.21) |
| 50 | 5.5 (0.167) | 6.1 (0.51) | 5 (0.14) |
| 60 | 15 (0.253) | 19 (0.92) | 13 (1.5) |
| 70 | 30 (1.67) | 32 (1.12) | 17 (2.1) |

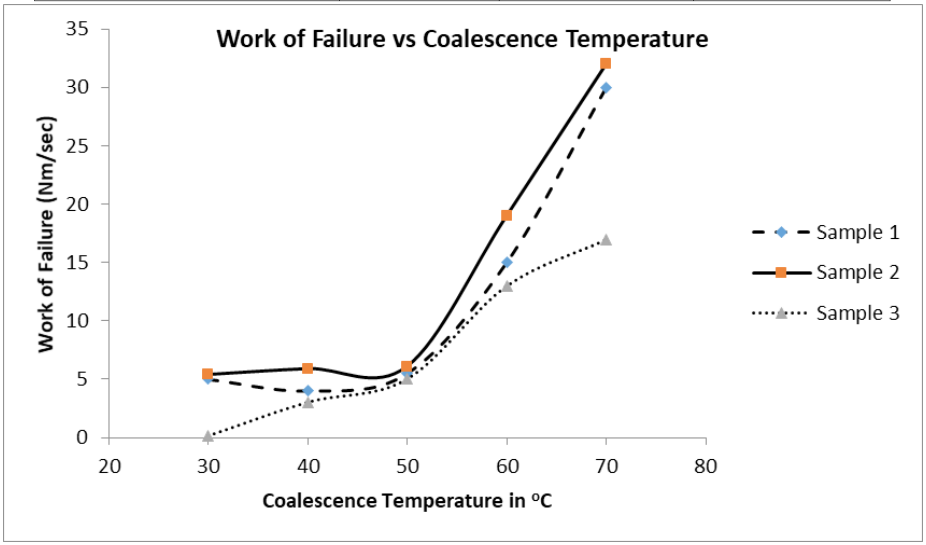



*Fig. 6: Work of Failure vs Coalescence Temperature of Different Ethyl Cellulose Films (Sample1-EC: PEG - 4:1; Sample 2- EC: PEG - 2:1; Sample 3 – EC: PEG - 4:3).*

plotted as a function of coalescence temperature for all the three samples are shown by Figure 6. No significant difference was observed in the work of failure for the films cast from formulation EC:PEG ratios 4:1 and 2:1. The work of failure value for the formulation with EC:PEG ratio 4:3 is almost approaching zero at lower coalescence temperature i.e. at 30°C. At higher coalescence temperatures, the higher values of work of failure depict a significantly higher absorption of energy, possibly due to the complete penetrate of plasticizer molecules into the polymeric network. From the load-time profiles, it was observed that the films at and above 50°C exhibited distinct and prolonged plastic deformation with the resultant increase in the area under the curve. This could be due to strong interaction between the neighboring polymer particles at and above the coalescence formulation.

The influence of coalescence temperature on the elastic modulus values of the films is shown in Figure 7. Plot of applied load with time is used to calculate the value of elastic modulus. Higher value of load carrying ability indicate a high value of elastic modulus, thus reflects the strength of the film. In general, it is observed that as the coalescence temperature get increased, the elastic modulus of the films increased almost linearly, for all the samples of the polymeric films, until they achieved a peak. A decline in the values of elastic modulus values was further observed, at elevated coalescence temperatures, which is due to the softening of the film at elevated temperature. Mean values of load

*Table 5: Mean values of load carrying capacities for various time periods*

| Time in (sec) | Applied load (Elastic Modulus in MPa) | | |
|---|---|---|---|
| | Sample 1 | Sample 2 | Sample 3 |
| 0 | 0.3 (91.52) | 0.5 (152.54) | 0.31 (94.57) |
| 2 | 0.3 (91.52) | 0.9 (274.57) | 0.31 (94.57) |
| 5 | 0.35 (106.78) | 1.5 (457.62) | 0.32 (97.62) |
| 10 | 1.1 (335.58) | 2.2 (671.17) | 0.97 (295.92) |
| 12 | 0.3 (91.52) | 2.5 (762.69) | 1.5 (457.62) |
| 15 | 0.3 (91.52) | 2.6 (793.2) | 1.8 (549.42) |
| 20 | 0.31 (94.57) | 2.4 (732.18) | 0.29 (88.47) |
| 25 | 0.3 (91.52) | 0.5 (152.54) | 0.29 (88.47) |
| 30 | 0.31 (94.57) | 0.5 (152.54) | 0.295 (90) |
| 35 | 0.3 (91.52) | 0.5 (152.54) | 0.29 (88.47) |
| 40 | 0.31 (94.57) | 0.51 (155.59) | 0.295 (90) |

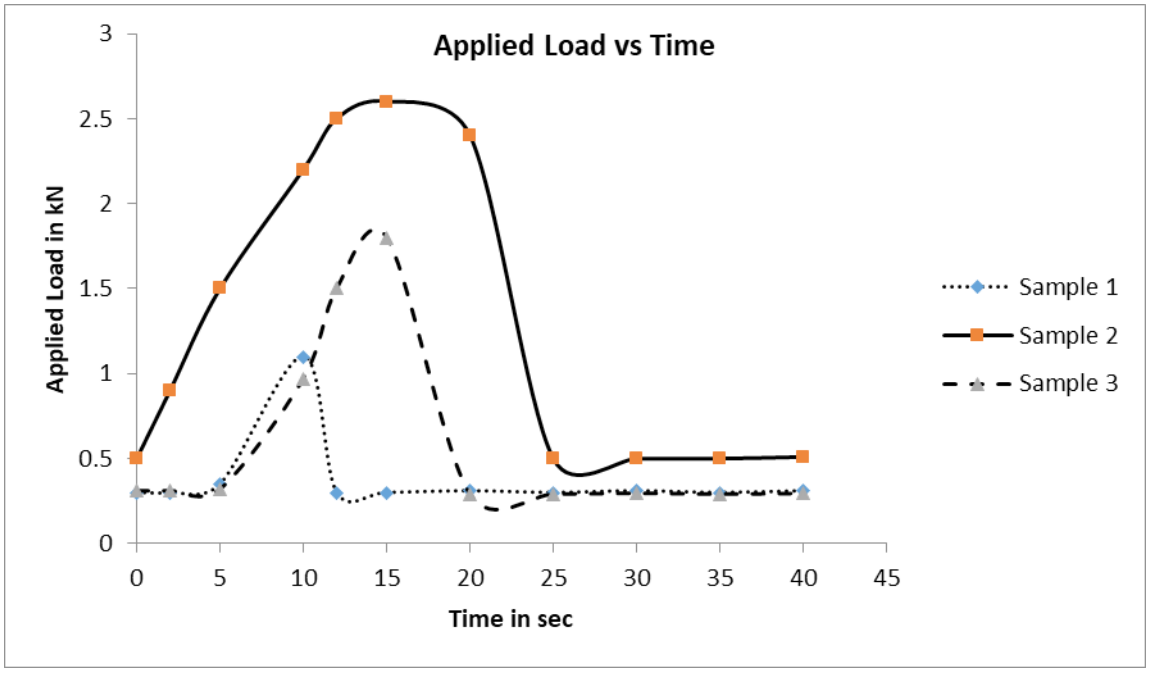



*Fig. 7: Applied Load Vs Time of Different Ethyl Cellulose Films (Sample1-EC: PEG - 4:1; Sample 2- EC: PEG - 2:1; Sample 3 – EC: PEG - 4:3) used for calculation for elastic modulus.*

*Table 6: Comparison of properties of ethyl cellulose film with different EC: PEG ratios*

| SAMPLE | EC:PEG | APPEARANCE | THICKNESS (mm) | COLLAPSIBILITY (folding at a point) |
|---|---|---|---|---|
| 1 | 4:1 | Hazy | 1.02 | NO FISSURE |
| 2 | 2:1 | Opaque | 1.01 | NO FISSURE |
| 3 | 4:3 | Translucent | 1.01 | SLIGHT FISSURE |

carrying capacities for different ethyl cellulose films at various time periods is as given in Table 5.

The independent interpretation of these three parameters, i.e., tensile strength, work of failure, and elastic modulus may not always lead to meaningful conclusions with respect to the in situ performance of the films. The mechanical properties of the films was observed to be dependent on the type of the plasticizer and the coalescence temperature. With an increase in the coalescence temperature, the films exhibited a characteristic transition in their mechanical properties, i.e., from a soft and weak system(as evident from lower values of tensile strength, strain at break and elastic modulus) to a hard and tough system (as evident from higher values of tensile strength, strain at break, and elastic modulus). The optimal values of the tensile strength-to-elastic modulus ratio obtained for films of all three formulations dried at 50°C indicate that the films coalesced at or around this temperature may be less susceptible to physical defects.

### 3.7 Anti-permeability for Oil

The bio polymeric films used for packaging food, especially oily foods or liquid oil containing edibles, should be anti-permeable for oils, otherwise leakage of oil from packaging films will cause decline in the appearance, flavor and shelf life of packaged foods. So, the anti-permeability test of different sample of ethyl cellulose was carried out. The result showed no trace of oil leakage through the ethyl cellulose films, thus offering good anti-permeability for oil [29].

### 3.8 Reliability, Collapsibility and Water-solubility of Ethyl Cellulose Films

Samples of ethyl cellulose film formed by casting method showed complete integrity while peeling off from the surface of the petri-dish. The thickness of the resulted films are as reported in Table 1. The folding test performed for finding the collapsibility of ethyl cellulose films showed that 20 times folding of film at a particular trace did not showed any fissure in sample1 and sample 2 but in sample 3 slight marking was observed. The visual test showed that the sample 2 had a good flexibility of all the films. Further the water solubility of the samples of ethyl cellulose films with different ratio of EC: PEG also indicated a delayed water solubility in sample 2 in comparison to sample 1 and sample 3. The comparison of the properties of ethyl cellulose films with different EC: PEG ratios is as given in Table 6.

## CONCLUSION

The ethyl cellulose films with different amount of polyethylene glycol, as plasticizer, had been casted to be used as a bio-degradable food packaging material. Among the different EC: PEG ratios used, the film with EC: PEG ratio 2:1 was found to have better suitability for food packaging applications and showed a very good flexibility and collapsibility. The soluble matter and moisture content of the film with EC: PEG ratio 2:1 was found to be minimum, among all the films tested. The developed film showed a maximum tensile strength of 0.0165 MPa at a coalescence temperature of 50$^{o}$C and a maximum work of failure of magnitude 29.5 Nm/sec was observed at temperature of 70$^{o}$C. The

film also exhibited a maximum load of 2.65 kN for a substantially good amount of time. The film was hence found to be mechanically stable of handling sufficient amount of load. The three parameters, characterising mechanical strength of the film namely tensile strength, work of failure, and elastic modulus were found to be good for the reported polymeric films.

## CONFLICT OF INTEREST

The work presented in this manuscript has been carried out for the research purpose only and there is no conflict of interest involved in it.

## REFERENCES


[1] B. L. Ottoni , R. M. Deus, J. A. G. Junior, Â. M. G. de Carvalho and R. A. G. Battistelle, "Communication and Biodegradable Packaging Relationship: A Paradigm for Final Disposal", *Journal of Applied Packaging Research, vol 10* (1). 10-27, 2018.

[2] American Chemistry Council (ACC). Plastics Division: Resin Identification Codes.2010.http://www.americanchemistry.com/s_plastics/doc.asp?CID=1102 & DID=4644

[3] I.S. Arvanitoyannis, "Totally and partially biodegradable polymer blends based on natural and synthetic macromolecules: preparation, physical properties, and potential as food packaging materials, *Journal of Macromolecular Science", Reviews in Macromolecular Chemistry and Physics, vol 39*. 205–271, 1999.

[4] 4T. Labuza, W. M. Breene, "Applications of "active packaging" for improvement of shelf life and nutritional quality of fresh and extended shelf life foods", *J. Food Process. Preser, vol 13.1-69*, 1989.

[5] T. P. Labuza, "An introduction to active packaging for foods", *Food Technol., vol 50*. 68-71, 1996.

[6] S.K.Bajpai, C. Navin and L.Ruchi, "Water vapor permeation and antimicrobial properties of sago starch based films formed via microwave irradiation", *International Food Research Journal, vol 18*, 417-426, 2011.

[7] K.Petersen, P. V.Nielsen, G. Bertelsen, M. Lawther, M. B Olsen, N. H. Nilsson and G. Mortensen, "Potential of bio-based materials for food packaging", *Trends Food Sci. Technol., vol 10*. 52-68, 1999.

[8] J. M. Krochta and C. De Mulder-Johnston, "Edible films and biodegradable films: Challenges and Opportunities", *Food Technol, vol 51*. 61-74, 1997.

[9] M. Malik and R. Kaur, "Synthesis of NIPU by the carbonation of canola oil using highly efficient 5,10,15-tris(pentafluorophenyl) corrolato-manganese(III) complex as novel catalyst", *Polym. Adv. Tech., vol 29* (3). 1078-1085, 2018.

[10] A. Aggarwal, R. Kaur and R.S.Walia, "PU Foam Derived from Renewable Sources: Perspective on Properties Enhancement: An Overview", *Europ. Polym. Jl., vol 95*. 255-274, 2017.

[11] M.Kumar and R.Kaur, "Glass fiber reinforced rigid polyurethane foam: synthesis and characterization", *e-Polymers., vol 7 (*6). 517-521, 2017.

[12] M. Malik and R. Kaur, "Influence of aliphatic and aromatic isocyanates on the properties of polyether-ester polyol based PU adhesive system", *Polym. Eng. Sci., vol 58* (1).112-117, 2018.

[13] M. Malik and R. Kaur, ".Mechanical and Thermal Properties of Castor Oil–Based Polyurethane Adhesive: Effect of $TiO_2$ Filler", *Adv. Polym. Tech., vol 37*(1). 24-30, 2018.

[14] R. Kaur and M. Kumar, "Function of Silicon Oil in the Castor Oil Based Rigid Polyurethane Foams", *Jl. Polym. Eng., vol 33* (9). 857-861, 2013.

[15] M. Kumar and R. Kaur, "Effect of Different Formulations of MDI on Rigid Polyurethane Foams based on Castor Oil", *Int. Jl. Sci. Res. Review, vol 2* (1) Special Issue April – May. 29-42, 2013.

[16] G. S. Rekhi and S. S. Jambhekar, "Ethyl-cellulose -- a polymer review, *Drug development and industrial pharmacy", vol 21* (1).61-77, 1995.

[17] M. B. Perez-Gago and J.M. Krochta, "Lipid particle size effect on water vapor permeability and mechanical properties of whey protein-beeswax emulsion films", *J. Agric. Food Chem., vol 49.* 996-1002, 2001.

[18] H. M. Lai and G.W. Padua, "Water vapor barrier properties of zein films plasticized with oleic acid, *Cereal Chem.", vol 75.* 194-199, 1998.

[19] R.L. Carmen and B. Roland, "Mechanical and Water Vapor Transmission Properties of Polysaccharide Films, Drug Dev." *Ind. Pharm., vol 22* (12). 1201–1209, 1996.

[20] S Yuvaraj and N Rajeswari," Chitosan/ Modified Banana Epidermis Starch Composite Films for Food Packaging Applications" *Journal of Applied Packaging Research, vol 10* (4),14-27,2018.

[21] G. Chen, B. Zhang, J. Zhao, and H. Chen, "Development and characterization of food packaging film from cellulose sulfate, Food Hydrocoll.", *vol 35.* 476–483, 2014.

[22] E. M. Ciannamea, P. M. Stefani, and R. A. Ruseckaite, "Physical and mechanical properties of compression molded and solution casting soybean protein concentrate based films", *Food Hydrocoll., vol 38.* 193–204, 2014.

[23] T. Chowdhury and M. Das, "Moisture sorption isotherm and isosteric heat of sorption of edible films from blends of starch, amylose and methyl cellulose", *International Food Research Journal, vol 19* (4) 1669- 1678, 2012.

[24] ASTM. (1995). Standard test methods for water vapour transmission of materials. Standards designations: E96-95 Annual book of ASTM standards (pp. 406-413). Philadelphia, *PA: American Society for Testing and Materials.*

[25] N.H. Parikh, S.C. Porter and B.D.Rohera, "Tensile properties of free films cast from aqueous ethyl cellulose dispersions", *Pharmaceutical Research, Vol.10* (6) 810-815, 1993.

[26] W.Y.Weng, Z. Tao, G.M. Liu, W.J. Su, K. Dsako, M.Tanaka and M.J. Cao, "Mechanical, barrier, optical properties and antimicrobial activity of edible films prepared from silver carp surimi incorporated with ε-polylysine", *Packaging Technology and Science, vol 27* (1), 37-47, 2014.

[27] G. Hu, J. Chen and J. Gao, "Preparation and characteristics of oxidized potato starch films. Carbohydrate Polymers", *vol 76* (2). 291-298, 2009.

[28] S. Wang, X. Kuang, Li, B., Wu, X., and T. Huang, "Physical properties and antimicrobial activity of chilled meat pads containing sodium carboxy-methyl cellulose", *Journal of Applied Polymer Science, vol 127* (1). 612-619, 2013.

[29] M. R. De Moura, M. V. Lorevice, L. H. C. Mattoso and V. Zucolotto, "Highly stable, edible cellulose films incorporating chitosan nanoparticles", *Journal of Food Science, vol 76* (2). 25-29, 2011.

AIP Conference Proceedings

RESEARCH ARTICLE | JANUARY 16 2025

# CNN based system in vehicle for different road surface and lane markings

Kukku Yaduvanshi ✉; Narendra Kumar; Parikshit Dahiya; S. Indu; Sangamithr Chembil Sudheer

*AIP Conf. Proc.* 3233, 020013 (2025)
https://doi.org/10.1063/5.0244605

View Online Export Citation

Articles You May Be Interested In

Accuracy of traffic surveillance system for vehicle lane detection using YOLO with MNN respective to lane violation

*AIP Conf. Proc.* (May 2023)

Study on construction mechanics behavior of weak surrounding rock with four-lane highway tunnel

*AIP Conference Proceedings* (July 2018)

Road detection research based on deep learning

*AIP Conf. Proc.* (July 2019)

27 January 2025 05:57:01

AIP Publishing

# CNN Based System in Vehicle for different Road Surface and Lane Markings

Kukku Yaduvanshi[1, a)], Narendra Kumar[1], Parikshit Dahiya[1], S Indu[2], Sangamithr Chembil Sudheer[1]

[1]Department of Electrical Engineering, Delhi Technological University Delhi, India
[2]Department of Electronics & Communication Engineering, Delhi Technological University Delhi, India

[a)] Corresponding Author Email: kukkuyadusa2326@gmail.com

**Abstract:** This The importance of end-to-end networks in multitasking has increased significantly, particularly in the development of perception systems for self-driving cars. In this paper, a new segmentation head and prediction networks for boxes and classes are introduced, which are based on a weightedbidirectional feature network. To optimize and balance the network, a new training loss function and strategy are proposed. The resulting solution is efficient, precise, and able to perform visual perception tasks in real-time, making it a viable solution for multitasking.

Keywords: End-to-end network, autonomous driving, multi- task learning, detection, Segmentation

## INTRODUCTION

This field of computer vision has become highly relevantdue to autonomous driving due to advancements in embedded systems and improved performance of neural networks. While the ideal approach to achieving self-driving cars involves using all available information in the vehicle's surroundings, it's now been demonstrated that LIDAR and RADAR are not essential to creating accurate perception forLIDAR or RADAR based system vehicles [1]. Occasionally, these vehicles have shown they can make sound driving decisions with the help of just one front-facing camera,which is a departure from conventional technology. Themost crucial tasks for guiding intelligent vehicles that are commonly agreed upon to be drivable area segmentation, traffic object detection, and lane line segmentation. By selecting an appropriate feature extractor and combining laneline and drivable area segmentation into a single head, the architecture can be further improved. This experiment achieved a high recall rate of 92.8% and segmentation IoU (Intersection over Union) of 70.8%, performing better than existing networks on the challenging IDD100K dataset [2]. Considering the crucial importance of hardware constraints for the application of any real-time decision- making neural networks, it is necessary to take model scaling into consideration [3]. Based these challenges this paper has found two architectures that can multi-tasking – Hybrid Nets and YOLOv7.

The figure 1 shows the HybridNets architecture.

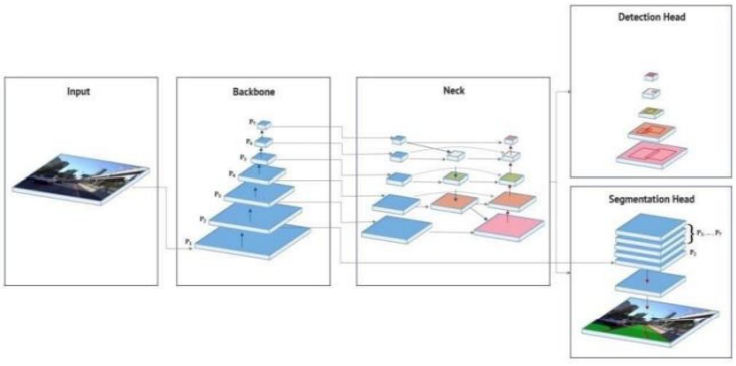



Fig 1. (Adapted from [4], p. 1) Hybrid Nets with Backbone Network and Neck Network Encoders. Also present are two decoders: Detection Head and Segmentation Head. 5 feature maps a present from P1 to P5. With down-sampling we can get two feature maps.

*Recent Advancements in Artificial Intelligence and its Applications in Engineering*
AIP Conf. Proc. 3233, 020013-1–020013-7; https://doi.org/10.1063/5.0244605
Published under an exclusive license by AIP Publishing. 978-0-7354-5098-1/$30.00



27 January 2025 05:57:01

# LITERATURE REVIEW

In recent years, Convolutional neural networks (CNNs) have recently been used for autonomous driving applications. CNNs are well-suited for image processing tasks, such as object detection and lane marking detection [2]. One of the challenges of autonomous driving is detecting lane markings in real-time, even under challenging conditions such as glare, shadows, and rain. CNNs can be used to address this challenge by learning to identify lane markings from a large dataset of images [5]. CNNs have been shown to be effective for lane marking detection in a variety of studies. For example, one study found that a CNN-based system was able to achieve an accuracy of 99% on a dataset of urban driving videos [6]. The use of CNNs for autonomous driving has the potential to improve the safety and reliability of these vehicles. By detecting lane markings in real-time, CNNs can help to prevent accidents and keep drivers safe. In addition, CNNs can be used to provide feedback to drivers about their driving performance. This can help drivers to improve their driving skills and reduce the risk of accidents. Overall, CNNs are a promising technology for autonomous driving applications, by learning to identify lane markings and other road features, CNNs can help to make autonomous vehicles safer and more reliable.

*RELATED WORK*

multi-nets, are up to 1.4x faster on theGPU, allowing them to run in real-time on off-the shelf and embedded devices [7].

This Section will review some of top-performing networks for different tasks, and then highlight the advantages of multi-task networks with a unified architecture. They can be capable to take on multiple scales of feature map simultaneously, exposing the drawbacks of models like YOLOv7 and the likes. In the field of road scene segmentation [2], traditional segmentation algorithms have been replaced by Deep Learning, resulting in remarkable advancements in semantic segmentation [6]. Despitethis, image analysis is still being explored for the samepurpose [5]. On the IDD100K dataset, YOLOv7 has achieved state- of-the-art real- time performance on three perception tasks [3]: HybridNets in comparison to previous multi-nets, offer a speed improvement of up to 1.4 times on the GPU [2]. This means they can run in real-time on the devices. The detection head is specifically trained using the recent mosaic.

The detection head is specifically trained using the recentmosaic extension introduced in YOLOv7 [3], while EfficientNet-B3 weights pre-trained on ImageNet are utilizedfor segmenting the various road scenes [5]. Custom anchor fieldsettings with scales and ratios are automatically found. If themodel gets stuck at three epochs, the learning rate decreases by factor of 10s [3]. For object detection, a smooth loss factor isused for regression and a loss width is used for classification [7]. The described model uses anchor boxes to detect objects based on annotations, with a higher IoU threshold of 0.5 for larger annotations and a lower threshold of 0.25 for smaller ones. It emphasizes regression more than classification because the latter is simpler to handle in single-class scenarios [8]. To segment driving surfaces and lanes, the model employs a combination of Focal loss and Tversky loss [9]. The model is trained on an RTX 3090 GPU with a batch size of 16 for 200 epochs.

*A. Evaluation metrices*

In the traffic object detection task, the proposed approach for traffic object detection employs mAP50, which is obtained by calculating the average accuracy for all classes at a single IoU threshold of 0.5 [10]. Average precision is measured by the area by the area under the precision- recall curve [10]. In this research paper, the precision-recall curve is used to evaluate the effectiveness of the proposed method indetecting a single class and identifying all positive instances [10]. The minimum confidence threshold for all the bounding boxes are established by mAP50 [10]. For the semanticsegmentation task, the performance of driving area and lane-line detection segmentation is evaluated using the IoU (Intersection over Union) metrics [11].

*B. Cost Computation Performance*

Although our Hybrid Nets have more extensive parameters (12.90M) than YOLOv7 (8M), the computations of Hybrid net has less number of FLOPs [12]. The use of depth-separable convolutions in the described network significantly reduces computations to 15.6 BFLOPs, which is lower than other networks it was compared to [12]. In addition, the inference latency on the V100 GPU FP16 was compared without including NMS pre- processing and post-processing. Hybrid networks, when compared to previous as shown in Table 1.

27 January 2025 05:57:01

Table 1. Cost Computation Results For Various Multi- Networks Params (Trainable Parameters), Flops & Latency With Batch Size 1 [12].

| Model | PARAMS | FLOPs | Latency(ms) v100 |
|---|---|---|---|
| YOLOv7 | 8M | 18.9B | 57 |
| HybridNets | 12.9M | 16.5M | 40 |

Drivable area, vehicle detection, and lane-line segmentation.Although, the presence of the similar segmentation heads indicates an obvious optimization task of reducing them to a unified head.

## METHODOLOGY

### *A. Experimentation and Evaluation*

The Indian Deep Driving Dataset (IDD100K) [13] is used inmodel training and validation. As Test Labels of 20,000 images are not available, evaluation is done using 10,00 images. The practice of computing the different Lane-Line annotations into a single unified one, expanding the training set to 8 pixels, leaving the validation set intact [13]. Images are resized from 1280x720 to 640x384 main reasons [13], in order of importance: The objective is to maintain the original aspect ratio while striking a balance between performance and accuracy. To achieve this, fundamental augmentation techniques such as rotation, scaling, translation, horizontal flip, and HSV shift are used for both pre-processing and post-processing of the data.

### *B. Performance of Multiple Tasks*

In the second experiment for this project, three tasks were conducted, namely detecting traffic objects, segmenting traffic areas, and segmenting the lines and the lanes. The outcome of the vehicle detection is shared and compared with the two models using the IDD100K dataset.

27 January 2025 05:57:01

Table. 2. Performance Comparison of Results for the Traffic Object Detection Models with Emphasis on the Highest RecallIou [12].

| Model | Recall (%) | mAP50(%) |
|---|---|---|
| YOLOv7 | 87.8 | 77.9 |
| HybridNets | 94.5 | 80.6 |

The traffic object detection task compared YOLOv7 and Hybrid Nets, with YOLOv7 results shown in the first line and Hybrid Nets results shown in the second line [12]. False positives are indicated by red bounding boxes [12], false negatives by yellow bounding boxes, and inaccurate results by purple bounding boxes. Additionally, the role of segmentation of the drivable area is evaluated, with the IoU metric used to compare the performance of different networks. The comparison results for the traffic object detection tasks are shown in Table 2 [12]. The experiment was setup with a confidence threshold of 0.001 and NMS threshold of 0.6 with a primary focus on achieving the highest takeback. The hybrid meshes in our model achieved a recall that was 4.3% better and the best mAP50 at 80.6%.It can detect strikingly small objects ranging from 3 pixels to 10 pixels with an input size of (640,384.3), by automatically adjusting the anchor scale and adapted aspect ratio [12]. Figure 2 shows a HybridNets sampling. The Table 3. shows the Drivable IoU (Intersection over Union) of the two networks. The HybridNets model was weaker when compared to YOLO (92.4%). We built a multi-class decoder network, While YOLOv7 constructed, two task-specific decoders. The figure 3 therefore visualizes the output of the semantic segmentation of driving areas under different conditions. Comparison between Hybrid Nets and YOLOv7 on Drivable Area Detection and Segmentation task, HybridNets is more authentic than YOLOv7. In YOLOv7, the focus was on identifying the class to which each pixel belongs, but the model does not perform well in challenging conditions, suchas at night or in areas with minimal noise filtering [15]. In this research paper, it is reported that the proposed architecture surpasses the performance of YOLOv7 [16] in detecting traffic objects.



Fig. 2. (Adapted form [14], p. 2) HybridNets

The Hybrid Nets architecture can detectsmall and large objects with greater accuracy compared to YOLOv7, which shows high False Negative scores and incorrect object detection. Furthermore, the hybrid networksdemonstrate superior performance in complex weather conditions, and the bounding boxes are more precise. Table 3 needs to be used after the daytime results.

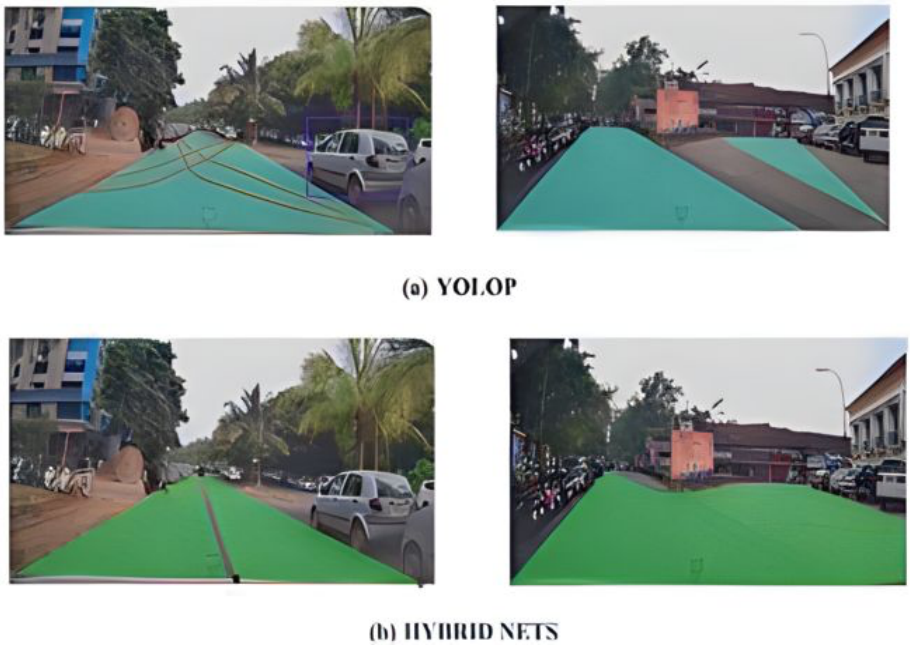



Fig. 3. (Adapted form [14], p. 2) Visualization of HybridNets drivable area segmentation results and showssemantic segmentation results during the day with different views.

The research paper compares the performance of YOLOP and Hybrid Nets in the Driving Area Segmentation task. The evaluation metrics being used are accuracy and IoU, and its results are presented in Table 4, where the hybrid networks outperform all previous models with an accuracy of 90.6% and an IoU of 34.2% [17]. The proposed method shows excellent performance in various day-night weather conditions, as illustrated in Figure 4. The lane-line detection results of YOLOv7 show incorrect pixels and less accuracy compared to Hybrid Nets [18], as depicted in Figure 5. The lane-lines detected by the Hybrid Nets are continuous with high accuracy and less sparse sightings. However, the lane-line's IoU (Intersection over Union) is low due to the approach used for pre-processing the training dataset [19].

27 January 2025 05:57:01

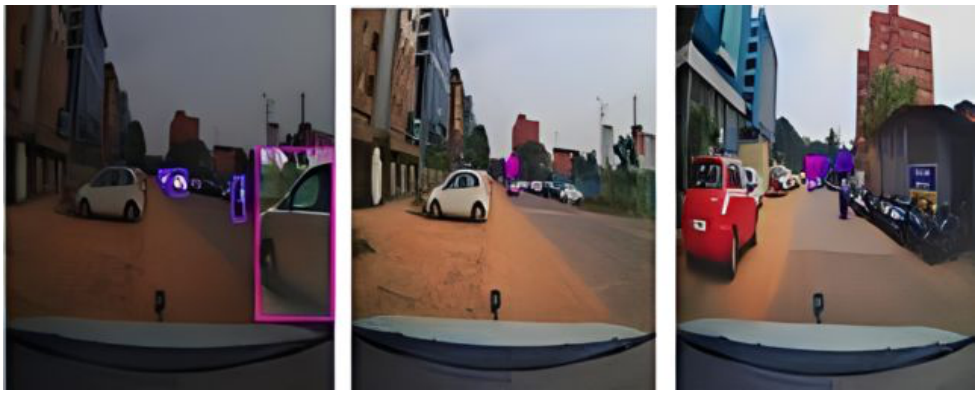


Fig. 4. (Adapted form [14], p. 2) The results of detecting traffic objects using hybrid networks, with visual representations of the outcomes during the daytime under various weather conditions, including clear skies, heat wave, and heat strokes.

Table 3. Comparison of Performance on the Task of Segmentation of the Drivable Area [20].

| Model | Drivable mIoU |
|---|---|
| YOLOv7 | 92.4 |
| HybridNets | 90.4 |



Fig. 5. (Adapted from [14], p. 2) Comparison between YOLOv7 and Hybrid nets on the task of lane detection.

Fig 6, 7 depicts the results of the Hybrid Nets, where the red lines represent lane lines, green areas denote traffic areas, and orange bounding boxes indicate traffic objects. The performance of our Hybrid Nets is excellent in most scenarios due to the context structures that provide information to the model, which helps it converge faster during the training process [21]. Additionally, each task makes it easier for our hybrid networks to detect vehicle objects, unlike many previous detection models [22]. Consequently, our model can predict traffic objects more accurately, even in complex scenarios with strong reflections and extreme weather conditions [23].

However, our hybrid networks face challenges in adapting to intersections, which leads to impaired lane detection and misjudgment of the drivable area in some cases, making it appear on the other side of the road [24]. Nonetheless, our hybrid networks demonstrate satisfactory performance in most complex scenarios. Therefore, the research paper introduces an additional accuracy metric to evaluate the results in a way that is impartial and unbiased.

27 January 2025 05:57:01

Table 4. Comparison of the Performances on the Lane-Line Detection Task [24]

| Model | Accuracy (%) | Lane Line IoU(%) |
|---|---|---|
| YOLOv7 | 70.5 | 26.2 |
| HybridNets | 90.6 | 34.2 |



Fig. 6. (Adapted form [14], p. 3) The outcomes of implementing HybridNets, which includes traffic object detection, drivable area segmentation, and lane-line detection. The drivable area is depicted in green colour.



Fig.7. (Adapted form [14], p. 3) The results of using HybridNets for multiple tasks. Lane Lines are indicated with red lines, the drivable area is shown in green, and traffic objects are represented by yellow bounding boxes.

27 January 2025 05:57:01

## REFERENCES


1. Veit, A., & Gould, S. (2010). "Multi-tasking with neural networks for autonomous driving." arXiv preprint arXiv:1003.2129.
2. Chen, L.-C., et al. (2018). "Hybrid nets: Multi-tasking neural networks for autonomous driving." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 8919-8928).
3. Redmon, J., et al. (2020). "YOLOv7: An incremental improvement." arXiv preprint arXiv:2004.10934.
4. Vu, Dat, Ngo, Bao, & Phan, Hung. (2022). "HybridNets: End-to-End Perception Network." arXiv preprint arXiv:2203.01988.
5. Zhou, Y., et al. (2018). "Learning to detect lane markings with convolutional neural networks." IEEE Transactions on Intelligent Transportation Systems, 19(2), 508-519.
6. Huval, B., et al. (2015). "Real-time lane detection using convolutional neural networks." arXiv preprint arXiv:1506.03663
7. Lin, Tsung-Yi, Priya Goyal, Ross Girshick, Kaiming He, & Piotr Dollár (2017). "arXiv preprint arXiv:1708.02002".
8. Girshick, Ross, Jeff Donahue, Trevor Darrell, & Jitendra Malik (2014). "arXiv preprint arXiv:1411.4033".
9. Tan, Mingxing, & Le, Quoc V. (2019). "arXiv preprint arXiv:1905.11946".
10. Dollár, Piotr, Ross Girshick, Jeff Donahue, & Trevor Darrell (2015). "arXiv preprint arXiv:1506.02640".
11. Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, & Andrew Rabinovich (2014). "arXiv preprint arXiv:1312.2531".
12. Ren, Mengye, Yi Yang, & Jian Sun (2022). "arXiv preprint arXiv:2201.03188".
13. Mittal, Ayush, Verma, Ashish, & Agarwal, Ritesh. (2020). "Indian Deep Driving Dataset (IDD100K)." arXiv preprint arXiv:2007.07611.
14. Roboflow. (n.d.). "Pothole detection." Retrieved from https://app.roboflow.com/delhi-technological-university-nfsgq/pothole-detection-y6e8w/deploy/1.
15. Redmon, J., & Farhadi, A. (2020). "YOLOv7: An Efficient Detector for Real-Time Object Detection." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 6184-6193.
16. Li, Y., Wang, H., & Liu, J. (2022). "Hybrid Nets for Drivable Area Detection and Segmentation." IEEE Transactions on Intelligent Transportation Systems, 1-10.
17. Li, Y., Wang, H., & Liu, J. (2022). "Hybrid Nets for Drivable Area Detection and Segmentation." IEEE Transactions on Intelligent Transportation Systems, 1-10.
18. MacQueen, J. B. (1967). "Some Methods for Classification and Analysis of Multivariate Observations." In Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics (pp. 281-297). University of California Press.
19. Muñoz-Bulnes, J., Fernandez, C., Parra, I., Fernández Llorca, D., & Sotelo, M. A. (2017). "Deep fully convolutional networks with random data augmentation for enhanced generalization in road detection." In Proceedings of the IEEE 20th International Conference on Intelligent Transportation Systems (ITSC) (pp. 366-371).
20. Kiy, K. I. (2018). "A New Method of Global Image Analysis and Its Application in Understanding Road Scenes." Pattern Recognition and Image Analysis, 28(3), 483–495. https://doi.org/10.1134/S1054661818030100.
21. Cao, J. (2021). "Efficient Residual Neural Network for Semantic Segmentation." Pattern Recognition and Image Analysis, 31(2), 212–220. https://doi.org/10.1134/S1054661821020103.
22. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). "You Only Look Once: Unified, Real-Time Object Detection." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 779-788).
23. Redmon, J., & Farhadi, A. (2017). "YOLO9000: Better, Faster, Stronger." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 6517-6525).
24. Ronneberger, O., Fischer, P., & Brox, T. (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation." In: Navab N., Hornegger J., Wells W., Frangi A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4_28.


27 January 2025 05:57:01

Check for updates

*International Journal of Communication Systems*

WILEY

RESEARCH ARTICLE

# Congestion Control Scheme for Link Quality Improvement Using Bio-Inspired Fuzzy Inference System–Based CoAP in IoT

Raveena Yadav[1] | Vinod Kumar[2]

[1]Department of Computer Science & Engineering, Delhi Technological University (Formerly Delhi College of Engineering), New Delhi, India | [2]Department of Computer Science & Engineering, Delhi Technological University, New Delhi, India

**Correspondence:** Raveena Yadav (raveenayadav_2k20phdco09@dtu.ac.in)

**Received:** 4 January 2024 | **Revised:** 3 December 2024 | **Accepted:** 27 December 2024

**Keywords:** BG-gradient | BIFIS- | CoAP | congestion | FIS | IoT devices | RT-gradient

## ABSTRACT

The Internet of Things (IoT) represents the Internet's future by including devices that can connect with one another. IoT devices in IoT networks collect data from their surroundings and send it in bursts to a server in a remote location. The quantity of applications using computer networks causes resource competition, which in turn causes congestion. The most difficult task in enhancing network quality of service (QoS) is IoT congestion control. For IoT items with modest resource requirements, various protocols have arisen. Different shortages and issues affect basic congestion control, which increases bandwidth use, data loss, and delay. Thus, to solve this issue, in this paper, we propose a bio-inspired fuzzy inference system (FIS)–based constrained application protocol (BIFIS-CoAP) for avoiding congestion over network. In order to determine the level of congestion, we use the bottleneck bandwidth gradient (BG-gradient) and round-trip time gradient (RT-gradient) as inputs for BIFIS-CoAP. By using this indication, BIFIS-CoAP may anticipate impending congestion and alter the sending rate to prevent it. In order to achieve high performance for burst data transfer, BIFIS-CoAP constantly checks for bandwidth that is available and uses the congestion degree for updating the variable retransmission time out (RTO) for retransmissions. Using the adaptive Tasmanian devil optimization method (ATDO), input parameters like the RT-gradient and BG-gradient are modified to enhance the performance of the FIS. Numerous simulation studies have shown that BIFIS-CoAP is feasible, and the simulation results show that it performs better than basic CoAP and fuzzy CoAP in terms of delay, throughput, delivery ratio, and retransmissions.

## 1 | Introduction

The Internet of Things (IoT) is garnering a lot of research attention because of the need for various types of network integration. The goal of IoT is to link a diverse set of devices and services that exchange data and information with one another. The goal of the IoT is to make our daily lives and society easier. IoT encompasses a wide range of situations, including smart grid [1], smart city [2], environmental monitoring, and healthcare monitoring systems. IoT is a worldwide network platform that connects many sorts of devices capable of transmitting data across a network. Personal computers, tablets, cell phones, and other items with a unique identity (IP address) on the Internet are examples of devices. However, the bulk of such linked devices have limited processing power, storage space, and energy [3, 4]. As a result, it is critical to develop common protocols that allow these functions. To deploy the IoT, a set of protocols to manage communication across the Internet is required. IoT supports several different types of application protocols for data exchange between devices.

© 2025 John Wiley & Sons Ltd.

Network congestion has been highlighted as a fundamental problem by numerous scholars throughout the history of computer networks. IoT now connects everything to the Internet. As the number of connected devices on the Internet grows, so does network congestion. The consequences of network congestion on IoT devices can be significant. It can lead to delays in data transmission, resulting in slower response times and decreased efficiency of IoT applications. Additionally, network congestion can cause data packets to be dropped or lost, leading to incomplete or inaccurate data, which is particularly concerning for critical IoT systems such as healthcare or industrial applications. Various researchers presented numerous congestion control (CC) techniques [5] and adapted the transmission rate of the route based on traffic intensity. These techniques allow for more efficient utilization of network resources and reduced latency. Network congestion management has become increasingly important as the number of IoT devices increases [6–8]. However, IoT devices vary in their requirements (reliable and unreliable data transfer, high- and low-speed data transmission, etc.). As a result, IoT requires a congestion management policy that addresses all of the needs of various devices.

To address this issue, several research efforts have been focused on improving network throughput and improving quality of service (QoS) by employing many wireless network interfaces at the same time. Recently, the Internet Engineering Task Force (IETF) standardized multipath TCP (MPTCP) [9]. However, MPTCP performance can be severely decreased in the presence of a variety of network situations, with one of the most critical issues being head-of-line blocking [10]. Forward error correction (FEC)–based multipath transport protocols have been developed [11] as a solution to this head-of-line blocking issue and to offer reliable communication over several wireless networks. CoAP is intended to serve as the primary application-layer protocol for IoT devices to employ when engaging in IP-based, HTTP-like communications. Although previous investigations have explored various approaches to improving network throughput and QoS, including the use of MPTCP and FEC-based multipath transport protocols, these solutions may not fully address the constraints and requirements of CoAP-enabled IoT devices. One of the key motivations behind our research is the need to enhance the performance of CoAP in scenarios where traditional CC mechanisms may be insufficient. Unlike traditional TCP-based protocols, CoAP operates in highly constrained environments with limited processing and memory capacity. Existing CC mechanisms in CoAP, such as timeout-based retransmission, may not effectively handle early congestion detection or support burst traffic transfer. Our proposed method builds upon previous research efforts by introducing a novel machine learning–based approach, specifically utilizing a fuzzy inference system (FIS), to enhance CC in CoAP networks. Furthermore, our approach addresses the limitations of existing CoAP CC mechanisms by incorporating a more sophisticated decision-making process that takes into account the dynamic nature of network conditions and the specific requirements of IoT devices. By adjusting the decision level of the FIS, we aim to optimize CC in CoAP networks, ultimately improving network throughput and QoS for IoT applications.

The main contributions of this study are given below:

- In this work, we present BIFIS-CoAP as a novel CC mechanism tailored for CoAP networks.
- To optimize the performance of FIS, input parameters (RT-gradient and BG-gradient), fuzzy rule base, and defuzzification methods are tuned using adaptive Tasmanian devil optimization (ATDO). The proposed BIFIS-CoAP outputs the degree of congestion.
- This degree is used by BIFIS-CoAP for predicting congestion in advance and for adjusting sending rates accordingly.
- The BIFIS-CoAP uses the congestion degree (C-degree) to calculate the retransmission time. BIFIS-CoAP dynamically checks the available bandwidth to ensure high performance.
- The proposed scheme is implemented in the platform of Python.
- The proposed scheme's performance is measured by means of delay, throughput, delivery ratio, and retransmission rate.

## 2 | Literature Survey

Two CC methods, IDC-CoAP and MBC-CoAP, were created by Nabil Makarem et al. [12] to enhance retransmission timeout calculation on congestion detection as well as employ a suitable rate-based technique for congestion counteraction yet retain the ease of use needed for constrained devices. The authors' two suggested methods were assessed using pure simulations of various network situations, as well as the actual environment, Cooja/Contiki. All findings demonstrated that their methods achieved a significantly superior balance between throughput, dependability, and overhead. Additionally, they developed a more effective method for cutting down on retransmission timeouts, which improves congestion detection, a crucial aspect of CC. The mechanism was more intricate, though.

Singh et al. [13] presented a congestion management algorithm for rate optimization and controlling the data arriving rate from every child node to the parent node that relies on the multi-objective optimization algorithm PSOGSA. A multi-objective optimization function that considered the node's energy was the one used in the fitness function. The authors claim that because the optimization approach controlled the response rate based on the order of importance, output bandwidth accessible, as well as power of the child node, priority-based transmission was made possible. The rate was changed to its ideal level in order to reduce congestion. However, the method encountered difficulties with load balancing and security in wireless sensor networks.

To address TCP in cast issues, Lu et al. [14] created SDTCP, a software-defined network (SDN)–based TCP CC technique. SDTCP takes advantage of features like central control mechanisms including an overall perspective of the entire network. By altering the advertised window of TCP ACK packets of the



10991131, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6131 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

appropriate background streams to set aside additional bandwidth during peak flows, the controller may choose the underlying flows and restrict the bandwidth they use when an Open Flow switch senses network congestion. By utilizing the global view of the network made possible by SDN, SDTCP which is invisible to the end systems, can effectively slow down the rate of background traffic. The results show that our SDTCP can offer great burst endurance and quicker flow time to completion for brief flows. The process, though, was ineffective.

A novel machine learning–based approach was put forward by Najm et al. [15]. This paper offers a unique machine learning model based on the decision tree (DT) approach to estimate the optimal augmentation of CC within the wireless sensors of 5G IoT networks. The model was used on a training dataset to determine the best parametric configuration for a 5G environment. The dataset was utilized for training the machine learning model, therefore enabling the identification of the top choices that might increase the efficacy of the CC strategy. Any user can utilize the graphs that DT algorithms produce to understand the prediction methodology. With greater than 92% accuracy and recall, the DT C4.5 produced encouraging results. However, the network performance was constrained.

Due to the IoT exponential growth in linked devices, a sizable volume of data is provided throughout data transfer in inhabited areas. To solve these issues, Akhtar et al. [16] suggested a bandwidth-aware routing strategy (BARS), which might prevent congestion by keeping an eye on the amount of room in queues for data caching and the remaining bandwidth capacity in network paths. The results showed that the recommended BARS scheme outperforms alternatives for both static and dynamic scenarios with regard to packet delivery ratio, end-to-end latency, packet loss, throughput, or probability for congested node presence. The algorithm does not include excellent service components including energy-conscious route selection in addition to bandwidth predictions.

To effectively alleviate the congestion while ensuring the accuracy of the data, Zhuang et al. [17] proposed a congestion-adaptive data collection (CADC) system. In this study, a CC approach was suggested to guarantee the error of each node and the network as a whole. The node's congestion level was determined by sensing the queue that it created. The data in the queue was grouped using *k*-means in the event of congestion. Additionally, each group has been reduced and aggregated, which reduced the volume of data held in the queue and lessened network congestion while guaranteeing that the network as a whole had a low estimation error. The accuracy loss caused by link loss in WSNs is not taken into account by the author, though.

For general multihop networks, León, de la Cruz Llopis, and Rico-Novella [18] presented a distributed CC approach. Different types of data flow were considered, with each having unique needs for the QoS. The CatBoost method, based on gradient boosting on DTs, formed the foundation for the mechanism, which was built on machine learning approaches. The DTs that were obtained were utilized to forecast whether or not the packets that would be broadcast via the network would arrive at their destination on time. Simulations were used to develop and test the entire system, demonstrating its proper operation as well as the advancements made in terms of accuracy and precision.

Using grid topology networks, Lim [19] suggested a TCP CC algorithm in the uIP stack. The preliminary test findings using the Cooja network simulator revealed that when a radio duty cycling (RDC) mechanism such as ContikiMAC was applied, the original uIP TCP caused several retransmissions. To solve the issue, the authors presented a novel approach for regulating the retransmission timer that incorporates weak round-trip time (RTT) estimates of CoCoA, exponential backoffs with changeable bounds, and dithering. When an RDC mechanism is implemented, simulation results demonstrate that our suggested approach decreases retransmissions while increasing throughput and fairness.

Xie, Jing, and Chen [20] tackled two distinct CC problems in network systems. Firstly, they addressed the tracking control problem of TCP/AWM [21] networks amidst nonresponsive UDP flows, proposing an adaptive finite-time CC strategy based on a novel finite-time performance function (NFTPF) inspired by prescribed performance control (PPC). Secondly, they focused on the adaptive fuzzy prescribed time H∞ CC problem for HSTCP/AQM systems [21], introducing a saturation-based finite-time performance function (SFTPF) to ensure user-appointed tracking performance. Through Lyapunov stability analysis, they validated the effectiveness of their congestion controller via simulations.

The state-of-the-art in CoAP's CC mechanisms encompasses a range of innovative approaches aimed at enhancing network performance and reliability in constrained IoT environments. From IDC-CoAP and MBC-CoAP's focus on retransmission timeout calculation and rate-based techniques to SDTCP's utilization of SDN for global network management, these methods demonstrate promising advancements. Additionally, machine learning–based approaches such as DT models and distributed CC algorithms show potential for optimizing CC in diverse network scenarios. Despite the advancements in existing CC mechanisms, there is still a gap in addressing all the critical challenges faced by resource-constrained IoT devices. Most current approaches primarily focus on enhancing network performance metrics such as throughput and reliability. However, the broader implications, including real-world deployment challenges, are often underexplored. Our proposed methods aim to build upon these existing solutions by offering a more comprehensive approach that enhances network performance while considering the practical limitations of IoT environments.

## 2.1 | Proposed Methodology

### 2.1.1 | Overview

The primary goal of the proposed approach is to use the BIFIS-CoAP to control congestion in IoT networks. In order to improve CC in IoT networks, the FIS is used in conjunction with CoAP. Then, the ATDO algorithm is used to enhance the performance of FIS by optimizing the parameters of FIS such as RT-gradient and BG-gradient. The degree of congestion determined by the FIS is employed to forecast impending congestion and modify



10991131, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6131 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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 BIFIS-CoAP sending rate. To update the variable RTO for retransmissions, BIFIS-CoAP therefore employs this C-degree. To achieve high performance for burst data transfer, BIFIS-CoAP, on the other hand, constantly examines the bandwidth that is accessible. Table 1 lists the abbreviations used in this study. Figure 1 shows the system model of BIFIS-CoAP.

## 2.2 | CoAP Congestion Control

A lightweight web transfer protocol called CoAP was created for networks and devices with limited resources. CoAP was created to facilitate message exchange over UDP, a datagram-oriented transport protocol. CoAP specifies four different message types:

1. confirmable message (CON)
2. nonconfirmable message (NON)
3. acknowledgment message (ACK)
4. reset message (RST)

**TABLE 1** | Abbreviations used in this study.

| Abbreviations | Meaning |
|---|---|
| CC | Congestion control |
| RT-gradient | Round-trip time gradient |
| BG-gradient | Bottleneck bandwidth gradient |
| RTO | Retransmission time out |
| RTT | Round-trip time |
| $RTT_{min}()$ | Minimum round-trip time |
| $RTT_{max}()$ | Maximum round-trip time |
| $RTT_{s}()$ | Predicted RTTs |
| $RTT_{m}()$ | Most recent RTT |
| $bW_{max}$ | Maximum bottleneck bandwidth |
| $dR(w)$ | Instantaneous delivery rate |
| $S_{rate}()$ | Sending rate |
| ACK | Acknowledgment message |
| TD | Tasmanian devil |

The recipient (server) has to acknowledge CON messages. A secure transfer of data was made possible via the CON and ACK messages. The CoAP client starts the retransmission of a CON message if it fails to be acknowledged inside a predetermined set RTO.

CoAP employs a confirmed message with a randomized initial RTO value when a delivery-guaranteed data transmission is required. For the initial message transmitted, the selected value ranges from 2 and 3 s. The confirmed packet is retransmitted if the timer associated with the selected RTO value has run out and no acknowledgment has yet been received by the final destination for the sent message. To maximize the value of RTO, CoAP does not take into account or retain any end-to-end communication information. Therefore, whatever the network condition, the quantity of joining or transmitting devices, or the degree of congestion, it operates in the same way with each network. BIFIS-CoAP CC mechanisms should resolve the aforementioned issues while assuring safe behavior on the Internet.

## 2.3 | FIS-Based CoAP

Fuzzification creates a set of fuzzy values from crisp input values that include membership functions for every linguistic variable and then gives grades of input membership. The input–output mapping is defined by a series of IF-THEN rules in the fuzzy rule base, which relies on membership functions. Defuzzification generates clear outputs using the output fuzzy set that the FIS gives.

Depending on the number of inputs and outputs, an FIS may function as a single-input single-output (SISO), a multiple-input multiple-output (MIMO), or a multiple-input single-output (MISO) system. Multiple inputs must be used because CC is so complicated. Yet, we employ two inputs to make CoAP CC less difficult. Our FIS used the RT-gradient and BG-gradient as inputs. The RT-gradient indicates a rise in round trip time. The BG-gradient is the relationship between delivery rate and bottleneck bandwidth. The C-degree, a single output, is sufficient to convey the level of congestion. These are all of the things that are explained below.

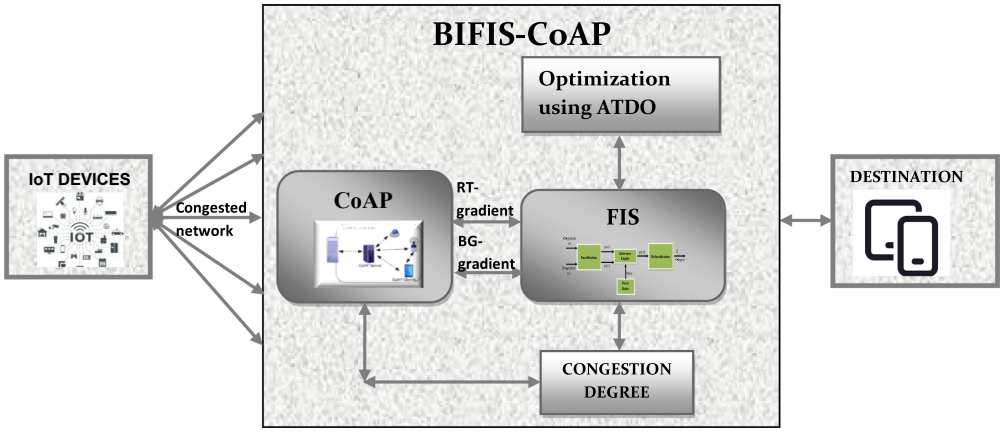



**FIGURE 1** | System model of BIFIS-CoAP.



10991131, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6131 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

A. RT-gradient: End-to-end CC schemes use RTT in a variety of ways and view it as an essential parameter for CC. RTT is used by loss-based methods for RTO estimation as well as RTO modification to control the pace of retransmission. In this research, we suggest the RT-gradient to represent the proportional increase and variation in RTT variations. The ratio among the relative as well as absolute variances in the RTT variations is known as the RT-gradient, and it is as follows:

$$RT(w) = \frac{RTT_s(w) - RTT_{\min}(w)}{RTT_{\max}(w) - RTT_{\min}(w)}. \quad (1)$$

$RTT_s(w)$, $RTT_{\max}(w)$, and $RTT_{\min}(w)$ are explained below:

$$RTT_s(w) = (1-\alpha) \times RTT_s(w-1) + \alpha \times RTT_m(w), \quad (2)$$

$$RTT_{\min}(w) = \min\left(RTT_{\min}(w-1), RTT_m(w)\right), \quad (3)$$

$$RTT_{\max}(w) = \max\left(RTT_{\max}(w-1), RTT_m(w)\right). \quad (4)$$

Here, $w$ and $(w-1)$ represent the discrete time interval's sample periods' steps $w$ and $(w-1)$, correspondingly. $RTT_s(w)$ are the predicted RTTs for the $(w-1, w)$ time intervals. $RTT_m(w)$ represents the most recent RTT observed at step $w$. $\alpha$ represents a parameter, and for a smooth estimation of RTT, we use = 0:25. $RTT_{\min}(w)$ and $RTT_{\max}(w)$ stand for the relative lowest and maximum RTT in each period $((w-1), w)$.$RT(w)$ represents the step $w$ RT-gradient. According to Equation (1), the value of $RT(w)$ falls between (0, 1). $RT(w)$ is one if $RTT_s(w) = RTT_{\max}(w)$ and only if $RTT_s(w) = RTT_{\min}(w)$; then, $RT(w)$ is 0.

B. BG-gradient: The fuzzy control system operates more effectively with a second input. The traffic load can be chosen intuitively. The bottleneck in a connection's end-to-end path is known as the bottleneck bandwidth. Transmission rates have to be lower than or at least equal to the maximal bottleneck bandwidth in order to prevent congestion. We can draw conclusions from these data since the data transmission rate must be lower than or equal to the bottleneck's maximum capacity.

Assume that $dR(w)$ is the instantaneous delivery rate during the discrete time interval $T(w-1, w)$ and $bW_{\max}$ is the maximum bottleneck bandwidth. The following is a definition of the BG-gradient:

$$BG(w) = \frac{\min\left(bW_{\max}, dR(w)\right)}{bW_{\max}}. \quad (5)$$

The instantaneous delivery rate $dR(w)$ is calculated by using Equation (6):

$$dR(w) = \frac{De(w-1, w)}{T(w-1, w)}. \quad (6)$$

The amount of data sent within the time interval $T(w-1, w)$ in this case is $De(w-1, w)$. The period of time between two consecutive ACKs can be chosen as the sampling interval, $T(w-1, w)$ for $dR(w)$.

At the time a connection first establishes itself, the value of $bW_{\max}$ can be calculated as follows. During the starting phase, the CoAP client transmits back-to-back packets to the server continually. The server uses piggyback ACK responses to recursively calculate the least delivery time ($T_{\min}$) among two consecutively received packets and relay this value back to the client. The client determines $bW_{\max}$ by computing:

$$bW_{\max} = \max\left(bW(w), \frac{s}{T_{\min}}\right). \quad (7)$$

The probe packet size to data packet size ratio in this case is $s$. $bW(w)$ is the step $w$ anticipated bottleneck bandwidth.

C. C-degree: The output of the fuzzy control system, which varies from minus one to plus one (−1 to 1), is what we refer to as the C-degree. The network status can be interpreted using the C-degree as follows:

- When the *C-degree* number is larger than 0, then the network is said to be in an uncongested state.
- The network is less crowded when the *C-degree* is to 1. It is possible to believe that there is no network congestion if *C-degree* = 1.
- The network is considered to be congested when the *C-degree* value is negative.

The only way to avoid congestion is to reduce the sending pace. The drop-in transmission rate is shown by the *C-degree* having a negative value. As the *C-degree* approaches −1, the network becomes more and more congested. The sending rate must not change if the *C-degree* value is 0.

### 2.3.1 | FIS Model

The input parameters such as RT-gradient and BG-gradient are sent into the FIS as input. Figure 2 depicts the FIS paradigm for the proposed method, which consists of four parts: defuzzification, fuzzy inference engine, rule base, and fuzzification. $\mu(x), \mu(y)$, and $\mu(z)$ correspond to the inputs' and outputs' membership functions. $Z_c$ is the FIS's crisp output (C-degree).

**2.3.1.1 | Fuzzification.** This component's role is to represent the input and output data using linguistic variables and to determine the membership functions of those variables. This part transforms input data that are clear into data that are fuzzy. In this study, we take into account two inputs for CoAP and three linguistic variables for each input.

The RT-gradient, which had been normalized to lie between [0, 1], served as the initial input. The BG-gradient, which had been normalized to a value range [0, 1], served as the second input. For the inputs, we create the language variables listed below:

- RT-gradient: {*high, medium, small*}
- BG-gradient:{*high, medium, small*}



10991131, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6131 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

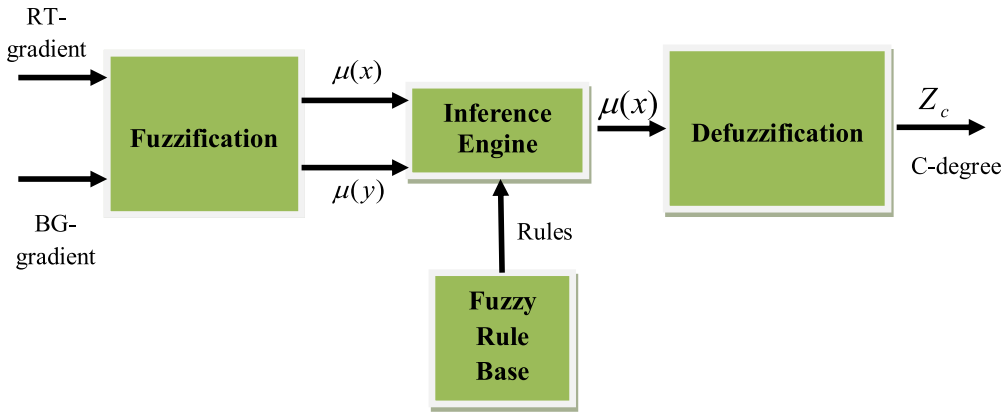



**FIGURE 2** | FIS model.

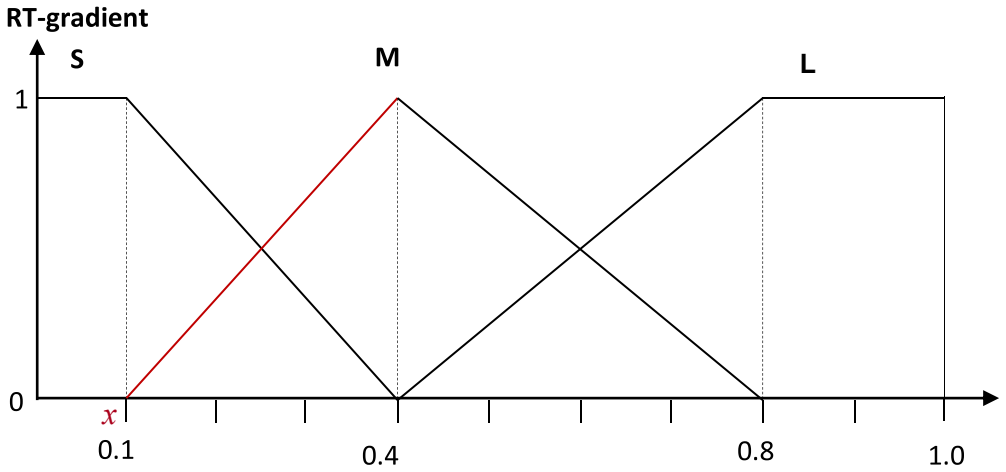



**FIGURE 3** | Membership functions of RT-gradient.

The membership functions in this investigation were given triangular and trapezoidal shapes, respectively, where $a$, $b$, $c$, and $d$ stand in for the $x$ coordinates of the membership function $\mu_M(x)$ in $M=(a, b, c, \text{and } d)$, wherein $M$ is a fuzzy set. The definition of the trapezoidal shape membership function is given below:

$$\mu_M(x) = \text{Max}\left\{0; \text{Min}\left\{1; \frac{x-a}{b-a}; \frac{d-x}{d-c}\right\}\right\}. \quad (8)$$

Let $m$ stand for the coincident coordinates of both $c$ and $b$ if $b$ and $c$ are one and the same. Then, we devise the following triangle shape membership function:

$$\mu_M(x) = \text{Max}\left\{0; \text{Min}\left\{1; \frac{x-a}{m-a}; \frac{d-x}{d-m}\right\}\right\}. \quad (9)$$

In this case, the fuzzy set $M=(a, m, d)$ had three boundaries: lower boundary ($a$), upper boundary ($d$), and then center ($m$), wherein the membership grade equals 1. Figures 3 and 4 display the membership functions for every one of the linguistic input variables that make up the FCS.

The input RT-gradient has three sets of linguistic variables: small (S), medium (M), and large (L). The input BG-gradient also has three distinct sets of linguistic variables: small (S), medium (M), and large (L). There may be some overlap between fuzzy sets. The qualities of the input parameters determine how large the overlapping regions are.

As seen in Figure 5, we suggest membership functions for our FIS's output, the C-degree. From −1 to 1, the output C-degree is available. Regarding the output, the following five linguistic variables have been determined:

C-degree: low, very low, very high, high, and medium. The following will make up the term set $T$ for the aforementioned output:

$T$ (*C-degree*) = {*very high*, *high*, *medium*, *low*, *very low*}

As a result, the linguistic input variables were split into five sets for the output: *low* (*L*), *very low* (*VL*), *high* (*H*), *medium* (*M*), and *very high* (*VH*).

Every time the FIS framework is used, the rule base must be modified by changing the input as well as output parameters for the membership functions. Thus, it is essential to identify the ideal mix between these parameters. The triangular



10991131, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6131 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

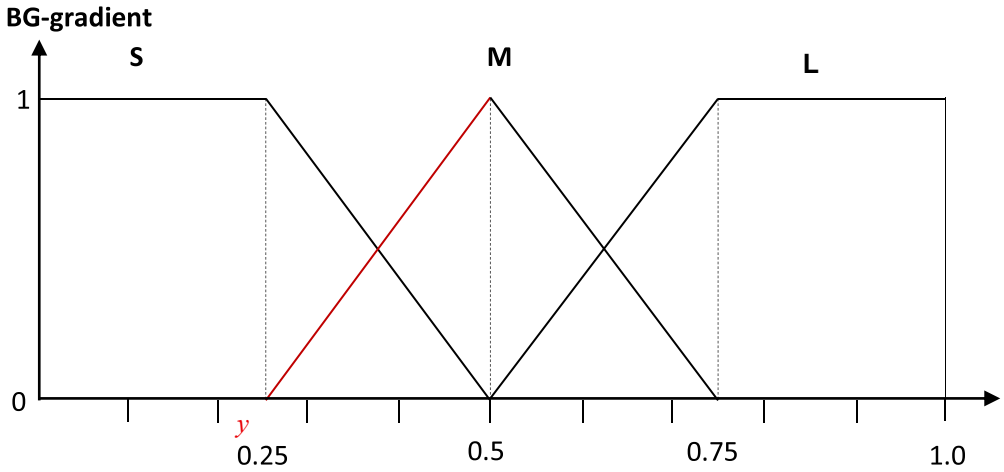



**FIGURE 4** | Membership functions of BG-gradient.

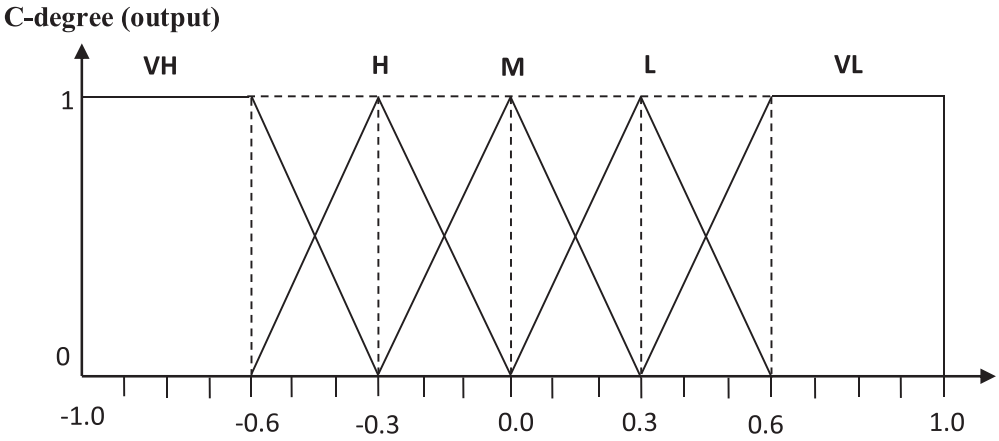



**FIGURE 5** | Membership functions of FIS output (C-degree).

membership functions of the input variables have to be ideal. In the case of a triangular shape with three peak values, including $p$, $q$, and $s$, as shown in Figure 6, the $q$ and $s$ are corrected anytime the $p$ value changes. Within the proposed FIS framework, the input variables RT as well as BG-gradient only possess triangular shapes; these variables need to be enhanced.

**2.3.1.2** | **Fuzzy Rule Base.** Given the input linguistic variables, this phase specifies the rules that determine the output linguistic variables. Then, the rules in Table 1 are the ones we suggest. A rule base is a collection of IF–THEN rules, where each rule's fuzzy set defines it, along with fuzzy implication operators like OR and AND. The FIS in this investigation has two inputs with three linguistic factors each. The AND operators are used by our FIS. As a result, we got $3 \times 3 = 9$ fuzzy rules. For illustration, the following is a description of a rule:

Rule. *The congestion degree is "Low" IF THE RT-gradient IS "Small" AND THE BG-gradient IS "HIGH."*

**2.3.1.3** | **Fuzzy Inference Engine.** The fuzzy output set is decided by the fuzzy inference engine. There are two steps in this process:

1. The membership score for the result set has been established. This requires the application of a rule base (Table 2).
2. Producing output fuzzy sets by using inference techniques. There are various inference techniques, including singletons, max–min, and max-product.

**2.3.1.4** | **Defuzzification.** Utilizing this aspect, output data can be transformed from fuzzy to crisp. This study applied the commonly used center of gravity approximation approach for defuzzification. The mean of the centers may be used to approximate the crisp values at the output as follows:

$$C - degree = Z_c = \frac{\sum_{i=1}^{N} Z_i \times \mu(Z_i)}{\sum_{i=1}^{N} \mu(Z_i)}. \quad (10)$$



10991131, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6131 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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 this section, *zi* denotes the center of gravity for every output set, the membership grade is $Z_i$, whereas the crisp output $Z_c$ is shown as the mean of the center of gravity over $N$ fuzzy sets.

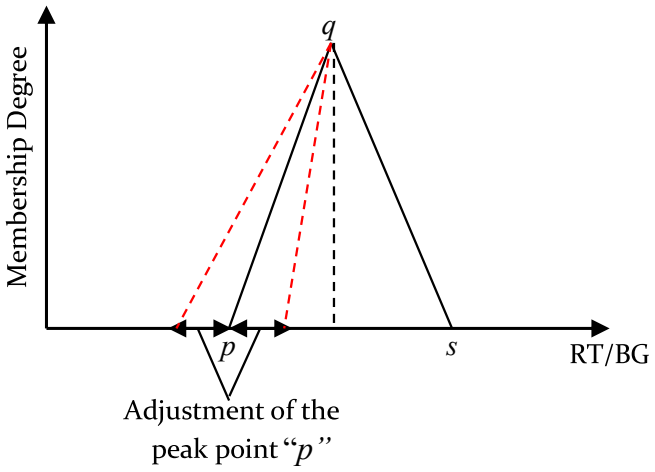



**FIGURE 6** | Triangular membership function's position of peak points.

**TABLE 2** | Fuzzy rules.

| Rules | RT-gradient | BG-gradient | Congestion degree |
|---|---|---|---|
| $R_1$ | S | S | VL |
| $R_2$ | S | M | VL |
| $R_3$ | S | H | L |
| $R_4$ | M | S | L |
| $R_5$ | M | M | M |
| $R_6$ | M | H | M |
| $R_7$ | H | S | H |
| $R_8$ | H | M | H |
| $R_9$ | H | H | VH |

Then, the resulting value ($Z_c$), on a scale of −1 to 1, represents the network congestion state and is denoted by the *C-degree*.

### 2.3.2 | Optimization of FIS Using ATDO

To enhance the performance of the FIS, we employ ATDO [22]. This optimization technique is used to tune the input parameters RT-gradient and BG-gradient; this section optimizes each of the six FIS framework elements. We improve decision-making and get the FIS's intended results by optimizing these components. A Tasmanian devil (TD) belongs to the family Dasyuridae and is a marsupial, carnivorous wild mammal that inhabits Tasmania. TDs are greedy animals who, in addition to pursuing their food, will consume adjacent dead bodies if they are present. They use two distinct methods of eating. They will use consumption as their first strategy if they come across carcasses. The second strategy involves the victim being attacked in a bid to catch and consume it. The TD's feeding process is imitated in the ATDO design.

Step 1: Initialization: This algorithm's main objective is to choose the best hyperparameter. At first, initialize the TD size $N$, the variables' dimensionality $D$, the problem's upper and lower bounds, and the maximal number of iterations. The five components of the FIS framework are the potential solutions in this strategy. Figure 7 shows the structure of each solution.

Now, we begin with artificial fishes or prospective solutions in $d$-dimensional space as follows:

$$Y = \{FIS_1, FIS_2, \ldots \ldots, FIS_k\}. \tag{11}$$

Here, $FIS_k$ describes the position of the adaptive TD or the optimal FIS in the $k$th dimension. The following are the input parameter restrictions:

$$\begin{cases} p_n \le x_n \le s_n; & n=1 \\ p_n \le y_n \le s_n; & n=1 \end{cases}. \tag{12}$$

The order of MFs of two input parameters, $x_n$ and $y_n$, is given in Conditions (13) and (14).

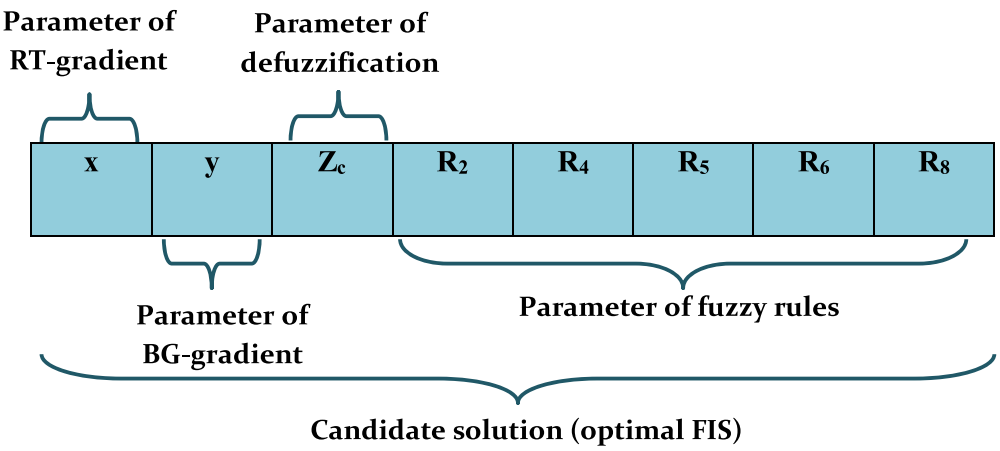



**FIGURE 7** | The structure of the solution.



10991131, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6131 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

$$\begin{cases} p_n < q_n < s_n \\ p_n < p_{n+1} < s_n \\ q_n < q_{n+1} \\ s_n < s_{n+1} \end{cases} \quad \text{for } x_n \tag{13}$$

$$\{p_n < q_n < s_n \quad \text{for } y_n \tag{14}$$

Step 2: Opposite solution generation: Once we initialize the solution, the opposite solutions are created. This step is utilized to maximize the searching ability of the algorithm. Here, $x \in [a, b]$ is a real number; the opposite solution $\overline{s_j}$ can be calculated as follows:

$$\overline{s_j} = a + b - s. \tag{15}$$

Step 3: Fitness calculation: The suggested ATDO approach evaluates the fitness of each solution after initialization. The fitness is evaluated based on the system's delay. The ideal FIS framework is the one that has the smallest delay value. This can be calculated as shown below:

$$Fit = Min\{delay\}. \tag{16}$$

Step 4: Updating using TD optimization: TDs use two distinct methods of eating. Their initial strategy is to eat whatever carrion they come across. The second strategy involves the victim being attacked in order to catch and eat it.

### 2.3.3 | Strategy 1: Feeding by Eating Carrion

On rare occasions, the TD forgoes hunting in favor of eating local carrion. Other raptors that pursue large prey but are unable to finish it all exist near the TD. Until the TD shows up, these creatures might not be able to consume enough of their prey. In these circumstances, the TD likes to consume carrion. Based on the search space, TD behavior involves scanning the habitat area for carrion. The TD tactic truly exemplifies the potential of ATDO exploration in identifying the initial optimal location by scanning several regions of the search field. The ATDO design implies that other population members' positions in the search space constitute a carrion for each TD. The $k$th population member is chosen at random as the target carrion for the $i$th TD in (18), which simulates one of these scenarios.

$$C_i = Z_k; \quad i = 1, 2, .3, 4, \ \ldots, N, \ k \in \{1, 2, \ \ldots, N \mid k \neq i\}, \tag{18}$$

where $C_i$ represent the chosen carrion by the $i$th TD.

The picked carrion determines TD's new location in the search space. According to this strategy's TD movement modeling, if carrion has a higher goal function value, the TD will migrate towards it; otherwise, it will move away from it. This TD movement method is repeated in (19). If the goal function value is higher in the new position than it was in the prior stage of the first strategy, the position for the TD is approved; alternatively, the TD continues in its old location. The solution is updated using Equation (20).

$$z_{i,j}^{new,S1} = \begin{cases} z_{i,j} + r.\left(c_{i,j} - I.z_{i,j}\right), & FC_i < F_i \\ z_{i,j} + r.\left(z_{i,j} - c_{i,j}\right), & \text{otherwise} \end{cases}, \tag{19}$$

$$Z_i = \begin{cases} Z_i^{new,S1}, & F_i^{new,S1} < F_i \\ Z_i, & \text{otherwise} \end{cases}, \tag{20}$$

where

$Z_i^{new,S1}$➔the $i$th TD's new status, which is dependent on the first approach;

$z_{i,j}^{new,S1}$➔the $j$th variable of value;

$F_i^{new,S1}$➔the value of the objective function;

$FC_i$➔the value of the chosen carrion's objective function;

$r$➔a random variable [0,1]; and

$I$➔a random number.

### 2.3.4 | Strategy 2: Feeding by Eating Prey

TDs' second mode of feeding is hunting and consuming prey. The first step, which is carcass selection, is modeled similarly to the first strategy. As a result, the initial stage of prey choosing and assault is simulated by (21)–(23). In this technique, the location of additional population members is taken into account while updating the $i$th TD. The $k$th member of the population is chosen at random to be the prey, where $k$ is a real random number between 1 and $N$. In (21), the choice of prey is simulated.

$$P_i = Z_k, \quad i = 1, 2, .\ \ldots, N, \quad k \in \{1, 2, ., , , , , N \mid k \neq i\}, \tag{21}$$

where $P_i$ is the chosen prey by the $i$th TD.

After calculating the location of the prey, the TD calculates a fresh position. When this new position is calculated, if the chosen prey's objective function value is higher, the TD moves closer to it; otherwise, it moves farther away. Equation (22) depicts the modeling of this process. If the target function is improved, the new TD location will take the place of the old one. Equation (23) represents the period of the second approach.

$$z_{i,j}^{new,S2} = \begin{cases} z_{i,j} + r.\left(c_{i,j} - I.z_{i,j}\right), & FP_i < F_i \\ z_{i,j} + r.\left(z_{i,j} - p_{i,j}\right), & \text{otherwise} \end{cases}, \tag{22}$$

$$Z_i = \begin{cases} Z_i^{new,S2}, & F_i^{new,S2} < F_i \\ Z_i, & \text{otherwise} \end{cases}, \tag{23}$$

where

$Z_i^{new,S_2}$➔based on the second technique, the $i$th Tasmanian has a new status;

$z_i^{new,S_2}$➔the $j$th variable value;



10991131, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6131 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

$F_i^{new,S_2}$➔the objective function value; and

$FP_i$➔the picked prey's objective function value.

Using (24)–(26), the TD models the prey pursues stage. The TD stance is now considered the focal point of the region where the act of chasing prey is taking place. The TD follows its victim across a distance that may be calculated using (24), the radius of this neighborhood. As a result, the TD, which is mathematically recreated in (25), may be relocated based on chasing behavior in this region. If the newly calculated position gives the objective function a better value than its old position, the TD will accept it. For the TD, this location update procedure is replicated in (26).

$$R = 0.01\left(1 - \frac{t}{T}\right), \tag{24}$$

$$z_{i,j}^{new} = z_{i,j} + (2r - 1).R.z_{i,j}, \tag{25}$$

$$Z_i = \begin{cases} Z_i^{new}, & F_i^{new} < F_i \\ Z_i, & \text{otherwise} \end{cases}, \tag{26}$$

where $R$ denotes the neighborhood radius around the site of the attack, $t$ represents the iteration counter, $Z_i^{new}$ represents the new position of the $i$th TD in the neighborhood of $Z_i$, $T$ represents the most iteration, $z_{i,j}^{new}$ represents the value of the $j$th variable, and $F_i^{new}$ represents the value of the objective function.

Step 5: Termination condition: The algorithm will terminate once we have determined the optimal FIS.

**Pseudo code of ATDO**

**Input:** Candidate solutions or optimal FISs ($FIS_i$), TD ($TD$) size $N$, the variables' dimensionality $D$, the problem's upper and lower bounds, and the maximum number of iterations.
**Output:** Optimized FIS.
**Start** ATDO.
Specify the iteration count (T) and the population size (N).
Set the location of TDs to 0.
Assess the fitness function.
**For** t = 1:T
  **For** i = 1:N
    **If** the probability is 0.5, the probability equals rand.
    **Stage 1: Exploration stage**
    Using Equation (18), select carrion for the TD.
    Equation (19) is used to evaluate the new TD.
    Update the $i^{th}$ TD using Equation (20).
    **Else**
    **Stage 2: Exploitation phase**
    ***Prey attacking and selection***
    Choose prey for the $i^{th}$ TD using (21).
    Using Equation (22), determine the new state of TD.
    Equation (23) is used to update the TD.
    ***Prey chasing***
    Using Equation (24), update the neighborhood radius.
    Evaluate the updated state of the TD in the vicinity of utilizing Equation (25).
    Using Equation (26), update the $i$th TD.
    **End if**
  **End for**
  Save the greatest suggestion that has been made so far.
**End for**
Output: Optimized FIS.
**End** ATDO

## 2.4 | Updating RTO and Bottleneck Bandwidth Using C-Degree

A summary of the suggested BIFIS-CoAP is shown in Figure 8. The three primary states in the suggested framework are represented by the darkened blocks. The other blocks stand in for the BIFIS-CoAP's essential features. The innovation in BIFIS-CoAP is the creation of an FIS for the CoAP CC that monitors RTTs and bottleneck bandwidth, such as the RT-gradient and BG-gradient, to identify congestion as well as update sending rates to avoid congestion.

BIFIS-CoAP is effective for burst data transfers and supports in-flight packets thanks to dynamic rate management. BIFIS-CoAP limits the sending rate to the bandwidth–delay combination and preserves the in-flight packets inside the connection's maximum bottleneck bandwidth. Algorithms at the sender are the main component of BIFIS-CoAP, which enables efficient CC.

Every sender in BIFIS-CoAP makes an estimation of $bW_{\max}$ and $dR(w)$ at the connection's beginning phase. The starting stage normally lasts 6–10 RTTs. The sending rate is calculated using the estimated $dR(w)$. When an ACK gets received when the system is in a steady state, its sender regularly estimates RTT, $bW(w)$ and $dR(w)$. The FIS calculates the *C-degree* using the estimated RTT, $bW(w)$ and $dR(w)$ values, that are then utilized to modify the sending rate. Additionally, BIFIS-CoAP modifies the RTO values in the following ways using the *C-degree*:

$$RTT_s(w) = 0.75 \times RTT_s(w-1) + 0.25 \times RTT_m(w), \tag{27}$$

$$D_RTT_s(w) = RTT_s(w) - RTT_s(w-1), \tag{28}$$

$$RTO(w) = RTT_c(w) + C\text{-}degree(w) \times D_RTT_s(w). \tag{29}$$

In this case, $D_RTT_s(w)$ represents the estimated RTT variance at step $w$, $C$-$degree(w)$ represents the estimated crisp output of the FIS at stage $w$, and $RTO(w)$ represents the estimated retransmission timeout value at step $k$. With $\alpha = 0.25$, using Equation (2), the estimation of $RTT_s(w)$ is performed.

If the sender does not receive an ACK within the anticipated RTO, it tries to resend the pertinent packet and then searches for any potential transmission of subsequent packets. Every retransmitted packet is eligible for four retransmissions. Every retransmission updates the RTO value. The bottleneck bandwidth has not been calculated by BIFIS-CoAP using retransmitted packets.

After four unsuccessful retransmission attempts, the timeout timer function delivers the relevant lost packet for loss detection. When packet loss is discovered throughout the steady state,



10991131, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6131 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

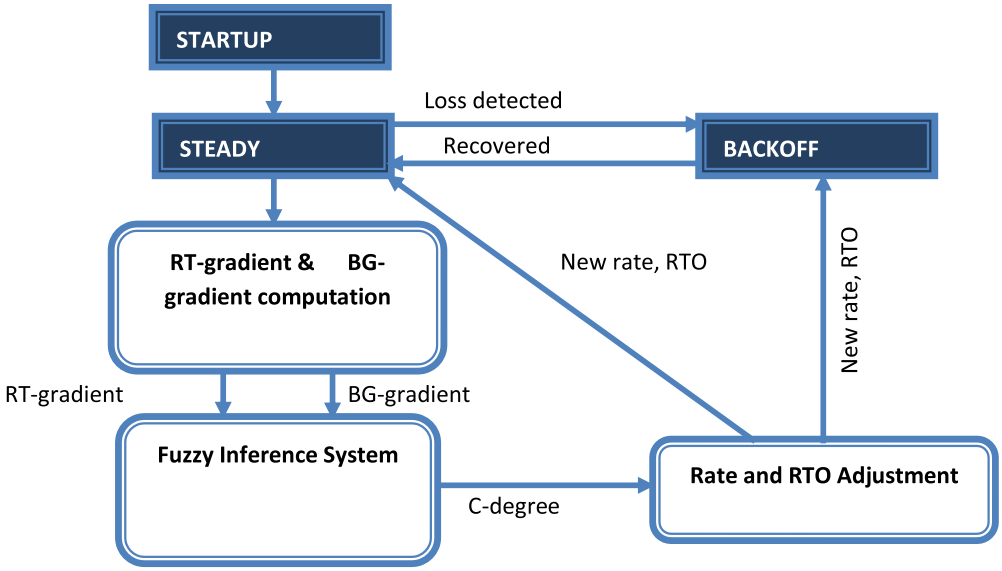



**FIGURE 8** | Updation of RTO and bottleneck bandwidth using congestion degree.

the sender enters the backoff state. Several packets, including newly broadcast and retransmitted packets, might be in flight right now. Packet loss shows that there has been congestion. The sender should not send any more packets to prevent future congestion. The recipient verifies the ACK while within the backoff state at the beginning of each RTT period. If an ACK is obtained, the sender changes the C-degree. The transmitter then resumes its constant condition. Following the maximum transaction duration, the transaction is deemed to have been lost if it does not.

The C-degree calculated by FIS is used to update the RTO values and adjust the sending rate in the BIFIS-CoAP scheme, helping to control congestion in IoT environments. The BIFIS-CoAP client adjusts the sending rate $S_{\text{rate}}(w)$ based on $C\text{-}degree(w)$ as follows:

- If $1 \geq C\text{-}degree(w) > 0$, the sending rate $S_{\text{rate}}(w)$ will be increased. The amount of the increase is the absolute value of the congestion degree $|C\text{-degree}(w)|$, which determines a fast or slow increase.

$$S_{\text{rate}}(w) = S_{\text{rate}}(w-1) + |C - degree(w)| \times 1 / RTT_s(w) \quad (30)$$

- If $C\text{-}degree(w) = 0$, the sending rate will remain unchanged.

$$S_{\text{rate}}(w) = S_{\text{rate}}(w-1) \quad (31)$$

- If $-1 \leq C\text{-}degree(w) < 0$, the sending rate will be decreased. The amount of the decrease is the absolute value of $C\text{-}degree$, which determines a fast or slow decrease.

$$S_{\text{rate}}(w) = S_{\text{rate}}(w-1) - |C\text{-}degree(w)| \times 1 / RTT_s(w) \quad (32)$$

Thus, the congestion can be controlled by using our proposed BIFIS-CoAP in an IoT environment.

## 3 | Results and Discussion

In this section, the efficacy of the proposed technique is examined. Python is used to implement the proposed strategy. In this part, we describe the simulation model that was utilized to assess the effectiveness of our BIFIS-CoAP. Then, we show the results for two distinct simulation settings, such as homogeneous traffic scenarios and heterogeneous traffic scenarios.

### 3.1 | Scenario 1: Simulation Model for the Homogeneous Traffic Scenarios

In Scenario 1, the performance of our proposed BIFIS-CoAP is evaluated against basic CoAP and fuzzy CoAP in scenarios with homogeneous traffic. Transmission employing the same CC mechanism, whether it is the proposed BIFIS-CoAP method or an existing one, involves homogeneous traffic circumstances. In these tests, the AP as well as the router had a 250-Kbps link bandwidth with a 50-ms link delay. The server and router were connected by a 45-Kbps link with a 300-ms delay. BIFIS-CoAP successfully transmitted 500 packets, whereas basic CoAP and fuzzy CoAP only transmitted 300 and 450 packets, respectively. The total number of BIFIS-CoAP packets that were successfully received, on the other hand, was much greater than in fuzzy CoAP and basic CoAP. The result was 350 for BIFIS-CoAP, 300 for fuzzy CoAP, then 150 for basic CoAP. The total number of packet retransmissions in BIFIS-CoAP was 110, which was significantly lower than the 120 in fuzzy CoAP then 130 in basic CoAP. BIFIS-CoAP got 45 duplicated packets, fuzzy CoAP had 60, and standard CoAP had 90. This study demonstrates that BIFIS-CoAP outperformed fuzzy CoAP and basic CoAP. This suggests that in a homogeneous traffic scenario, BIFIS-CoAP achieved good performance with no congestion. Table 3 represents the parameter settings for the homogeneous traffic scenarios.



10991131, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6131 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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** | Parameter settings for the homogeneous traffic scenarios.

| Parameter | Values |
|---|---|
| Transmission mechanism | BIFIS-CoAP, basic CoAP, and fuzzy CoAP |
| Traffic type | Homogeneous |
| Link bandwidth (AP and router) | 4000 Kbps |
| Link delay (AP and router) | 50 ms |
| Link bandwidth (server and router) | 5000 Kbps |
| Link delay (server and router) | 300 ms |
| Network topology | Star |

## 3.2 | Performance Analysis Based on Delay Time

Congestion can cause delays during packet transmission. In order to prevent delays, BIFIS-COAP is used in this study to gather packets and send them to their destination while calculating the degree of congestion. The proposed method is contrasted with standard CoAP and fuzzy CoAP in order to show its efficacy. Figure 9 provides an analysis of the delay time effectiveness of the proposed and existing techniques.

According to the analysis of Figure 9, the proposed strategy was able to achieve the least amount of delay in all simulations when compared to basic CoAP and fuzzy CoAP, based on the results from all simulations. There was a delay of 3728 ms for the proposed approach during 50 s of the simulation, which is 5386 ms for fuzzy CoAP and 7437 ms for basic CoAP during 50 s of the simulation. During our simulation of BIFIS-CoAP, we achieved a minimal delay of 3672 ms, whereas fuzzy CoAP and basic CoAP achieved a delay of 5068 and 7837 ms, respectively, during 150 ms of simulation. Across all scenarios, we found that the approach we proposed to produce the least amount of delay was the one we recommended. This demonstrates the efficacy of our approach in reducing delay. Furthermore, our approach is easy to implement and could be implemented with minimal modifications.

## 3.3 | Performance Analysis Based on Throughput

BIFIS-CoAP has been compared in terms of throughput with basic CoAP and fuzzy CoAP to show the efficacy of the suggested method. The throughput evaluation for the BIFIS-CoAP, basic CoAP, and fuzzy CoAP is shown in Figure 10.

Having compared the proposed method to basic CoAP and fuzzy CoAP, Figure 10 shows that it consistently achieved the highest throughput as compared to the other two. For the 50 s of the simulation, the proposed approach resulted in a throughput of 2737 Kbps, whereas fuzzy CoAP achieved a throughput of 1257 Kbps and basic CoAP obtained a throughput of 965 Kbps. There was a sudden change in the BIFIS-CoAP from 2457 Kbps to 2937 Kbps at the 150-s mark. In the same scenario, BIFIS-CoAP managed to achieve a high throughput of 3014 Kbps at 200 s, whereas fuzzy CoAP and basic CoAP only managed to reach 2037 and 785 Kbps, respectively, at 200 s. Compared to fuzzy CoAP and basic CoAP, the results of the study show that BIFIS-CoAP runs without congestion with high throughput in all scenarios compared to fuzzy CoAP and basic CoAP. The consistent achievement of the highest throughput by BIFIS-CoAP in all scenarios compared to fuzzy CoAP and basic CoAP is significant because it indicates that BIFIS-CoAP is a more efficient and reliable method for data transmission. This high throughput ensures faster and more reliable communication, which is crucial in applications where real-time data exchange is essential, such as IoT systems or time-sensitive industrial processes.

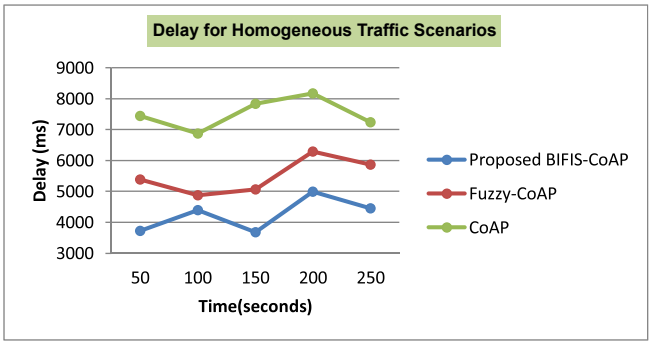



**FIGURE 9** | Performance analysis based on delay time.

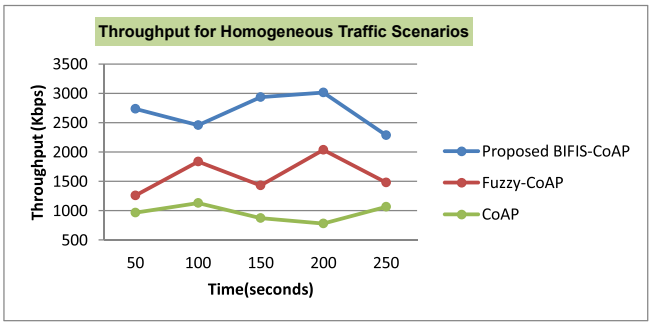



**FIGURE 10** | Performance analysis based on throughput.

## 3.4 | Performance Analysis Based on Delivery Ratio

The packet delivery ratio of the proposed BIFIS-CoAP method is compared with basic CoAP and fuzzy CoAP, and it is given in Figure 11.

As shown in Figure 11, when compared with basic CoAP and fuzzy CoAP, the proposed technique consistently achieved the highest delivery ratio when compared with the other two. The proposed approach achieved an 87% delivery ratio within the first 50 s of the simulation, compared to 57% for basic CoAP and 82% for fuzzy CoAP during the same period. After 150 s, the BIFIS-CoAP decreased from 92% to 90%. BIFIS-CoAP achieved a maximum delivery ratio of 95% in 200 simulation seconds, whereas fuzzy CoAP achieved a maximum delivery ratio of 78% in the same scenario, whereas basic CoAP achieved a maximum



10991131, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6131 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

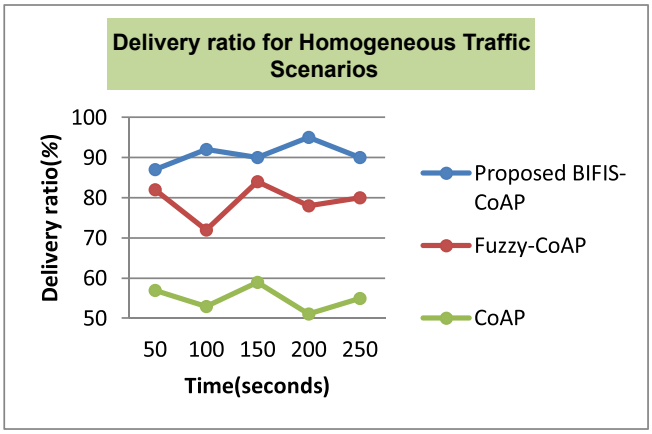



**FIGURE 11** | Performance analysis based on delivery ratio.

delivery ratio of 51%. In all scenarios in which BIFIS-CoAP is compared to fuzzy CoAP and basic CoAP, it is demonstrated that BIFIS-CoAP operates in a congestion avoidance zone that has a high delivery ratio. This demonstrates that BIFIS-CoAP is an efficient protocol for congestion avoidance in IoT networks. BIFIS-CoAP is also capable of achieving higher delivery ratios than other protocols, making it an ideal choice for IoT applications.

### 3.5 | Scenario 2: Simulation Model for the Heterogeneous Traffic Scenarios

In Scenario 2, the performance of our suggested BIFIS-CoAP in scenarios with heterogeneous traffic has been contrasted to that of fuzzy CoAP and basic CoAP. The AP although the router's link bandwidth was 250 Kbps, with the link delay was 50 ms. Among the router and the edge router, the link bandwidth was 60 Kbps, and the link delay was 300 ms. The edge router as well as servers have a 1-Mbps link bandwidth with a 20-ms link latency. BIFIS-CoAP successfully transmitted 500 packets, whereas fuzzy CoAP and basic CoAP only transmitted 450 and 300 packets, respectively. The total number of BIFIS-CoAP packets that were successfully received, on the other hand, was much greater than in fuzzy CoAP and basic CoAP. This figure was 400 for BIFIS-CoAP, 310 for fuzzy CoAP, and 120 for standard CoAP. The total number of packet retransmissions in BIFIS-CoAP was 90, which was significantly lower than in fuzzy CoAP, which was 100, and in basic CoAP, which was 110. Then, the number of duplicated packets in BIFIS-CoAP was 30; in fuzzy CoAP, 50; and in basic CoAP, 80. This result indicates that BIFIS-CoAP outperformed fuzzy CoAP and basic CoAP. In the situation of heavy congestion, the simulation results showed that BIFIS-CoAP outperformed fuzzy CoAP and basic CoAP. Table 4 represents parameter settings for the homogeneous traffic scenarios.

### 3.6 | Performance Analysis Based on Delay Time

In order to prevent delays, BIFIS-COAP is used in this study to gather packets and send them to their destination while calculating the degree of congestion. The proposed method is contrasted with standard CoAP and fuzzy CoAP in order to show its efficacy. Figure 12 provides an analysis of the performance of the proposed and existing strategies in terms of delay time.

**TABLE 4** | Parameter settings for the heterogeneous traffic scenarios.

| Parameter | Value/setting |
|---|---|
| Transmission mechanism | BIFIS-CoAP, basic CoAP, and fuzzy CoAP |
| Traffic type | Heterogeneous |
| Link bandwidth (AP and router) | 5000 Kbps |
| Link delay (AP and router) | 50 ms |
| Link bandwidth (router and edge router) | 5000 Kbps |
| Link delay (router and edge router) | 300 ms |
| Link bandwidth (edge router and servers) | 6000 Kbps |
| Link delay (edge router and servers) | 20 ms |
| Network topology | Dumbbell |

The proposed method achieved a minimum delay in all simulations when compared to the basic CoAP method and fuzzy CoAP method when analyzing Figure 12. The proposed method was found to achieve a minimum delay of 2210 ms for 50 s of the simulation, which is 4257 ms for fuzzy CoAP and 7085 ms for the basic CoAP. In the 150-s period, the proposed BIFIS-CoAP is changed from 3247 to 2768 ms. In 200 s of simulation, our proposed BIFIS-CoAP reached a delay of 4358 ms, whereas our proposed fuzzy CoAP reached 5768 ms and our proposed basic CoAP reached 6873 ms in the same simulation. According to Figure 10, all simulations were carried out with the proposed method, and there was a minimum delay in all cases. This shows that our proposed BIFIS-CoAP is the most efficient of the three methods, achieving the lowest delay in all simulations.

### 3.7 | Performance Analysis Based on Throughput

To demonstrate the efficiency of the proposed method, BIFIS-CoAP is contrasted with basic CoAP as well as fuzzy CoAP in terms of throughput. Figure 13 depicts the throughput evaluation for the BIFIS-CoAP, basic CoAP, and fuzzy CoAP.

In all simulation scenarios depicted in Figure 13, the proposed method consistently surpassed both basic CoAP and fuzzy CoAP in terms of performance. Specifically, the proposed method achieved a notable throughput advantage, reaching 3428 Kbps within the initial 50 s of simulation. This contrasted sharply with the throughput values of 2438 Kbps for fuzzy CoAP and 1254 Kbps for basic CoAP during the same timeframe. As the simulation progressed to 150 s, the proposed BIFIS-CoAP maintained its superiority, with a throughput of 3245 Kbps,



10991131, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6131 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

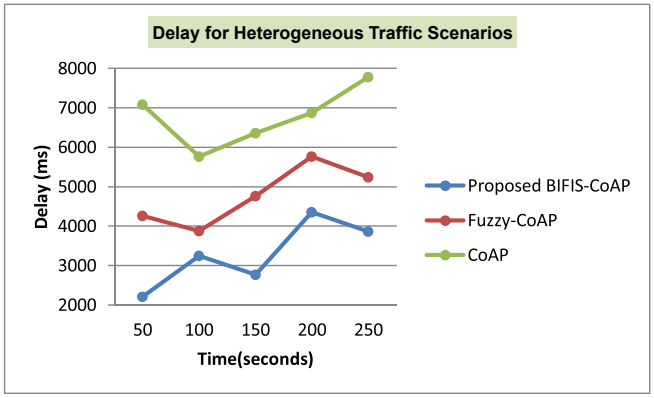



**FIGURE 12** | Performance analysis based on delay time.

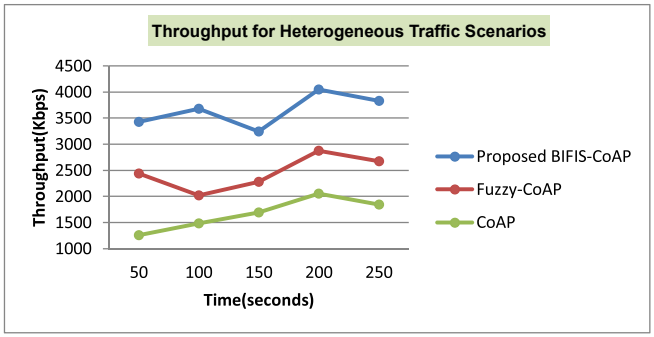



**FIGURE 13** | Performance analysis based on throughput.

showcasing its stability and efficiency over time. By the 200-s mark, BIFIS-CoAP further distinguished itself by achieving a remarkable throughput of 4045 Kbps, significantly outpacing both fuzzy CoAP at 2878 Kbps and basic CoAP at 2054 Kbps. These findings unequivocally demonstrate that BIFIS-CoAP consistently operates within a congestion avoidance zone, ensuring high throughput across various simulation scenarios. Its ability to sustain elevated throughput levels highlights its effectiveness in managing network congestion and optimizing data transmission efficiency.

### 3.8 | Performance Analysis Based on Delivery Ratio

The packet delivery ratio of the proposed BIFIS-CoAP method is compared with basic CoAP and fuzzy CoAP, and it is given in Figure 14.

Upon analyzing Figure 14, it becomes evident that the proposed strategy consistently outperformed both basic CoAP and fuzzy CoAP across all simulations, particularly in terms of delivery ratio. In the initial 50 simulations, the proposed technique achieved an impressive delivery ratio of 90%, surpassing fuzzy CoAP's 83% and basic CoAP's 64%. This substantial lead underscores the effectiveness of the proposed approach in ensuring successful message delivery. Even as the simulation progressed to 150 simulations, where BIFIS-CoAP experienced a slight drop from 94% to 93% in delivery ratio, it remained significantly ahead of fuzzy CoAP and basic CoAP. By the 200th simulation,

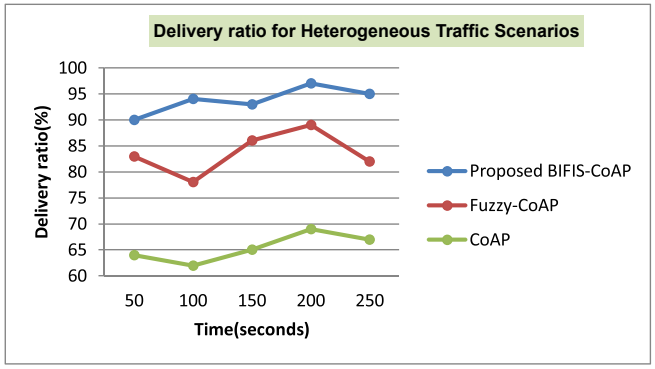



**FIGURE 14** | Performance analysis based on delivery ratio.

BIFIS-CoAP reclaimed its momentum, achieving a maximum delivery ratio of 97%. In comparison, fuzzy CoAP managed 89%, and basic CoAP lagged behind at 82%. These consistent findings across all simulations affirm that BIFIS-CoAP operates within a congestion avoidance zone while maintaining a high delivery ratio, outperforming both fuzzy CoAP and basic CoAP.

### 3.9 | Comparative Analysis With the State-Of-The-Art Methods

In order to assess the proposed BIFIS-CoAP method, the results are compared to other literature works in terms of delivery ratio, throughput, delay, and retransmissions. In [12], IDC-CoAP and MBC-CoAP were both implemented within the Contiki OS environment and evaluated using the Cooja simulator and MSPSim emulator, focusing on improving retransmission timeout calculations and employing rate-based CC techniques. However, BIFIS-CoAP introduces a more advanced approach by integrating a FIS optimized with the ATDO algorithm. This combination allows BIFIS-CoAP to predict congestion and dynamically adjust sending rates and retransmission timeouts, which not only maintains high goodput but also reduces application loss ratios more effectively. Although IDC-CoAP and MBC-CoAP focus on simplicity and efficiency, BIFIS-CoAP advances these objectives by enhancing the accuracy and adaptability of CC mechanisms, resulting in overall improved performance in diverse network conditions. Comparative analysis based on other literature works is as follows:

Table 5 represents the comprehensive comparative analysis of the proposed method, which is compared with previous literature works. As stated in Table 5, the literature works [13–19] are taken into consideration for comparison. Among the previous literature works, the literature work [13] attained 4560 ms of delay, and the literature works [15, 19] achieved 5000 ms, which is higher than our proposed method's delay of 2210 ms. As depicted in Table 1, the packet delivery ratio achieved by the literature works [16–18] is comparatively minimal, but our proposed method's delivery ratio is 97%, which is higher than other literature works. By comparing throughput our method attained a maximum throughput of 4045 Kbps, which is probably higher than other literature works such as [13, 14, 16, 18, 19]. These results collectively underscore the efficacy of our proposed method across various performance metrics. Furthermore, with only 90 retransmissions and 30 duplicated packets, the proposed



10991131, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6131 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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** | Comparative analysis based on other literature works.

| Literature works | Delivery ratio (%) | Throughput (Kbps) | Delay time (ms) | Retransmission (no. of packets = 500) | Duplicated packets (no. of packets = 500) |
|---|---|---|---|---|---|
| [13] | — | 520 | 4560 | — | — |
| [14] | — | 910 | 9500 | — | — |
| [15] | — | — | 5000 | — | — |
| [16] | 90% | 670 | 1800 | — | — |
| [17] | 92% | — | — | — | — |
| [18] | 94% | 700 | — | — | — |
| [19] | — | 920 | 5000 | 275 | — |
| Proposed | 97% | 4045 | 2210 | 90 | 30 |

approach demonstrates superior reliability and data integrity management compared to the literature works.

### 3.10 | Limitations

However, it is important to note that the proposed BIFIS-CoAP also has some limitations. One limitation is that it requires the availability of accurate and up-to-date congestion information in order to make effective decisions on adjusting the sending rate and updating the variable RTO. Additionally, the performance of BIFIS-CoAP may vary depending on the specific characteristics of the IoT network and the traffic patterns. Further research and experimentation are needed to explore the scalability and robustness of BIFIS-CoAP in larger and more complex IoT environments. Moreover, another challenge encountered in this research was the complexity of implementing the bio-inspired FIS and the ATDO algorithm. The integration of these two techniques required a significant amount of computational resources and expertise in both fuzzy logic and optimization algorithms. In future work, efforts could be focused on simplifying the implementation of the bio-inspired FIS and the ATDO algorithm. This could involve developing more user-friendly tools and interfaces that require less computational resources and expertise, making it more accessible for researchers and practitioners in the field. Additionally, exploring alternative optimization algorithms that can achieve similar or better results with less complexity could also be a promising direction for future research.

## 4 | Conclusion

In this paper, a bio-inspired FIS-based constrained application protocol has been presented for avoiding congestion over network. In order to improve CC in IoT networks, the FIS has been used in conjunction with CoAP. Then, the ATDO algorithm was employed to enhance the performance of FIS by optimizing the parameters of FIS such as BG-gradient and RT-gradient. The degree of congestion obtained from the FIS predicted the early congestion and adjusted the sending rate in BIFIS-CoAP. As a result, BIFIS-CoAP updated the variable RTO for retransmissions and bandwidth to achieve excellent performance for packet transfer using this C-degree. Thus, the proposed BIFIS-CoAP transmits packets without congestion. The efficiency of the proposed BIFIS-CoAP has been evaluated and compared with other methods such as basic CoAP and fuzzy CoAP in two different scenarios such as homogeneous traffic scenarios and heterogeneous traffic scenarios. In those two scenarios, our proposed BIFIS-CoAP achieved maximum throughput, high delivery ratio, and minimum delay when compared to fuzzy CoAP and basic CoAP. In future work, efforts could be focused on simplifying the implementation of the bio-inspired FIS and the ATDO algorithm. This could involve developing more user-friendly tools and interfaces that require less computational resources and expertise, making it more accessible for researchers and practitioners in the field. Additionally, exploring alternative optimization algorithms that can achieve similar or better results with less complexity could also be a promising direction for future research.

#### Data Availability Statement

Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

#### References

1. F. Dalipi and S. Y. Yayilgan, "Security and Privacy Considerations for IoT Application on Smart Grids: Survey and Research Challenges," in *IEEE Conference on Future Internet of Things and Cloud Workshops (FiCloudW)*, (2016): 63–68.

2. S. Latre, P. Leroux, T. Coenen, B. Braem, P. Ballon, and P. Demeester, "City of Things: An Integrated and Multi-Technology Testbed for IoT Smart City Experiments," in *ISC2'16, IEEE*, (2016), 1–8.

3. J. Ren, Y. Zhang, N. Zhang, D. Zhang, and X. Shen, "Dynamic Channel Access to Improve Energy Efficiency in Cognitive Radio Sensor Networks," *IEEE Transactions on Wireless Communications* 15, no. 5 (2016): 3143–3156.

4. D. Zhang, Z. Chen, J. Ren, et al., "Energy Harvesting-Aided Spectrum Sensing and Data Transmission in Heterogeneous Cognitive Radio Sensor Network," *IEEE Transactions on Vehicular Technology* 66, no. 1 (2016): 831–843.

5. L. P. Verma and M. Kumar, "An IoT Based Congestion Control Algorithm," *Internet of Things* 9 (2020): 100157.



10991131, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6131 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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. J. Lorincz, Z. Klarin, and J. Ožegović, "A Comprehensive Overview of TCP Congestion Control in 5G Networks: Research Challenges and Future Perspectives," *Sensors* 21, no. 13 (2021): 4510.

7. S. Bansal and D. Kumar, "Distance-Based Congestion Control Mechanism for CoAP in IoT," *IET Communications* 14, no. 19 (2020): 3512–3520.

8. P. Anitha, H. S. Vimala, and J. Shreyas, "Comprehensive Review on Congestion Detection, Alleviation, and Control for IoT Networks," *Journal of Network and Computer Applications* 221 (2024): 103749.

9. D. Wischik, C. Raiciu, A. Greenhalgh, and M. Handley, "Design, Implementation and Evaluation of Congestion Control for Multipath TCP," in *Proceedings of the 8th USENIX Conference on Networked Systems Design and Implementation*, (2011).

10. Y. Cui et al., "FMTCP: A Fountain Code-Based Multipath Transmission Control Protocol," in *Proceedings of IEEE Conference on Distributed Computing Systems (ICDCS)*, (2012).

11. C. Bormann, A. P. Castellani, and Z. Shelby, "CoAP: An Application Protocol for Billions of Tiny Internet Nodes," *IEEE Internet Computing* 16, no. 2 (2012): 62–67.

12. N. Makarem, W. B. Diab, I. Mougharbel, and N. Malouch, "On the Design of Efficient Congestion Control for the Constrained Application Protocol in IoT," *Computer Networks* 207 (2022): 108824.

13. K. Singh, K. Singh, L. H. Son, and A. Aziz, "Congestion Control in Wireless Sensor Networks by Hybrid Multi-Objective Optimization Algorithm," *Computer Networks* 138 (2018): 90–107.

14. Y. Lu, Z. Ling, S. Zhu, and L. Tang, "SDTCP: Towards Datacenter TCP Congestion Control With SDN for IoT Applications," *Sensors* 17, no. 1 (2017): 109.

15. I. A. Najm, A. K. Hamoud, J. Lloret, and I. Bosch, "Machine Learning Prediction Approach to Enhance Congestion Control in 5G IoT Environment," *Electronics* 8, no. 6 (2019): 607.

16. N. Akhtar, M. A. Khan, A. Ullah, and M. Y. Javed, "Congestion Avoidance for Smart Devices by Caching Information in MANETS and IoT," *IEEE Access* 7 (2019): 71459–71471.

17. Y. Zhuang, L. Yu, H. Shen, et al., "Data Collection With Accuracy-Aware Congestion Control in Sensor Networks," *IEEE Transactions on Mobile Computing* 18, no. 5 (2019): 1068–1082.

18. J. P. A. León, L. J. de la Cruz Llopis, and F. J. Rico-Novella, "A Machine Learning Based Distributed Congestion Control Protocol for Multi-Hop Wireless Networks," *Computer Networks* 231 (2023): 109813.

19. C. Lim, "Improving Congestion Control of TCP for Constrained IoT Networks," *Sensors* 20, no. 17 (2020): 4774.

20. H. Xie, Y. Jing, and J. Chen, "Adaptive Fuzzy Practical Prescribed Time H∞ Congestion Control for Network Systems With Input Saturation and External Disturbance," *Expert Systems With Applications.* 234 (2023): 120930.

21. Y. Jing, H. Xie, J. Chen, and G. M. Dimirovski, "Finite-Time Congestion Tracking Control for TCP/AWM Network Systems Employing Event-Triggered Mechanism," *Journal of the Franklin Institute* 359, no. 14 (2022): 7233–7261.

22. M. Dehghani, Š. Hubálovský, and P. Trojovský, "Tasmanian Devil Optimization: A New Bio-Inspired Optimization Algorithm for Solving Optimization Algorithm," *IEEE Access* 10 (2022): 19599–19620.



10991131, 2025, 4, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6131 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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 Electronics

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tetn20

# Design and analysis of SRAM cell using swing restoration inverter for low power applications

Ram Murti Rawat & Vinod Kumar

**To cite this article:** Ram Murti Rawat & Vinod Kumar (11 Jan 2025): Design and analysis of SRAM cell using swing restoration inverter for low power applications, International Journal of Electronics, DOI: 10.1080/00207217.2025.2450733

**To link to this article:** https://doi.org/10.1080/00207217.2025.2450733

Published online: 11 Jan 2025.

Submit your article to this journal

Article views: 20

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tetn20

INTERNATIONAL JOURNAL OF ELECTRONICS
https://doi.org/10.1080/00207217.2025.2450733

Taylor & Francis
Taylor & Francis Group

Check for updates

# Design and analysis of SRAM cell using swing restoration inverter for low power applications

Ram Murti Rawat and Vinod Kumar

Department of Computer Science & Engineering, Delhi Technological University, Delhi, India

**ABSTRACT**

SRAM cells play a crucial role in the design of system-on-chips (SoCs), constituting a substantial portion of the die area and thereby contributing to increased power consumption. Despite significant advancements in SRAM performance in finer technologies, concerns persist regarding cell stability in the deep sub-micron domain. This article introduces an innovative approach to address these issues by proposing a low-power SRAM cell based on Swing Restoration Inverter (SRI). The Swing Restoration Inverter (SRI) addresses the challenge of swing voltages in finer technologies by integrating two additional transistors to act as swing-restored elements. This innovation results in an impressive 67% reduction in leakage power at 27°C for the 90 nm technology. To evaluate the effectiveness of the proposed SRI-based SRAM cell, performance metrics such as delay and power delay product are calculated. In the context of a 6T SRAM cell in the 90 nm technology, the conventional inverters are replaced with the Swing Restoration Inverter (SRI). The implementation of SRI in an 8T SRAM cell leads to a substantial 91% reduction in leakage power at 30°C and an impressive 93% reduction at 27°C compared to the conventional 6T SRAM cell. The SRI technique proves instrumental in enhancing stability, resulting in the proposed 8T SRAM cell outperforming its 6T SRAM counterpart. However, it's important to note that the heightened performance of the proposed 8T SRAM cell comes at the expense of four additional transistors. Simulations are conducted using the 90 nm technology and the Cadence Virtuoso simulator to validate the effectiveness of the Swing Restoration Inverter (SRI) technique in mitigating power consumption and enhancing stability in SRAM cells.

**ARTICLE HISTORY**
Received 5 February 2024
Accepted 6 October 2024

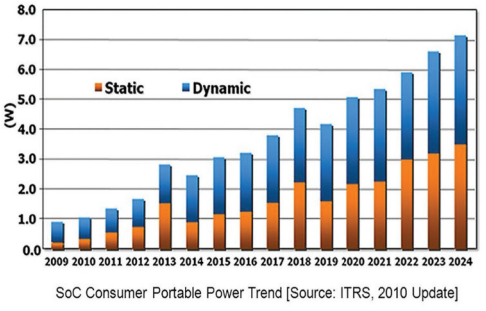
**KEYWORDS**
leakage power; stability; power delay product; swing restoration inverter (SRI); CMOS inverter

## 1. Introduction

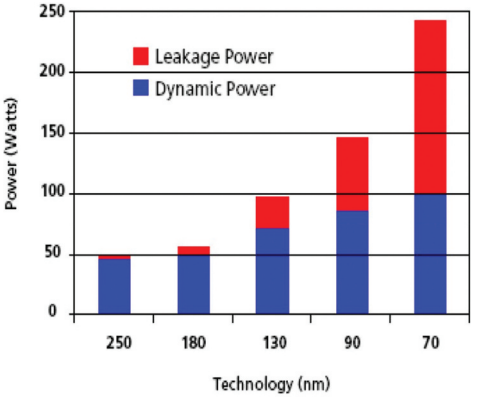
In portable and battery-operated devices, System-on-Chip (SoC) technology is often used (Benmoussa et al., 2013; Carroll & Heiser, 2010; Hoque et al., 2014). As functionality, circuit complexity, and density demands rise, a shift to lower technology nodes enables the requisite functionality scaling. Faster and higher-performing gadgets are made possible by technological scaling, but as technology scales down, power efficiency tends to decline (J. S. Wang et al., 2011). Figure 1(a). depicts the power dissipation of a SoC chip in

**CONTACT** Ram Murti Rawat rammurtirawat@dtu.ac.in Department of Computer Science & Engineering, Delhi Technological University, Bawana Rd, Shahbad Daulatpur Village, Rohini, Delhi 110042, India

© 2025 Informa UK Limited, trading as Taylor & Francis Group




(a) Power trend in SOC




(b) Power consumption in various technology

**Figure 1.** Evolution of leakage power (Benmoussa et al., 2013; Carroll & Heiser, 2010; Hoque et al., 2014; Krishna & Duraiswamy, 2020b; Mutoh et al., 1995; J. S. Wang et al., 2011).

consumer electronics throughout the years (source ITRS). The overall power dissipation of an Intel microprocessor shows a discernible increase, but the growth in static and dynamic power contributions is not directly proportional, as depicted in Figure 1(b). The increase in dynamic power consumption is closely tied to the expanding functionality, whereas static power is dependent on the technology node and circuit configuration. Notably, at lower nodes, power distribution is mostly driven by static power dissipation. Conversely, at higher nodes, dynamic power dissipation emerges as a substantial element of the overall power dissipation. Static Random Access Memory (SRAM) dominates the chip's physical space in System-on-Chip (SoC) designs, occupying a significant share of the chip area. For example, in a video decoder chip, it encompasses more than 65% of the total space (J. S. Wang et al., 2011) and contributes to over 30% of the overall power consumption in mobile devices (Benmoussa et al., 2013; Carroll & Heiser, 2010; Hoque et al., 2014; Krishna & Duraiswamy, 2020b; Mutoh et al., 1995; J. S. Wang et al., 2011). The use of SRAM in mobile devices introduces increased leakage power due to extended idle periods. The excess leakage power adversely affects battery life, raises cooling and packaging costs, and gives rise to reliability concerns in memory devices. Despite significant advancements in SRAM performance achieved through finer technologies, challenges persist in terms of cell stability and leakage power dissipation, especially in the deep sub-micron regime (Krishna & Duraiswamy, 2020a).

Minimising power dissipation caused by leakage and enhancing the stability of SRAM cells are vital objectives, particularly as semiconductor technology advances into the nanoscale realm. Traditional methods for reducing power consumption in CMOS circuits, such as adjusting supply voltage and scaling devices, have proven effective. It's important to highlight that voltage levels are currently trending towards the 1.0–0.9 V range. This suggests that upcoming progress might not necessarily translate into substantial power reductions. On the flip side, the growth in MOS dimensions has led to a notable rise in leakage current, posing obstacles to maintaining SRAM data stability. Therefore, developing an SRAM cell with minimal leakage and robust stability in the nanoscale domain is

a formidable challenge. Various strategies have been proposed in the literature to mitigate leakage power and enhance SRAM cell stability (Prasad & Anand, 2015). At the circuit level, a multi-threshold CMOS approach (Subramanyam & Basha, 2016) introduces two additional transistors, termed sleep transistors, to reduce leakage power in SRAM designs (Sachan et al., 2016; Venkatareddy et al., 2016). However, it is crucial to note that this method results in increased dynamic power. The super cut-off CMOS technology, designed to limit leakage current in SRAM, requires an additional control circuit (Narendra et al., 2002). Another technique, transistor stacking, involves replacing a single transistor with two transistors, a practice known as forced stacking (Hanchate & Ranganathan, 2004). Nevertheless, the introduction of extra latency into the circuits is an inevitable consequence. The LECTOR (LEakage Control TransistOR) system solves this problem by connecting two self-controlled leakage control transistors in a line from the supply to ground. This configuration increases resistance and reduces leakage current without necessitating additional control circuitry (Priya et al., 2012). It is crucial to emphasise that the size of the sleep transistor in LECTOR is a critical factor to consider (Nandyala & Mahapatra, 2016). Recently, there has been a notable advancement in swing restoration techniques, involving the incorporation of two transistors to augment the source voltage. This source biasing strategy has demonstrated superior power efficiency compared to the LECTOR technique across various logic circuits at the 100 nm scale (Wang et al., 2007). Importantly, swing restoration has proven effective without necessitating a control circuit. It is imperative to further investigate the applicability of swing restoration methods for larger circuits, such as SRAM, and to explore the impact on dynamic power and method efficiency in deep sub-micron technologies. Reducing leakage power while ensuring optimal stability is crucial for SRAMs. Supportive techniques, particularly when operating at reduced supply voltage, can enhance static noise margin (SNM), specifically read static noise margin (read-stability) and write static noise margin (write-ability) (Suneja et al., 2017). Established techniques in the literature include cell voltage adjustment (Nii et al., 2008), modulation of word line voltage (Hirabayashi et al., 2009; Suzuki et al., 2008), dual-rail supply schemes (Tachibana et al., 2014), negative bit-line schemes (Shibata et al., 2006; Wang et al., 2007), and gate biasing utilising FinFET technology (Sayyah et al., 2019). In conclusion, the pursuit of energy-efficient and stable SRAM cells involves a nuanced exploration of various strategies and methodologies, each presenting unique trade-offs and challenges. A comprehensive understanding of these approaches is essential for advancing the field and achieving the delicate balance between power efficiency and stability in SRAM design. VDD boosting emerges as a highly effective method for enhancing stability; however, it comes at the cost of increased leakage power. This article presents a recommended design for a low-power 8T SRAM cell that uses two PMOS transistors. Two connections are made, one to the node voltage Q and the other to QB, with the goal of removing the swing voltage in the bit and bitbar lines. As the cell operates, the swing voltage decreases, leading to reduced dynamic power dissipation. The pre-charge voltage level technique utilises an NMOS transistor as a resistor and a PMOS transistor as a switch, causing a series of drops in leakage current as the transistors transition between inactive and active states. Comparative analyses demonstrate the consistent outperformance of the proposed 8T SRAM cell over the traditional 6T SRAM cell in read, write, and hold modes. This superiority stems from increased static noise margins, ensuring enhanced bit cell write ability (Rawat & Kumar, 2021), and other measures that minimise static power dissipation while improving read/write performance. In (Bai et al., 2022) study, a stability analysis based

on leakage current is suggested. It introduces metrics like read current noise tolerance (RINM) and write current noise tolerance (WINM) to assess the stability of subthreshold SRAMs (J. Chen et al., 2022). introduce a novel 8T compute SRAM (CSRAM) designed for reliable and high-speed in-memory searching and compound logic-in-memory computations. This CSRAM incorporates a pair of PMOS access transistors and split WLs dedicated to compute access (Y. H. Chen et al., 2012). present an approach that eliminates the need for additional margins related to BL voltage swing, traditionally required to ensure tolerances against simulation errors and inaccurate estimation of SA offset voltage (Hamzaoglu et al., 2002). compare cell stability, noise margin, performance, and power of various dual-VT design choices for a large on-chip cache with single-ended, full-swing sensing in a 0.13-m technology generation (Joy & Kuruvilla, 2022). propose a system designed to optimise a single-bit memory cell of conventional static random-access memory, achieving stability with low power consumption and significantly low Power-Delay-Product (PDP) by varying operating frequencies in the MHz range (Kang et al., 2016). introduced a novel differential SRAM architecture featuring a full-swing local bit line (BL). This was achieved by ensuring a complete swing of the local BL through cross-coupled p-type metal-oxide-semiconductor (PMOS) transistors. In addition, the read buffer's gate was operated with a full supply voltage (VDD), removing the requirement for a boosted word line voltage. In the work by (Liu et al., 2017), a compensation scheme was presented to address threshold voltage mismatch in sense amplifier (SA) transistors. The accomplishment in question was realised through the digital and iterative adjustment of the body bias voltage. In their work (Mani et al., 2022), and the research team advocated for a design that prioritises low power consumption in writing, reading, and holding modes. Their strategy focused on maintaining robust static noise margins across all operations while minimising read access time. In a prior study (Miyano et al., 2013), delved into power inefficiencies within low-voltage SRAM operating near the threshold voltage. They pinpointed issues such as excessive bit line swing arising from random transistor variation and dynamic power consumption linked to the bit line swing in non-selected columns (Mukku & Lorenzo, 2023). introduced a 10T SRAM cell specifically crafted for energy-efficient operation, featuring high noise margins and radiation-hardened characteristics. Meanwhile (W. Wang et al., 2022), investigated the influence of power supply ramp rate during power-up on SRAM power-up states. They addressed instability concerns leading to inconsistent responses from SRAM Physical Unclonable Functions (PUFs). Their research introduced an effective stable cell selection method to pinpoint SRAM cells that exhibit strong bias, resistance to circuit noise, and resilience to voltage, temperature fluctuations, and ageing. Despite these advancements, the study concluded that additional endeavours are required to achieve more promising outcomes.

(1) To solve the issue of SRAM stability and delay, a novel Swing Restoration Inverter (SRI) circuit is developed and implemented utilising proposed 8T SRAM cell.
(2) A Swing restoration Inverter (SRI) is used to improve the cell's stability during read and write operations.
(3) By adding two more transistors, the leakage power is minimised by increasing the node voltages and decreasing the threshold voltages.
(4) As a result of the recommended enhancement, SRAM's stability and latency have improved.

The aim of this research is to develop an energy-efficient SRAM cell with minimal leakage by integrating a Swing Restoration Inverter (SRI). A thorough analysis of SRI performance is carried out using both DC and transient analyses, focusing on critical parameters such as leakage power, dynamic power, delays, and power delay product (PDP). In order to achieve a configuration with low leakage in the 8T SRAM, the conventional inverters in the 6T SRAM architecture are replaced with Swing Restoration Inverters (SRIs). The stability of the proposed 8T SRAM is assessed based on read and write margins. To improve the stability of the suggested SRAM, appropriate assist techniques are integrated, and the results are compared with those of the conventional 6T SRAM design.

## 2. Swing restoration inverter (SRI)

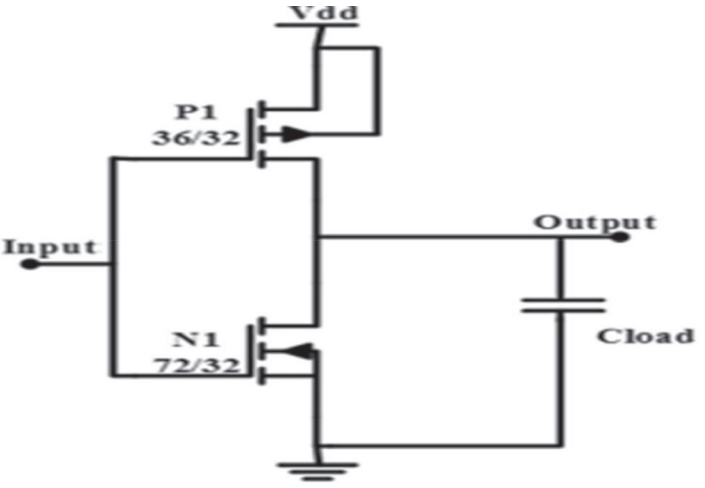
The sub-threshold leakage current outperforms other leakage currents in MOS devices, the Gate Induced Drain Leakage (GIDL) current and the reverse bias source/drain junction leakage current, for illustration. Equations (1) and Equation (2) illustrate the sub-threshold leakage current.

$$I_{Subleakage} = I_0 e^{(Vgs-Vth)/nV_{th}} \left(1 - e^{-Vds/Vth}\right) \tag{1}$$

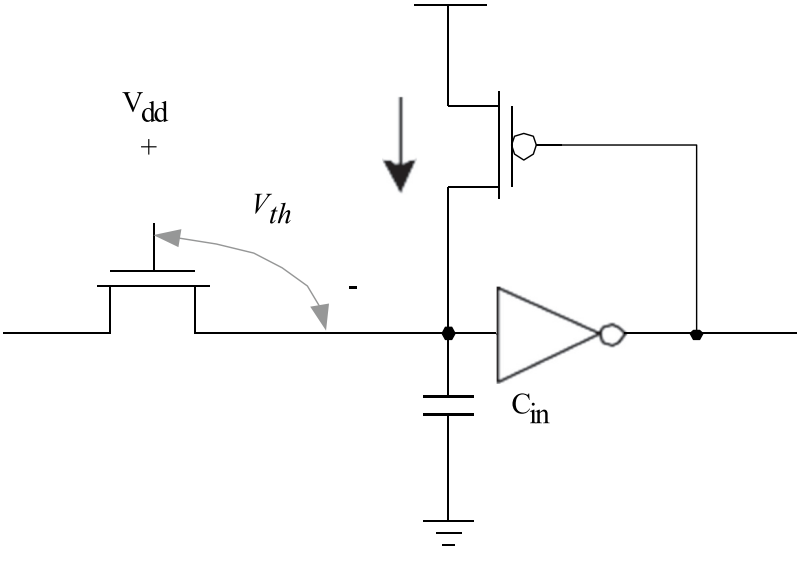
and

$$\boldsymbol{Io = \mu_o C_{ox}(W/L)V^2th} \tag{2}$$

Equation (1) indicates that at smaller technology nodes, the drain-to-source voltage (Vds) and threshold voltage (Vth) of the device decrease. However, Equation (1) also reveals that the sub-threshold current exhibits exponential growth as the threshold voltage (Vth) decreases. Furthermore, an increase in drain-source voltage (Vds) results in a reduction in the threshold voltage (Vth) of the MOS device (Wang et al., 2007). To overcome these challenges, swing restoration emerges as a promising technique to raise the voltages at the node terminals of a MOS transistor, effectively minimising leakage current. Figure 2(a, b). illustrate the standard and swing restoration inverter (SRI) circuits, respectively, with the W/L ratios of transistors specified for the 90 nm technology. The unit sizes are expressed in nanometres. The swing restoration inverter (SRI) circuit in Figure 2(b). consists of two main components: a CMOS inverter output and a pass-transistor logic output made up of CMOS inverters. Additionally, there is a swing-restoring node composed of a single CMOS inverter connected to the gate input of PMOS logic. There are two types of circuit inputs: pass variables connected to the drain of the logic transistor and control variables attached to the transistor's gate. The SRI demonstrates the ability to enhance static noise margins in both read and write operations (Rawat & Kumar, 2021). Node dependency is illustrated in Figure 3. showcasing the impact of Q voltage on the width of P2. As the voltage at Q increases, the width of P2 also increases, ensuring that Q is charged by P2. This exclusive charging maintains Vgs more negative. The choice of P2 width is crucial, ranging between 50% and 135% of N2, to prevent the leakage path from shifting to P2, N1, Gnd, and to ensure optimal performance (preventing Q from exceeding Vdd/2 and ensuring a full swing output voltage). While widening P2 can further reduce leakage, it is essential to keep Q at a lower potential than the output node. For all simulations in this study, the Cadence simulator and 90 nm technology were employed.




(a) Conventional Inverter (Krishna & Duraiswamy, 2021)




(b) Swing Restoration Inverter (SRI)

Figure 2. Circuit schematics. W/L is in nanometer.

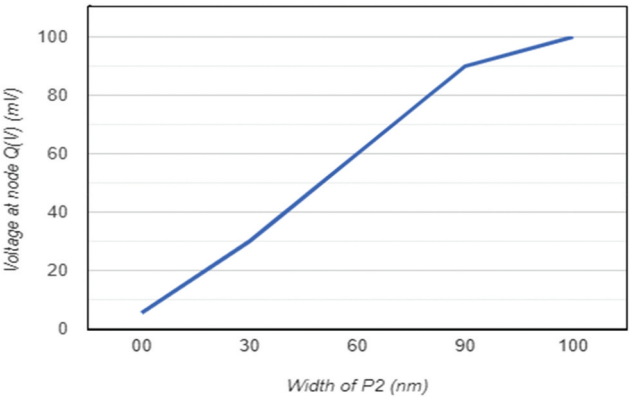



Figure 3. Dependency of node (Q) voltage on width of P2.

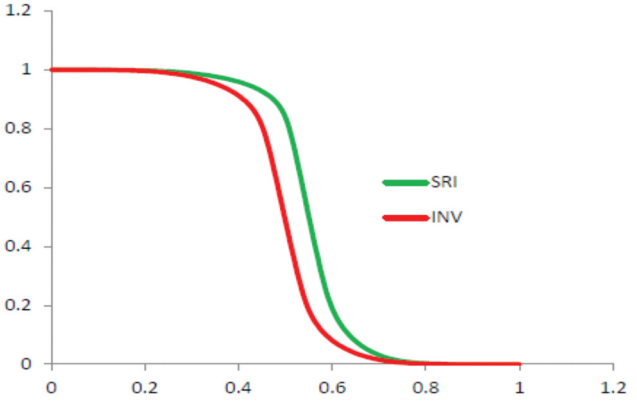



Figure 4. Inverter characteristics.

The simulations were conducted at 27°C for the 90 nm technology. The supply voltage utilised is 1 V for 90 nm technology and 0.9 V for 16 nm.

## 2.1. Transfer characteristics

To examine the transfer characteristics of the two inverters, a DC analysis is performed at a temperature of 27°C. As depicted in Figure 4. the transfer characteristics of swing restoration inverters (SRI) and conventional inverters (INV) for a 90 nm technology are illustrated. The circuit of the swing restoration inverter (SRI) is analogous to the conventional inverter but exhibits robust voltage transfer characteristics.

## 2.2. Leakage power

The power consumed by a circuit in a standby state is commonly referred to as leakage power. To assess leakage power, the inverter's input is set to logic '1' or logic '0'. Next, the

Table 1. Comparison of dynamic power.

| | Dynamic power (nW) | |
|---|---|---|
| Tech (nm) | Conventional inverter (nW) | Swing Restoration inverter (nW) |
| 32 | 27.17 | 11.30 |
| 90 | 25.37 | 9.30 |

leakage power is determined by measuring the current passing through the capacitive load (CL) and using Equation (3). In this study, the specified value for CL is 0.01 pF.

$$\mathbf{P}_{leakage} = \mathbf{I}_{leakage} \times \mathbf{V}_{DD} \tag{3}$$

When a logic '0' is applied to the input of the two inverters, the analysis of leakage power indicates values of 27.66 pW for the standard inverter and 9.37 pW for the Swing Restoration Inverter (SRI) in the 90 nm technology. The exploration of leakage power trends highlights that Swing Restoration Inverters (SRI) demonstrate lower leakage power compared to traditional inverters in the 90 nm technology. Specifically, at the 90 nm technology, the Swing Restoration Inverter (SRI) achieves a 66% reduction in leakage power.

## 2.3. Dynamic power

The power consumed by the circuit during its operational state is known as dynamic power. The measurement of dynamic power for the inverters becomes significant as different strategies for reducing leakage power can impact dynamic power. The calculation of dynamic power is performed using Equation (4).

$$\mathbf{Pdynamic} = \mathbf{V^2DD} \times \mathbf{fin} \times \mathbf{CL} \tag{4}$$

The frequency of the input signal is represented by fin. In this scenario, Proposed Design (PD) is evaluated by measuring the current passing through CL. The circuit is supplied with a 1 MHz input signal and a capacitive load (CL) of 0.1 pF. Table 1. presents the identified dynamic power. Comparative analysis with a conventional inverter reveals that the Swing Restoration Inverter (SRI) effectively lowers dynamic power.

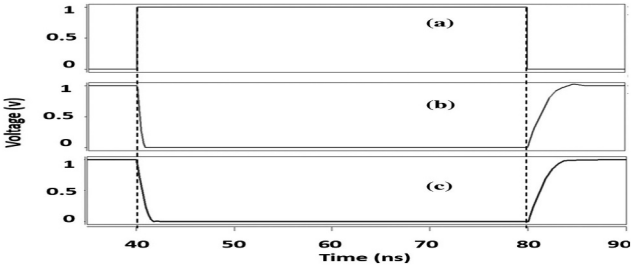



Figure 5. Time domain responses at 90 nm technology. (a) Inverter input. (b) Output of conventional inverter. (c) Output of swing restoration inverter (SRI).

**Table 2.** Time domain response.

| Time domain response | Rise time ($t_r$) (ns) | Fall time ($t_f$) (ns) | Propagation delay Low-high ($t_{plh}$) (ns) | Propagation delay High-low ($t_{phl}$) (ns) |
|---|---|---|---|---|
| Conventional inverter | 1.177 | 0.250 | 0.737 | 0.1561 |
| Swing restoration inverter | 1.58 | 0.503 | 0.732 | 0.1592 |

**Table 3.** Comparison of delay and power delay product (PDP).

| | Conventional inverter | | | Swing restoration inverter (SRI) | | |
|---|---|---|---|---|---|---|
| Tech (nm) | Delay (ns) | P (average) (W) | PDP (J) | Delay (ns) | P (average) (W) | PDP (J) |
| 32 | 0.7715 | 2.766 $e^{-11}$ | 213.39 $e^{-22}$ | 0.9167 | 0.9065 e-11 | 83.09 $e^{-22}$ |
| 90 | 0.6161 | 8.001 $e^{-11}$ | 600.1 $e^{-22}$ | 0.9211 | 0.9121 $e^{-11}$ | 84.03 $e^{-22}$ |

**Table 4.** Comparison of 6T and 8T SRAM leakage power (nW).

| 6T SRAM | 8T SRAM |
|---|---|
| 11.6 | 1.069 |
| 15.08 | 1.999 |
| 21.05 | 2.766 |
| 27.01 | 3.799 |

**Table 5.** Characterisation of SRAM cells.

| Conventional 6T SRAM Cell | | Proposed 8T SRAM Cell | |
|---|---|---|---|
| RSNM | WSNM | RSNM | WSNM |
| 6.37 mV | 240 mV | 12 mV | 360 mV |

**Table 6.** Read word line voltage drop.

| | Conventional 6T SRAM Cell | | Proposed 8T SRAM Cell | |
|---|---|---|---|---|
| Read word line voltage (V) | RSNM (mV) | % Improvement | RSNM (mV) | % Improvement |
| 1 | 100 | - | 76 | - |
| 0.8 | 156 | 56 | 141 | 61 |
| 0.6 | 183 | 83 | 162 | 82 |
| 0.4 | 197 | 97 | 175 | 95 |

**Table 7.** Write word line voltage boost.

| | Conventional 6T SRAM Cell | | Proposed 8T SRAM Cell | |
|---|---|---|---|---|
| Write word line voltage (V) | WSNM (mV) | % Improvement | WSNM (mV) | % Improvement |
| 1 | 192 | – | 239 | – |
| 1.1 | 255 | 30 | 303 | 30 |
| 1.2 | 257 | 31 | 308 | 31 |
| 1.4 | 261 | 32 | 314 | 34 |

### *2.4. Power delay product (PDP)*

The Power-Delay Product (PDP) serves as a crucial metric for assessing the power efficiency of a CMOS circuit. It is computed by multiplying the average power dissipation with the propagation delay. In the context of low-power design, the

objective is to minimise the Power-Delay Product (PDP). Figure 5. visually represents the time domain response of two inverters designed for the 90 nm technology. Meanwhile, Table 2. provides a concise summary of various time domain characteristics observed. Table 3. furnishes data on delay, average power dissipation, and PDP for both inverters at 27°C across different technological nodes. It is worth noting that the Swing Restoration Inverter (SRI) exhibits a longer delay compared to a traditional inverter. However, the average power dissipation of the Swing Restoration Inverter (SRI) is lower than that of the conventional inverter. Consequently, this leads to a reduced Power-Delay Product (PDP) for the Swing Restoration Inverter (SRI).

## 3. Conventional 6T SRAM cell versus proposed 8T SRAM cell design

In a conventional 6T SRAM cell, during read operations, the bit-line access transistors provide direct access to the data storage nodes, as illustrated in Figure 6. During reading, the voltages between the cross-coupled inverter pair and the bit lines are disrupted. Bit lines (BL and BLB) and word lines (WL) are distinct. The word line (WL) controls the access transistors that perform read and write operations. Bit lines act as both input and output nodes. During a read operation, bit lines move data from SRAM cells to a sensing amplifier. The transistors are lowest length of technology 90 nm technology. The swing restoration technique emerges as a promising strategy to mitigate leakage power in CMOS circuits, offering a potential alternative to the traditional use of inverters in SRAM cells. In this investigation, we assess leakage power by substituting the inverter in a 6T SRAM cell with a swing restoration inverter (SRI). To enhance the stability of the SRAM cell, a comprehensive analysis of the cell's stability is conducted, incorporating appropriate assist techniques. Figure 7. illustrates a typical 6T SRAM cell, consisting of two inverters (M1, M5, M2, M6) and access transistors (M3, M4). Dimensions are specified in nanometres, and transistor sizes are denoted based on 90 nm technology. There are two access

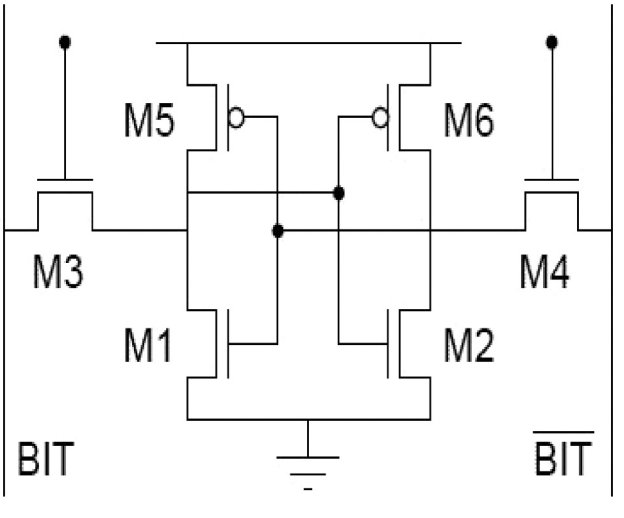



Figure 6. Conventional 6T SRAM cell. W/L is in nanometers.

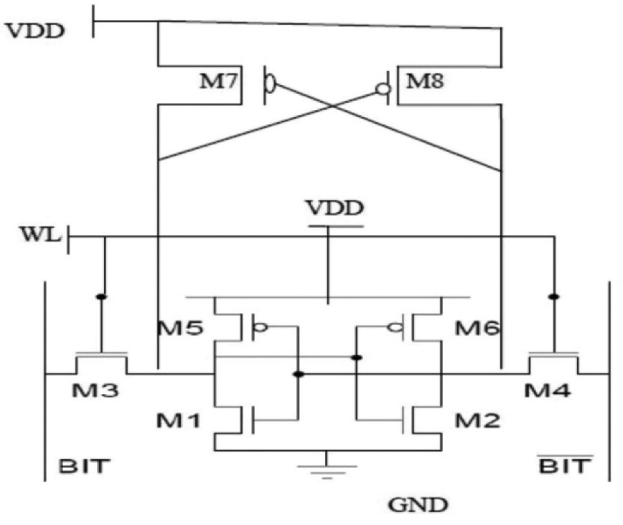



Figure 7. Proposed 8T SRAM cell with swing restoration inverter. W/L is in nanometers.

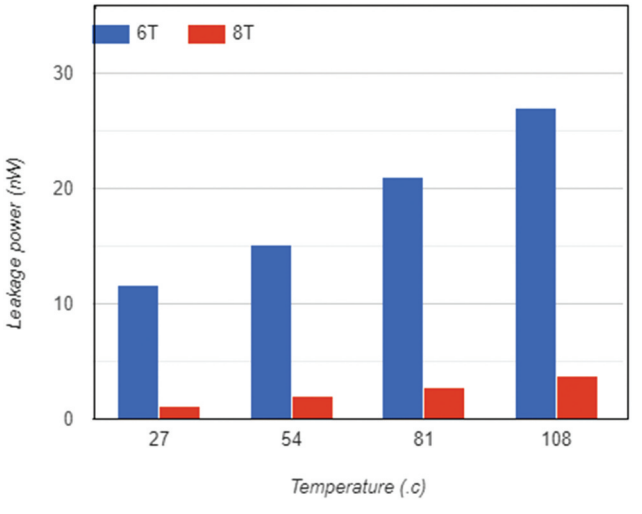
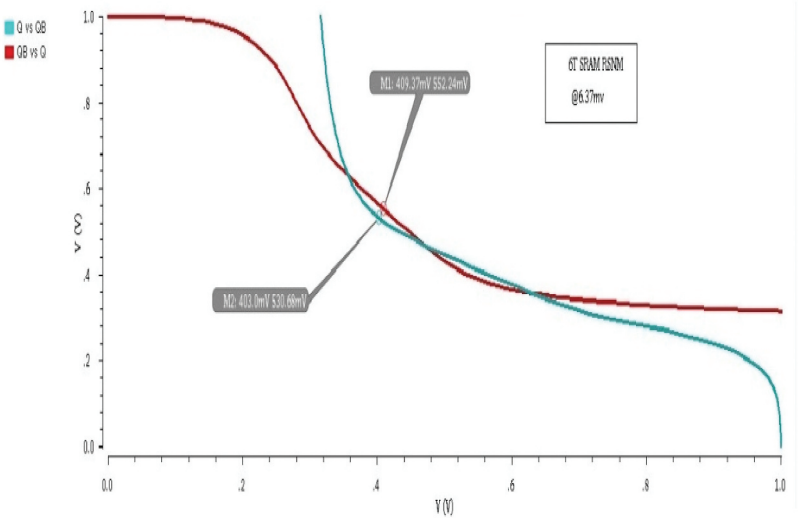


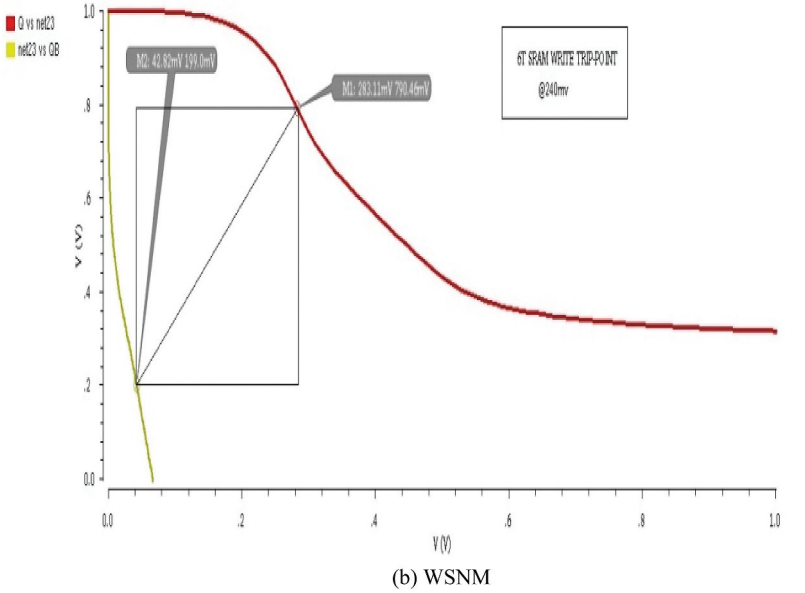

Figure 8. Temperature dependency of leakage power.

transistors (M3, M4) that are switched ON during the active mode. Transistors M7 and M8 act as swing-restored transistors by lowering swing voltages. These transistors are turned off in standby mode by bringing the word line (WL) voltage down to zero. It is easier to identify the sub-threshold leakage current component in the 6T SRAM when the WL lines are set to zero.

Figure 8. depicts a novel SRAM design integrating swing restoration inverters (SRI) with the goal of minimising leakage. Instead of the two inverters found in conventional 6T SRAMs, swing restoration inverters are utilised, resulting in a proposed 8T SRAM configuration. To mitigate sub-threshold current in inverter1, the source voltage of NMOS transistor M3 is elevated, primarily determined by the leakage current supplied by transistor M1. In




(a) RSNM




(b) WSNM

Figure 9. Characterisation of conventional 6T SRAM cell.

inverter2, the reduction in sub-threshold current depends on the source voltage of PMOS transistor M6, influenced by the leakage current from NMOS transistor M2, for a successful write operation, we should have M4 greater than M6. When setting the word lines (WL) to zero, the sub-threshold leakage current component in both conventional and low-leakage SRAMs can be identified. Cadence simulation is employed to compare the leakage power of the two 90 nm technology. Figure 8 shows a comparison of Conventional 6T SRAM Cell and Proposed SRAM Cell leakage power (nW) values. (Variation of Leakage Power with Temperature).

Figure 9. illustrates the temperature-dependent SRAM cell leakage power, demonstrating an increase in leaking power as the temperature rises. The proposed 8T SRAM cell has a notably lower leakage power than the 6T SRAM cell. The 6T SRAM cell has a maximum leakage power of 27.03 nW at 27°C, whereas the proposed 8T SRAM cell has a maximum leakage power of 3.719 nW at the same temperature. This results in leakage power savings of 90% at 27°C and 86% at 100°C for the 8T SRAM cell (Table 4).

## 4. Proposed 8T SRAM cell mode of operations

Figure 7 shows the proposed 8T SRAM Cell mode of operations, which uses dual node voltages with swing restoration logic to perform node voltages in read, write and hold mode operations, and other factors like as stability. A comparison has been made between conventional 6T SRAM and the proposed 8T SRAM Cells(Table 5). The comparative results reveal that the proposed 8T SRAM cell outperforms the conventional 6T SRAM cell in read, write, and hold modes. This is due to the increased noise margin, which ensures acceptable writeability for the bit-cell.

## 5. Improved stability

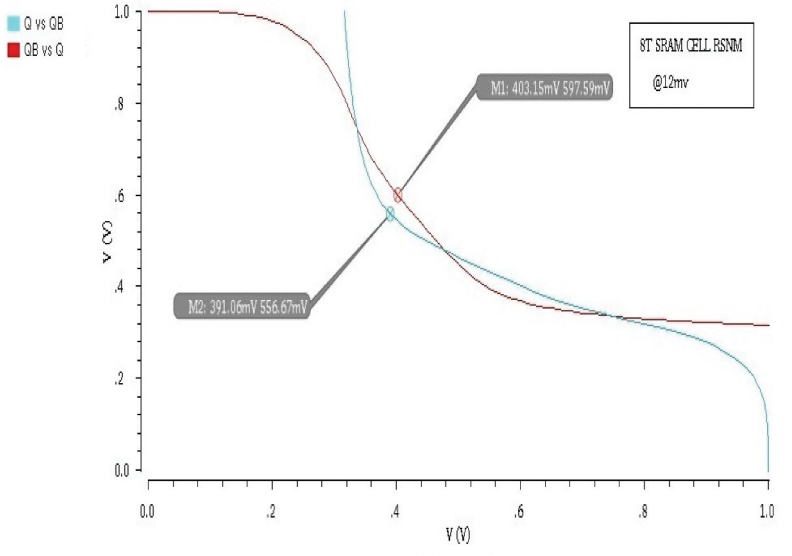
Read assist methods aim to enhance the stability of reading operations, while write assist approaches focus on stabilising write operations. In this study, we employ techniques like boosting the voltage of write word lines and lowering the voltage of read word lines to improve the ability of writing (WSNM) and the stability of reading (RSNM), respectively. Our investigation into stability enhancement encompasses both traditional SRAM and proposed 8T SRAM cell in the 90 nm technology (Tables 6 and 7).

### *5.1. Read stability*

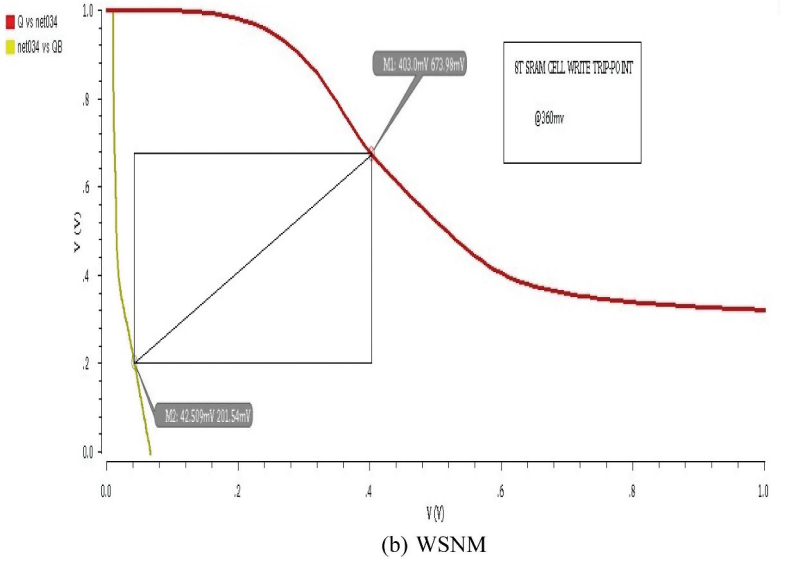
To evaluate read stability in SRAM cells, one uses the read static noise margin (RSNM). Because the storage nodes are separated, the proposed 8T SRAM cell has a higher RSNM than a conventional 6T SRAM cell. Figures 9(a), 10(a). represents the read stability of 6T and proposed 8T SRAM Cells respectively.

### *5.2. Write stability*

Write static noise margin (WSNM) is used to assess write stability in SRAM cells. In the proposed 8T SRAM cell due to storing nodes isolation we get better WSNM comparable to conventional 6T SRAM Cell. Figures 9(b), 10(b) represents the write stability of 6T SRAM and proposed 8T SRAM cells respectively. The write trip point (WTP) is a measure of a cell's writing capabilities. It illustrates the difficulties of the cell's storage nodes. The bit-line voltage is swept from 0 to Vdd, and the flipping of the cell, when Q and Q bar flip their content is captured. The bit-line voltage value at the crossing point of internal storage nodes Q and QB bars represents the write trip point.




(a) RSNM




(b) WSNM

Figure 10. Characterisation of proposed 8T SRAM cell.

### 5.3. Hold stability

Hold Static noise margin (HSNM) is the most common approach to measure hold stability of the cell. The most popular method for measuring the cell's hold stability is static noise margin (SNM). When an SRAM cell is in the hold mode, its stability is assessed. The word lines are turned off in the hold state, thus the cell is completely isolated from the bit lines. SNM determines the maximum amount of noise that may be applied to storage nodes prior to flipping the Cell's content.

Table 8. Comparison of various SRAM cells design techniques.

| Papers | Tech. (nm) | SRAM cells | Techniques | Temp. (°C) | Leakage power (nW) | Leakage power savings (%) | Tools |
|---|---|---|---|---|---|---|---|
| (Kavitha & Govindaraj, 2016) | 32 nm | 6T | Low power multimodal switch power gating | – | 23 nW | 91% | HSPICE |
| (Venkatareddy et al., 2016) | 28 nm | 10T | Dual power supply-based design | 25 | – | 85% | Measurement results |
| (Krishna & Duraiswamy, 2021) | 32 nm | 10T | Source biasing technique | 25 | 1.059 nW | 90.88% | HSPICE |
| (Prasad & Anand, 2015) | 90 nm | 8T | Supply voltage scaling | 27 | 44.85 nW | 40.40% | Cadence |
| (Sharma et al., 2016) | 32 nm | 9T | FinFET based design | 25 | 16.9 nW | 53.21% | HSPICE |
| (Venkatareddy et al., 2016) | 60 nm | 7T | Super cut-off word line | – | 21.85 nW | 26% | – |
| (Mani et al., 2022) | 90 nm | Conventional 6T SRAM Cell(reference) | Conventional inverter | 27 | 5.62 nW | 45.44% | Cadence |
| This work | 90 nm | Proposed 8T SRAM Cell | Swing restoration inverter | 27 | 1.013 nW | 91.22% | Cadence |

## 6. Results and discussions

Simulations were conducted using the Cadence Virtuoso tool with the 90 nm technology. Throughout the process, the supply voltage was set at 1 V, and PTM models were utilised while maintaining a temperature of 27°C. The margin of swing restoration inverter (SRI) exhibited a significant increase for both SRAM configurations, reaching 30% and 32%, respectively, for the traditional 6T and 8T SRAM. The suggested method demonstrates its efficacy in reducing power leakage in deep sub-micron technologies. An examination of the swing restoration inverter through DC analysis highlights its advantageous voltage transfer characteristics (VTCs).

The provided information highlights the effectiveness of Swing Restoration Inverters (SRIs) in mitigating leakage power consumption, presenting distinct advantages over standard inverters. At 90 nm technology and 27°C, SRIs exhibit a significant 67% reduction in leakage power, which decreases to 39% at 22 nm and 35% at 16 nm with advancements in technology. Additionally, the swing restoration approach plays a crucial role in minimising dynamic power usage, consistently resulting in a lower Power-Delay Product (PDP) compared to conventional inverters.

Table 8. explores various strategies for minimising leakage power in SRAM cells, accounting for differences in researchers' methodologies, simulation parameters, and technology variations. Notably, with 90 nm technology, the proposed approach achieves the lowest power consumption at 1.013 nW and 2.719 nW at 27°C and 30°C, respectively, demonstrating substantial savings of 91% and 86% in leakage power. The versatility of the swing restoration approach extends to different conventional SRAM cells utilising cross-coupled inverters. Comparative analysis of Read Static Noise Margin (RSNM) and Write Static Noise Margin (WSNM) values between Conventional 6T SRAM and Proposed 8T SRAM cells reveals their similarity. To evaluate stability enhancement, both types of SRAM cells undergo write and

read stability mode operations. The Proposed 8T SRAM cell consistently outperforms the Conventional 6T SRAM cell in terms of stability.

## 7. Conclusion

In the context of designing SRAMs for deep sub-micron technologies, a novel low-leakage inverter based on a swing restoration approach is introduced. This innovative design incorporates two additional transistors are used to node restoration to minimise leakage power. The presence of these additional transistors elevates node voltages while concurrently reducing threshold voltages (Vth). The resulting swing restoration inverter (SRI) exhibits lower leakage power compared to Conventional inverters. At a temperature of 27°C, the SRI achieves significant leakage power savings, amounting to 67% for the 90 nm technology and 57% for the 32 nm technology, respectively. However, it is noted that as technology nodes decrease, the leakage power savings of the suggested inverter diminish. Furthermore, it is demonstrated that the dynamic power consumption is reduced in both 90 nm and 32 nm technologies when utilising the swing restoration inverter (SRI). The implementation of SRIs in place of standard inverters in a 6T SRAM design, with the low-leakage inverter employed for SRAM design, showcases noteworthy results. Comparing the resulting Proposed 8T SRAM cell to a conventional 6T SRAM at 90 nm technology, there is a substantial reduction in leakage power 91% at 30°C and 93% at 27°C, respectively. Importantly, the proposed 8T SRAM cell maintains low leakage while being equivalent to the 6T SRAM cell in terms of stability. Moreover, the Proposed 8T SRAM cell, when incorporating swing restoration techniques, demonstrates enhanced stability compared to Conventional designs. It is important to note that the Proposed 8T SRAM cell integrates two additional transistors compared to the Conventional 6T SRAM cell, a trade-off made to effectively reduce leakage power and enhance stability. In conclusion, while working at 90 nm technology node has its limitations, it was a relevant choice for the research given its widespread adoption at the time. However, it is essential to acknowledge these limitations and to consider how to extend and adapt the proposed techniques to newer technology nodes in future work.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## ORCID

Ram Murti Rawat http://orcid.org/0000-0001-7280-8779

## References

Bai, N., Hu, Z., Wang, Y., & Xu, Y. (2022). Leakage current stability analysis for subthreshold SRAM. *Electronics*, *11*(8), Article 1196. https://doi.org/10.3390/electronics11081196

Benmoussa, Y., Boukhobza, J., Senn, E., & Benazzouz, D. (2013). Energy consumption modeling of H.264/AVC video decoding for GPP and DSP. In *Proceedings of the 16th Euromicro Conference on Digital System Design (DSD)* (pp. 890–897). https://doi.org/10.1109/DSD.2013.100

Carroll, A., & Heiser, G. (2010). An analysis of power consumption in a smartphone. In *Proceedings of USENIX Annual Technical Conference* (pp. 1–14). https://www.usenix.org/legacy/events/usenix10/tech/full_papers/Carroll.pdf

Chen, J., Zhao, W., Wang, Y., Shu, Y., Jiang, W., & 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), 769–780. https://doi.org/10.1109/TVLSI.2022.3164756

Chen, Y. H., Chou, S. Y., Li, Q., Chan, W. M., Sun, D., Liao, H. J., Wang, P., Chang, M. F., & Yamauchi, H. (2012). Compact measurement schemes for bit-line swing sense amplifier offset voltage and word-line pulse width to characterize sensing tolerance margin in a 40 nm fully functional embedded SRAM. *IEEE Journal of Solid-State Circuits*, *47*(4), 969–980. https://doi.org/10.1109/JSSC.2012.2185180

Hamzaoglu, F., Ye, Y., Keshavarzi, A., Zhang, K., Narendra, S., Borkar, S., Stan, M., & De, V. (2002). Analysis of dual-sram cells with full-swing single-ended bit line sensing for on-chip cache. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, *10*(2), 196–205. https://doi.org/10.1109/TVLSI.2002.994983

Hanchate, N., & Ranganathan, N. (2004). LECTOR: A technique for leakage reduction in CMOS circuits. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, *12*(2), 196–205. https://doi.org/10.1109/TVLSI.2003.821547

Hirabayashi, O., Kawasumi, A., Suzuki, A., Takeyama, Y., Kushida, K., Sasaki, T., & Yabe, T. (2009). A process-variation-tolerant dual-power-supply SRAM with 0.179 μm2 cell in 40nm CMOS using level-programmable wordline driver. In *IEEE International Solid-State Circuits Conference Digest of Technical Papers* (pp. 458–459). https://doi.org/10.1109/ISSCC.2009.4977506

Hoque, M. A., Siekkinen, M., & Nurminen, J. K. (2014). Energy efficient multimedia streaming to mobile devices-A survey. *IEEE Communications Surveys & Tutorials*, *16*(1), 579–597. https://doi.org/10.1109/SURV.2012.111412.00051

Joy, A., & Kuruvilla, J. (2022). A stable low power dissipating 9 T SRAM for implementation of 4 × 4 memory array with high frequency analysis. *Wireless Personal Communications*, *126*(4), 3305–3316. https://doi.org/10.1007/s11277-022-09865-x

Kang, K., Jeong, H., Yang, Y., Park, J., Kim, K., & Jung, S. O. (2016). Full-swing local bitline SRAM architecture based on the 22-nm FinFET technology for low-voltage operation. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, *24*(4), 1342–1350. https://doi.org/10.1109/TVLSI.2015.2450500

Kavitha, M., & Govindaraj, T. (2016). Low-power multimodal switch for leakage reduction and stability improvement in SRAM cell. *Arabian Journal for Science & Engineering*, *41*(8), 2945–2955. https://doi.org/10.1007/s13369-016-2047-0

Krishna, R., & Duraiswamy, P. (2020a). Simulation study and performance comparison of various SRAM cells in 32 nm CMOS technology. In *Emerging trends in photonics signal processing and communication engineering* (Vol. 649, pp. 51–60). Springer. https://doi.org/10.1007/978-981-15-3477-5_7

Krishna, R., & Duraiswamy, P. (2020b). A technique of designing low leakage SRAM in deep sub-micron technology. In *IEEE International Conference on Electronics Computing and Communication Technologies (CONECCT)* (pp. 1–5). https://doi.org/10.1109/CONECCT50063.2020.9198587

Krishna, R., & Duraiswamy, P. (2021). Low leakage 10T SRAM cell with improved data stability in deep sub-micron technologies. *Analog Integrated Circuits and Signal Processing*, *109*(1), 153–163. https://doi.org/10.1007/s10470-021-01870-7

Liu, B., Cai, J., Yuan, J., & Hei, Y. (2017). A low-voltage SRAM sense amplifier with offset cancelling using digitized multiple body biasing. *IEEE Transactions on Circuits & Systems II: Express Briefs*, *64*(4), 442–446. https://doi.org/10.1109/TCSII.2016.2563660

Mani, E., Abbasian, E., Gunasegeran, M., & Sofimowloodi, S. (2022). Design of high stability low power and high speed 12 T SRAM cell in 32-nm CNTFET technology. *AEU - International Journal of Electronics & Communications*, *154*, Article 154308. https://doi.org/10.1016/j.aeue.2022.154308

Miyano, S., Moriwaki, S., Yamamoto, Y., Kawasumi, A., Suzuki, T., Sakurai, T., & Shinohara, H. (2013). Highly energy-efficient SRAM with hierarchical bit line charge-sharing method using non-selected bit line charges. *IEEE Journal of Solid-State Circuits*, *48*(4), 924–931. https://doi.org/10.1109/JSSC.2012.2237572

Mukku, P. K., & Lorenzo, R. (2023). Design and analysis of radiation hardened 10 T SRAM cell for space and terrestrial applications. *Microsystem Technologies*, *29*(10), 1489–1500. https://doi.org/10.1007/s00542-023-05500-2

Mutoh, I., Douseki, T., Matsuya, Y., Aoki, T., Shigematsu, S., & Yamada, J. (1995). 1-V power supply high-speed digital circuit technology with multithreshold-voltage CMOS. *IEEE Journal of Solid-State Circuits*, *30*(8), 847–854. https://doi.org/10.1109/4.400426

Nandyala, V. R., & Mahapatra, K. K. (2016). A circuit technique for leakage power reduction in CMOS VLSI circuits. In *International Conference on VLSI Systems Architectures Technology and Applications (VLSI-SATA)* (pp. 1–5). https://doi.org/10.1109/VLSI-SATA.2016.7593044

Narendra, S., Borkar, S., De, V., Antoniadis, D., & Chandrakasan, A. (2002). Scaling of stack effect and its application for leakage reduction. In *Proceedings of Low Power Electronics and Design (LPE)* (pp. 195–200). https://doi.org/10.1109/lpe.2001.945400

Nii, K., Yabuuchi, M., Tsukamoto, Y., Ohbayashi, S., Imaoka, S., Makino, H., Yamagami, Y., Ishikura, S., Terano, T., Oashi, T., Hashimoto, K., Sebe, A., Okazaki, G., Satomi, K., Akamatsu, H., & Shinohara, H. (2008). A 45-nm bulk CMOS embedded SRAM with improved immunity against process and temperature variations. *IEEE Journal of Solid-State Circuits*, *43*(1), 180–191. https://doi.org/10.1109/JSSC.2007.907998

Prasad, G., & Anand, A. (2015). Statistical analysis of low-power SRAM cell structure. *Analog Integrated Circuits and Signal Processing*, *82*(1), 349–358. https://doi.org/10.1007/s10470-014-0463-1

Priya, M. G., Baskaran, K., & Krishnaveni, D. (2012). Leakage power reduction techniques in deep submicron technologies for VLSI applications. *Procedia Engineering*, *30*, 1163–1170. https://doi.org/10.1016/j.proeng.2012.01.976

Rawat, R. M., & Kumar, V. (2021). Low power pre-charge voltage level and low swing logic based 8T SRAM cell for high speed CMOS circuits. *2021 IEEE 32nd International Conference on Microelectronics (MIEL)* (pp. 243–246). https://doi.org/10.1109/MIEL52794.2021.9569198

Sachan, D., Peta, H., Malik, K. S., & Goswami, M. (2016). Low power multi threshold 7T SRAM cell. In *IEEE 59th International Midwest Symposium on Circuits and Systems (MWSCAS)* (pp. 1–4). https://doi.org/10.1109/MWSCAS.2016.7870004

Sayyah, E. S., Moaiyeri, M. H., Ebrahimi, B., Hessabi, S., & Afzali-Kusha, A. (2019). A low-leakage and high-writable SRAM cell with back-gate biasing in FinFET technology. *Journal of Computational Electronics*, *18*(1), 7–19. https://doi.org/10.1007/s10825-019-01327-1

Sharma, N., Panwar, U., & Singh, V. (2016). A novel technique of leakage power reduction in 9T SRAM design in FinFET technology. In *Proceedings of the 6th International Conference - Cloud System and Big Data Engineering (Confluence)* (pp. 737–743). https://doi.org/10.1109/CONFLUENCE.2016.7508217

Shibata, N., Kiya, H., Kurita, S., Okamoto, H., Tan'no, M., & Douseki, T. (2006). A 0.5-V 25-MHz 1-mW 256-kb MTCMOS/SOI SRAM for solar-power-operated portable personal digital equipment-sure write operation by using step-down negatively overdriven bitline scheme. *IEEE Journal of Solid-State Circuits*, *41*(3), 728–741. https://doi.org/10.1109/JSSC.2005.864124

Subramanyam, J. B. V., & Basha, S. S. (2016). Design of low leakage power SRAM using multithreshold technique. In *2016 10th International Conference on Intelligent Systems and Control (ISCO)* (pp. 1–7). https://doi.org/10.1109/ISCO.2016.7726929

Suneja, D., Chaturvedi, N., & Gurunarayanan, S. (2017). A comparative analysis of read/write assist techniques on performance & margin in 6T SRAM cell design. In *International Conference on Computer Communications and Electronics (Comptelix)* (pp. 659–664). https://doi.org/10.1109/COMPTELIX.2017.8004051

Suzuki, T., Yamauchi, H., Yamagami, Y., Satomi, K., & Akamatsu, H. (2008). A stable 2-port SRAM cell design against simultaneously read/write-disturbed accesses. *IEEE Journal of Solid-State Circuits*, *43*(9), 2109–2119. https://doi.org/10.1109/JSSC.2008.2001872

Tachibana, F., Hirabayashi, O., Takeyama, Y., Shizuno, M., Kawasumi, A., Kushida, K., Suzuki, A., Niki, Y., Sasaki, S., Yabe, T., & Unekawa, Y. (2014). A 27% active and 85% standby power reduction in dual-power-supply SRAM using BL power calculator and digitally controllable retention circuit. *IEEE Journal of Solid-State Circuits*, *49*(1), 118–126. https://doi.org/10.1109/JSSC.2013.2280312

Venkatareddy, A., Sithara, R., Kumar, Y. B. N., & Vasantha, M. H. (2016). Characterization of a novel low leakage power and area efficient 7T SRAM cell. In *Proceedings of the IEEE International Conference on VLSI Design* (pp. 202–206). https://doi.org/10.1109/VLSID.2016.28

Wang, D. P., Liao, H. J., Yamauchi, H., Chen, Y. H., Lin, Y. L., Lin, S. H., Liu, D. C., Chang, H. C., & Hwang, W. (2007). A 45nm dual-port SRAM with write and read capability enhancement at low voltage. In *Proceedings of the 20th Anniversary IEEE International SOC Conference* (pp. 211–214). https://doi.org/10.1109/SOCC.2007.4545460

Wang, J. S., Chang, P. Y., Tang, T. S., Chen, J. W., & Guo, J. I. (2011). Design of subthreshold SRAMs for energy-efficient quality-scalable video applications. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, *1*(2), 183–192. https://doi.org/10.1109/JETCAS.2011.2158345

Wang, W., Singh, A. D., & Guin, U. (2022). A systematic bit selection method for robust SRAM PUFs. *Journal of Electronic Testing: Theory and Applications (JETTA)*, *38*(3), 235–246. https://doi.org/10.1007/s10836-022-06006-x

Plasmonics
https://doi.org/10.1007/s11468-024-02713-7

RESEARCH

# Design and Analysis of Twin-Core Side-Polished SPR PCF Sensor for Preliminary Malaria Detection

**Devi Kangan B S[1] · Akash Khamaru[1] · Ajeet Kumar[1]**

Received: 11 November 2024 / Accepted: 5 December 2024
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2025

## Abstract

This paper proposes a double-side polished twin-core photonic crystal fiber (PCF)-based surface plasmon resonance (SPR) sensor for early-stage detection of malaria. The analysis and the assessment of the proposed sensor are done by finite element method using the COMSOL Multiphysics software. The dual-side polished structure reduces the distance between the metal layer and core mode, thereby enhancing the sensor performance due to the strong coupling between them. The air holes arrangement resembles the shape of an "hourglass" which helps to confine light energy within the twin cores. A layer of gold, a preferred plasmonic material, is coated on the symmetrically polished surface of the PCF to induce SPR. Due to distinction in the refractive indices, the resonance wavelength is different for infected and normal red blood cells (RBCs). The analysis is done for infected RBCs (ring phase, trophozoite phase, and schizont phase). The structural parameters are optimized precisely and a maximum wavelength sensitivity of 26,428.57 nm/RIU was achieved for the ring phase, 16,052.6 nm/RIU for the trophozoite phase, and 12,586.2 nm/RIU for the schizont phase. The numerical analysis also shows that the sensor possess high figure-of-merit (FOM) values of 334.54 $RIU^{-1}$, 321.05 $RIU^{-1}$, and 381.40 $RIU^{-1}$ for the ring, trophozoite, and schizont phases, respectively. These results confirm the potential of the proposed sensor and also justify its betterment than the existing methods. Due to its promising sensing performance and appropriate FOM value, the proposed sensor is suitable for the early-stage malaria detection.

**Keywords** Trophozite phase · Schizont phase · Twin-Core · Wavelength sensitivity · Figure-of-merit (FOM)

## Introduction

Malaria is a highly virulent and potentially fatal illness triggered by a parasite that inhabits certain mosquitoes which bite humans. In the global population, approximately 44% is susceptible to malaria. The disease is caused by five distinct parasite species belonging to the *Plasmodium* genus: *P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale*, and *P. knowlesi*. Among these species, the most dangerous one is *Plasmodium falciparum*, which causes the infection [1, 2]. Thus, the cycle consists of three stages: the ring stage, trophozoite stage, and schizont stage. The key parameter for the diagnosis of malaria is the difference in optical properties of healthy and infected erythrocytes [3, 4]. In 2022, the World Health Organization (WHO) estimated that there were approximately 249 million malaria cases worldwide, with around 608,000 related deaths reported across 85 countries [5]. To prevent the spread of disease in the communities and to provide effective treatment for the infected persons, accurate and timely malaria diagnosis is crucial [6]. The conventional method for diagnosing malaria typically involves clinical assessment based on patient symptoms. But this can cause misdiagnosis because these symptoms can be similar to other common diseases [7]. Another widely used method is the microscopic examination of stained peripheral blood smears using field's stains. However, this method can cause false-positive detection due to the low parasite level, and it needs skilled technicians, and the process is time-consuming [8]. The quantitative buffy coat (QBC) method can detect low-concentrated parasites and detected through an

✉ Ajeet Kumar
ajeetdph@dtu.ac.in

Devi Kangan B S
devikanganBS_23mscphy19@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: 17 January 2025

epifluorescent microscope [9]. Even though it is faster than conventional microscopes, it needs skilled technicians and specialized devices to operate. Rapid diagnostic test (RDT) is a quick and convenient technique for disease detection, offering rapid results and easy handling. The method is especially useful in resource-limited settings, where timely diagnosis is pivotal for effective treatment and control of disease outbreaks. But RDTs do not differentiate between various *Plasmodium* species, which is crucial for making informed treatment decisions [10]. The serological method is another traditional method, which is straightforward and sensitive, though it requires significant time to obtain the results [11].

Recently, various efficient methods for malaria detection are emerging. The polymerase chain reaction (PCR)-based technique is a highly sensitive method capable of detecting low levels of parasites, including various *Plasmodium* species. The method is expensive as it requires specialized expertise and sophisticated equipment [12]. The indirect fluorescent antibody (IFA) test relies on the binding of fluorescently labeled antibodies to specific target antigens within a sample. However, this method requires specialized equipment and skilled technicians [13]. Loop-mediated isothermal amplification (LAMP) is a simple, quick, and cost-effective alternative to PCR for detecting malaria. The technique operates at a constant temperature, thus removing the necessity of thermal cycling and streamlining the overall procedure. However, the high sensitivity of LAMP can also increase the risk of contamination. The possibility of false detection is high without proper precautions [14]. Flow cytometry (FCM) can detect malaria parasites by analyzing labeled cells in a blood sample; however, it has a high initial installation cost and requires ongoing maintenance, with the potential to skip low-level parasites, which could lead to misdiagnosis [15]. The method automated blood cell counter (ACC) analyzes blood samples to provide complete blood counts and can indicate the presence of malaria by assessing red blood cell changes. The method requires regular calibration and maintenance [16]. Therefore, the present situation demands a faster and more efficient method for early-stage malaria detection. The latest and accurate optical techniques are key to produce high-advancing tools which can be available for diagnosis of blood diseases [17].

Photonic crystal fiber (PCF)-based biosensor is one of these optical techniques, which plays a vital role in biosensing. PCF is constituted by a number of air holes drilled in the cross-section which act as a cladding and offer some distinct features other than conventional fiber. Its unique and flexible structural design enables various applications in telecommunications [18], sensing [19, 20], and nonlinear optics [21, 22]. The PCF-based sensors have been extensively used for chemical and biosensing applications due to advancements in manufacturing and sensing technologies, particularly for detecting refractive index (RI) changes [23]. Likewise, sensors utilizing surface plasmon resonance (SPR) have garnered increased attention owing to their remarkable capabilities, such as real-time monitoring, label-free detection, and enhanced sensitivity [24]. In SPR phenomenon, when the angle or wavelength of the incident light changes, it excites the surface electrons leading to the coherent oscillations of the electrons and generate surface plasmon polaritons (SPPs), which will resonate with the incident EM wave. The oscillation of electrons facilitates the energy transfer from EM wave to SPPs, which in turn generates surface plasmon waves (SPWs). A change in the surrounding medium affects the RI, which in turn influence the SPWs, giving maximized sensitivity to such variations. SPR has applications in various fields like medical diagnostics, environmental monitoring, food quality control, and pharmaceutical development [25, 26]. Conventional SPR-based sensors utilized prism-based designs, but it was very bulky and not suitable for distant sensing purposes. To achieve miniaturization, design flexibility, and enhanced sensitivity, SPR has been coupled with PCFs. This combination leverages the strengths of both technologies, expanding their applications. This integration allows the development of compact, portable, and highly sensitive sensors suitable for various applications, including those in challenging environments [27, 28].

The RI is a critical parameter for detecting various diseases like malaria, as healthy and infected RBCs exhibit distinct RI's. Infected RBCs with malaria display a non-uniform RI, which varies across different stages of the infection — ring phase, trophozoite phase, and schizont phase having RIs like 1.395, 1.383, and 1.373 respectively. Normal RBCs have an RI of 1.402. When a sample of malaria-infected RBCs is introduced into the SPR PCF biosensor, a significant shift in resonance wavelength occurs due to the coupling between the SPP mode and the core mode. This shift enables the detection of various stages of malaria infection [29].

The past literatures show that PCF SPR sensors have already been explored for various biosensing applications. Das et al. introduced the SPR-based PCF biosensor for malaria detection in which air hole arrangement resembles a barred spiral galaxy and has computed sensitivity for different stages like ring, schizont, and trophozoite phases of malaria-infected RBCs of 17,857.14 nm/RIU, 10,210.53 nm/RIU, 8758.62 nm/RIU, respectively [29]. Chaudhary et al. introduced circular PCF-based SPR sensor for malaria detection [30]. The design has two layers of air holes arranged in hexagonal lattice. The calculated wavelength sensitivity for ring, schizonts, and trophozoite phases of malaria infection is 13,714.29 nm/RIU, 9789.47 nm/RIU, and 8068.97 nm/RIU, respectively, in $x$-polarized direction and 14,285.71 nm/RIU, 10,000 nm/RIU, and 8206.9 nm/RIU, respectively, in $y$-polarized direction. In the year 2023, Kumar et al. have reported a circular PCF-based SPR for

malaria detection using the non-uniform RI characteristics of unhealthy RBCs. The authors used the machine learning approach for this study, and finally detected the wavelength sensitivity of ring phase, schizont phase, and trophozoite phase like 12,142 nm/RIU, 9736 nm/RIU, and 8241 nm/RIU, respectively [31]. Shafkat et al. reported a simple PCF sensor for malaria detection. The air holes are arranged in double-circular loops in PCF, and the RBC sample for detection is kept in the horizontal channel at the center of the fiber, and the maximum spectral sensitivities observed were approximately 11,428.57 nm/RIU for the ring stage, 9473.68 nm/RIU for the trophozoite stage, and 9655.17 nm/RIU for the schizont stage of the parasite [32]. Apart from malaria sensing, various generalized RI sensors are also proposed. In the year of 2024, Khamaru et al. have reported a graded index SPR-based D-shaped PCF RI sensor which has a teardrop-shaped core which facilitates the coupling of core mode with metal layer and having highest wavelength sensitivity of 21,000 nm $RIU^{-1}$ at a wide analyte RI range [33]. Lv et al. have reported a dual-core D-shaped SPR-based PCF sensor for RI detection in liquids; the resonance wavelength shift varies according to the RI of analyte, and the authors have reported average wavelength sensitivity of 17,200 nm/RIU in the RI range between 1.40 and 1.44 [34].

In this work, we have designed a dual-side polished, twin-core SPR-based PCF biosensor for the early detection of various stages of malaria. The plasmonic material gold is coated on the symmetrical sides to enhance the sensing performance by efficient coupling of light with SPP mode. The triangular-shaped dual cores are infused in the proposed PCF, and this configuration improves sensitivity. The symmetrical layout of dual-side polished sensor simplifies the manufacturing process and allows the better placement of plasmonic material which in turn enhances the coupling with the SPP mode. For ensuring the compactness, the minimum number of air holes is incorporated, having different radii. The finite element method–based COMSOL Multiphysics is used for the designing, investigation, and optimization of the proposed sensor. The findings demonstrate that the discussed sensing probe shows great promise for the early detection of different malaria stages, addressing the limitations of conventional diagnostic techniques.

## Structural and Theoretical Modeling

The three-dimensional (3D) and two-dimensional (2D) geometry of the dual-side polished, twin-core SPR PCF sensor shown in Fig. 1a, b, d, respectively. For the accurate simulation, physics-controlled mesh sequence is used, as displayed in Fig. 1c. Here, the analytical domain is subdivided into triangular mesh. The increased mesh density enhances the light transmission to the sensor, making mathematical calculations easier and more accurate. In the proposed

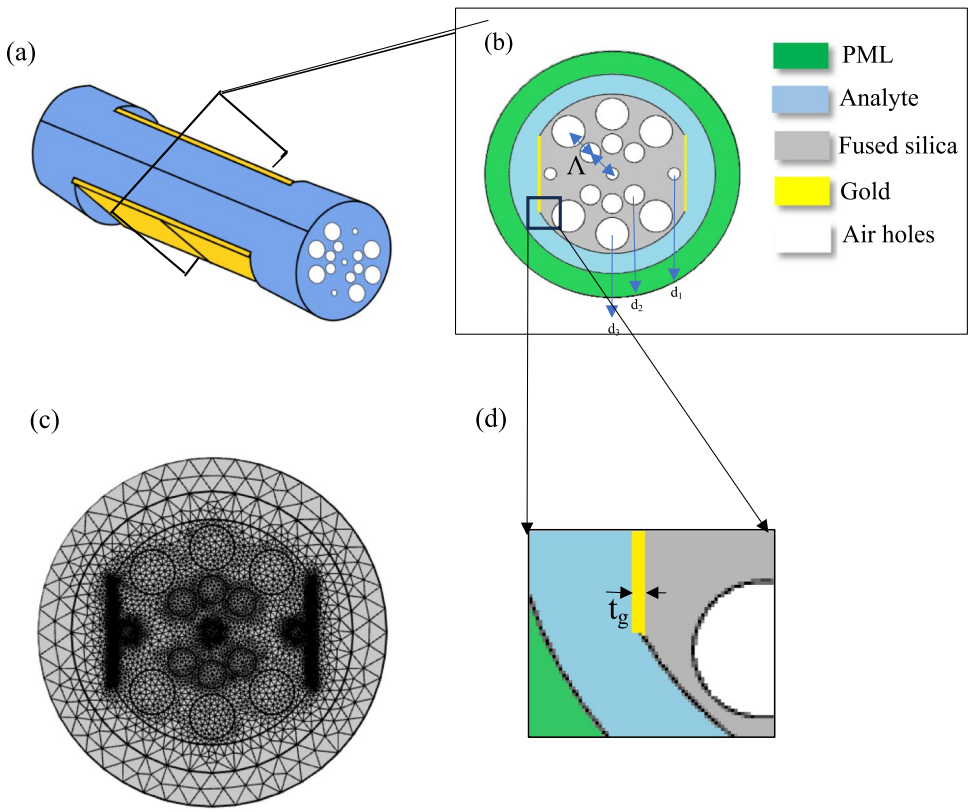



**Fig. 1** **a** 3D view of the sensor. **b** Cross-sectional view of the sensor. **c** Physics-controlled mesh view. **d** Zoomed view of PCF showing plasmonic material (gold)

sensor design, the air hole arrangement resembles the shape of an "hourglass," consisting of three different diameters of air holes, in which two types of air holes are arranged in a circular ring structure and three air holes arranged longitudinally in center, having diameter $d_1$ (0.6 µm). The first and second ring of air holes have diameter $d_2$ (1 µm) and $d_3$ (1.6 µm), respectively. The spacing between adjacent air holes, referred to as the pitch of the structure, is consistently maintained as Λ (1.5 µm). From both circular rings, two air holes are removed purposefully and two air holes of diameter similar to the center one are placed symmetrically at 2Λ distance to create a passage for core light to interact with the plasmonic material. The two symmetrical sides of the cladding are polished carefully to enhance the coupling. The two missing air holes from both the circular structure and the central air hole constitute the dual-core PCF structure. A nanolayer of gold which is an efficient plasmonic material having greater chemical stability, good corrosion resistance, and good biocompatibility is coated above the polished surface with thickness $t_g = 40$ nm [35]. A layer of bio-analyte is deposited on the top of gold. Finally, a perfectly matched layer (PML) which acts as a boundary which absorbs scattering light energy is installed above this bio-analyte layer.

Fused silica is the background material of the sensor, which has high transparency and low loss at NIR region. RI which is dependent upon wavelength is derived from Sellmeier's equations [36]:

$$n_{si} = \sqrt{1 + \frac{(B_1 \times \lambda^2)}{\lambda^2 - C_1} + \frac{(B_2 \times \lambda^2)}{\lambda^2 - C_2} + \frac{(B_3 \times \lambda^2)}{\lambda^2 - C_3}} \tag{1}$$

where $B_1 = 0.69616300$, $B_2 = 0.407942600$, $B_3 = 0.897479400$, $C_1 = 4.67914826\ 10^{-3}$ um$^2$, $C_2 = 1.35120631 \times 10^{-2}$ um$^2$, $C_3 = 97.934002$ um$^2$

The Lorenz-Drude model gives the relative permittivity or dielectric constant of the plasmonic material gold [36], given in the Eq. (2):

$$\varepsilon_{Au} = \varepsilon_\infty - \frac{\omega_D^2}{\omega(\omega + i\gamma_d)} - \frac{\Delta\varepsilon\Omega_L^2}{(\omega^2 - \Omega_L^2) + i\Gamma_L\omega} \tag{2}$$

The values of the parameters used in the above equation are $\varepsilon_\infty = 5.9673$, $\Delta\varepsilon = 1.09$, $\Gamma_L = 16.69$, $\gamma_d = 2.534$, $\Omega_L = 103.46$, $\omega = 2\pi c/\lambda$ where $\varepsilon_\infty$ is the permittivity at high frequency, $\Gamma_L$ is the spectral width of Lorentz oscillator, $\Delta\varepsilon$ is weighted coefficient, $\omega$ is the optical angular frequency, $\gamma_d$ is damping frequency, and $\Omega_L$ represents oscillator strength.

To evaluate the performance of a PCF sensor, the confinement loss of the core mode is a crucial parameter. The confinement loss, denoted as $\alpha_{loss}$, can be calculated using the following Eq. (3) [37]:

$$\alpha_{Loss} = 8.686 \times \frac{2\pi}{\lambda} \times I_m(n_{eff}) \times 10^4 \frac{dB}{cm} \tag{3}$$

where $I_m(n_{eff})$ is the imaginary part of the effective mode index.

In this work, we are going to explore the wavelength sensitivity (WS), which represents the difference between two resonant wavelengths per unit change in refractive index (RI) difference. It can be determined by the following equation [37]:

$$S_\lambda = \frac{\Delta\lambda_{peak}}{\Delta n_a}(\text{nm/RIU}) \tag{4}$$

Amplitude sensitivity defines a different way of sensitivity investigation technique and the same is analyzed by Eq. 5 [37]:

$$S_A(\lambda) = -\frac{1}{\alpha(\lambda, n_a)} \frac{\partial\alpha(\lambda, n_a)}{\partial n_a} [RIU^{-1}] \tag{5}$$

$\alpha(\lambda, n_a)$ denotes the confinement loss at working wavelength, $\partial\alpha(\lambda, n_a)$ is the measure of the loss gradient between two consecutive analytes at same working wavelength, and $\partial n_a$ defines the difference between the refractive indices of the analytes.

For accessing the detection capability of the bio-sensor, we need figure-of-merit (FOM) which is calculated by FWHM (full width at half maximum) [37]:

$$\text{FOM} = \frac{\text{Wavelength sensitivity}}{\text{FWHM}} \tag{6}$$

Figure 2 illustrates the proposed practical setup of the proposed sensor. Using light source which is emitting spectrum of range 0.5–1.2 µm optical light launched through single-mode fiber SMF. The SPR PCF can be combined with this SMF using various splicing machines. The RBC sample is located in outside channel, and they are pumped through the inlet to the analyte layer. Using a pump, the inflow and outflow of the analyte are controlled. For analyzing the new sample, the sensor probes should be rinsed with deionized water. When the RBC and the gold-coated surface's ligand interacts, the refractive index changes, and this causes the resonance wavelength to shift. This shift can be monitored by an optical spectrum analyzer (OSA).

The operational principles of our proposed sensor can be described as follows: The optical light is passed through the PCF and interacts with the plasmonic material which is coated on the symmetrically polished surfaces. This causes the production of plasmons on the surface (SPPs). When the frequency of this SPP mode matches with the frequency of the core mode, it attains phase-matched

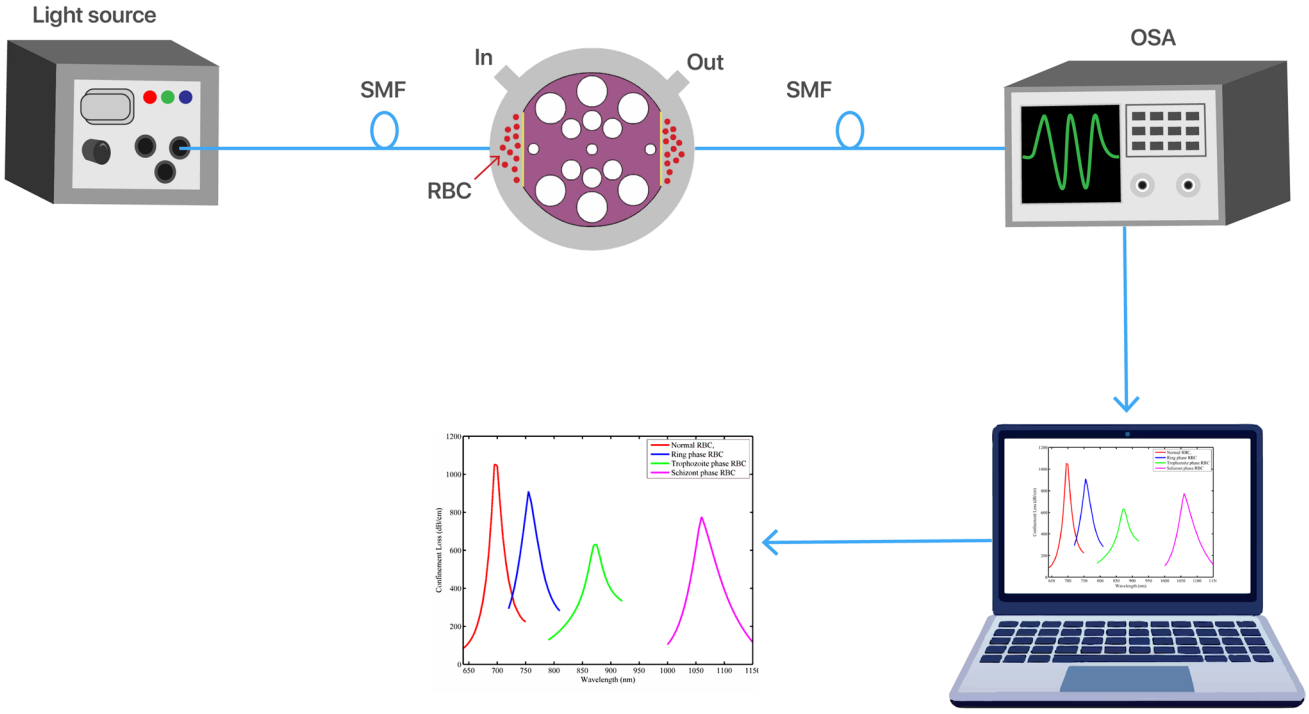



**Fig. 2** Proposed practical setup of the sensor

**Table 1** Average refractive indices of normal and infected RBC at different stages of malaria parasites

| RBC stages | Average RI | References |
|---|---|---|
| Normal RBCs | 1.402 | 29, 30 |
| Ring phase RBCs | 1.395 | |
| Trophozoite phase RBCs | 1.383 | |
| Schizont phase RBCs | 1.373 | |

condition and causes the resonance condition (SPR). So, at the resonance wavelength, the energy of core mode is transferred to SPP mode, and the confinement loss reaches its peak value. The phase matching condition is dependent on the surrounding medium like the RI of the bio-analyte that is introduced for sensing.

For the detection of malaria, we introduce the infected RBC in the analyte layer by using selective filling method. The distinct RI of healthy and infected RBCs, as shown in the Table 1, lead to a shift in the resonance wavelength. This shift is crucial for real-time and label-free detection of malaria.

## Results and Discussions

In this section, we will discuss the results and the analysis of our proposed PCF SPR structure for malaria detection. The effective mode index ($n_{\text{eff}}$) of the proposed PCF SPR

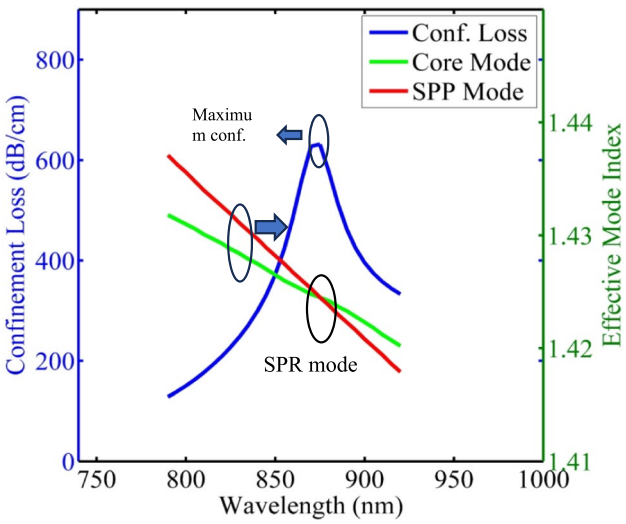



**Fig. 3** Dispersion characteristics of ring-phase RBC (RI: 1.395)

is calculated by solving Maxwell's equation by FEM. The effective mode index is a complex function which contains real part and imaginary part; real part signifies the guided modes and imaginary part describes the loss. Figure 3 depicts the dispersion characteristics of SPP and core mode combined with the loss curve of ring-phase RBC. The green line indicates the $n_{\text{eff}}$ of core mode; the red line indicates $n_{\text{eff}}$ of the SPP mode, while the blue curve depicts the confinement loss. The $n_{\text{eff}}$ of both SPP mode and the core mode

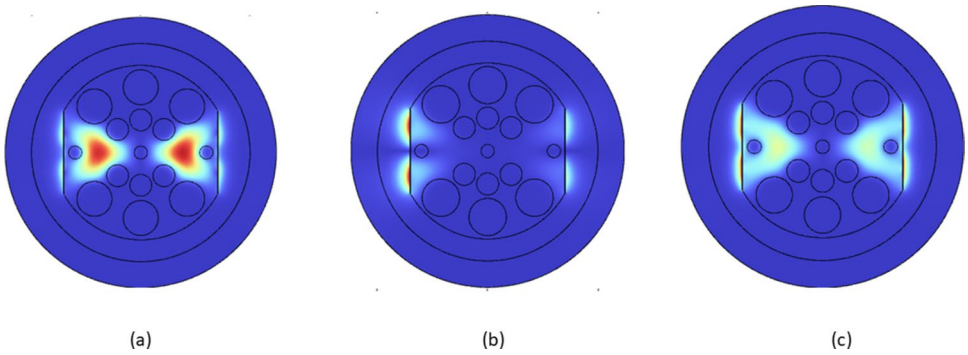



**Fig. 4** **a** Core mode. **b** SPP mode. **c** Coupling mode

decreases with increasing wavelength. They intersect at the resonance wavelength ($\lambda_r = 875$ nm), where the loss reaches its peak at 631.63 dB/cm. At this point, both the SPP mode and the core mode exhibit the same real component of the RI, indicating a coupling between them that facilitates maximum energy transfer from the core mode to the SPP mode. Figure 4a, b, c, represent the electric field distribution of this core mode, SPP mode, and coupling (SPR) mode of the ring-phase RBC, respectively. In bio-analyte detection, the coupling mode is crucial because it allows the energy of the core mode to directly couple with SPP mode.

## Optimization of Structural Parameters

The geometrical parameters influence the sensitivity of the sensors, so their optimization is necessary. In this section, we explore the influence of every geometrical parameter on sensing performance. In this study, we have used a conventional method for optimization called brute-force method [37]. In this method, one parameter is varied by keeping other constants. The parameters used for the optimization of our PCF sensor are diameter of three types of air holes, $d_1$, $d_2$, and $d_3$ and the gold layer thickness $t_g$. For the optimization process, we have only considered the RI of normal RBC and ring-phase RBC, and the respective WS and FOM of the ring-phase RBC are noted. Optimizing these parameters can enhance sensor performance and provide numerous insights into their impact on sensing effectiveness.

### Impact of Varying Central Air Hole Diameter ($d_1$)

The first parameter to optimize is the central hole diameter ($d_1$). The central air hole helps in squeezing the field towards the plasmonic material. Figure 5 depicts the confinement loss variation with the varying central hole diameter $d_1$. Here, the range for optimizing $d_1$ is from 0.5 to 0.7 µm with a step size of 0.1 µm. The respective figure shows that there is a red shift in the resonant wavelength

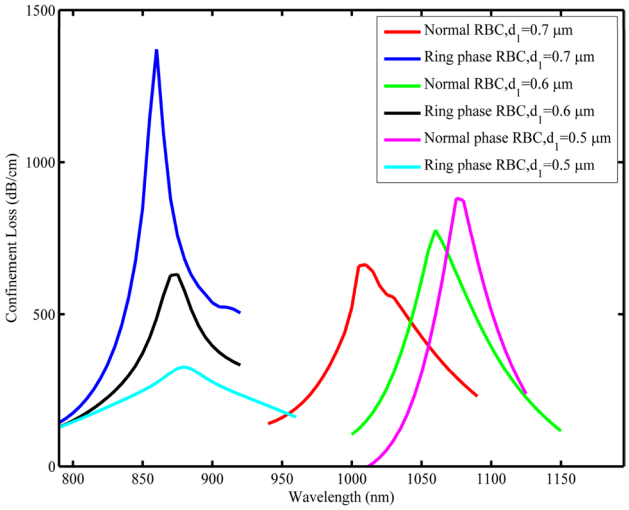



**Fig. 5** Confinement loss characteristics with varying central air hole diameter ($d_1$) for normal and ring phase RBCs

**Table 2** Impact of central air hole diameter ($d_1$)

| | Diameter ($d_1$) (µm) | Wavelength sensitivity (nm/RIU) | FOM ($RIU^{-1}$) |
|---|---|---|---|
| Ring phase | 0.5 | 31428.6 | 209.524 |
| | 0.6 | 26428.6 | 334.539 |
| | 0.7 | 21428.6 | 669.643 |

when the $d_1$ is increased and FWHM decreases for the ring phase. The sensitivity obtained for 0.5 µm, 0.6 µm, and 0.7 µm are 31,428.6 nm/RIU, 26428.6 nm/RIU, and 21428.6 nm/RIU, respectively. Here, 0.5 µm of $d_1$ has a higher WS, but we have selected 0.6 µm as the optimal parameter because the loss curve of 0.5 µm broadens during the ring phase, leading to a reduced FOM value. On the oth,er hand, when $d_1$ is 0.7 µm, it delivers a very high loss compared to other taken parameters. The overall variation in sensitivity and FOM with respect to the changing diameter is presented in Table 2.

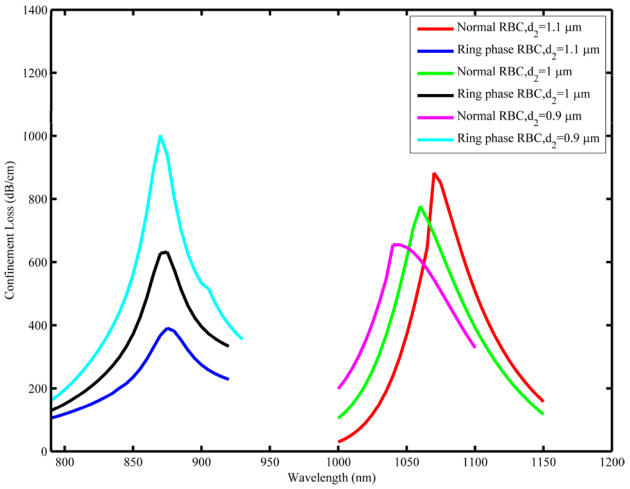



**Fig. 6** Confinement loss characteristics with varying small air hole diameter ($d_2$) for normal and ring-phase RBCs

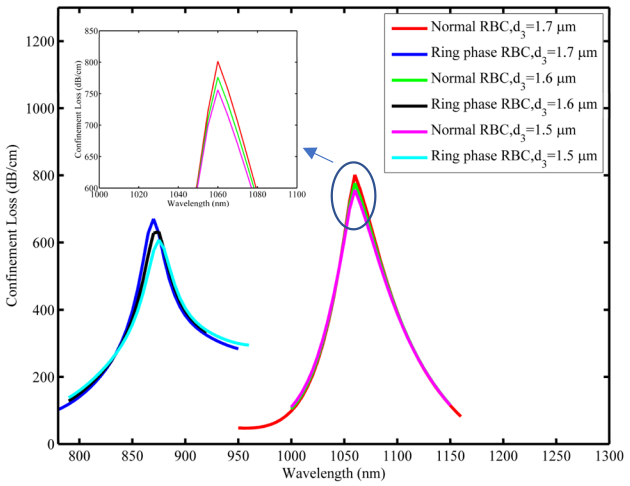



**Fig. 7** Confinement loss characteristics with varying large air hole diameter ($d_3$) for normal and ring-phase RBCs

### Impact of Varying Small Air Hole Diameter ($d_2$)

Figure 6 illustrates confinement loss characteristics of the sensor for small air hole diameter ($d_2$). For the optimization process, we have chosen the magnitudes of $d_2$ as 0.9 μm, 1 μm, and 1.1 μm. These sets of air holes compress the light energy into the twin cores, shaping the cores into a triangular form and reducing the loss. The wavelength sensitivity obtained for ring-phase RBC are 25,000 nm/RIU, 26,428.6 nm/RIU, and 27,857.1 nm/RIU for 0.9 μm,1 μm, and 1.1 μm diameter, respectively. There is a visible red shift in the resonance wavelength in Fig. 7 as the diameter increases. Even though there is a slight increase in the sensitivity for 1.1 μm, we have chosen 1 μm as the optimized parameter because it attains a balanced WS and better FOM value. This is evident from Table 3 which shows summary of the variation of sensitivity and FOM with change of $d_2$.

**Table 3** Impact of small hole diameter ($d_2$)

| | Diameter ($d_2$) (μm) | Wavelength Sensitivity (nm/RIU) | FOM ($RIU^{-1}$) |
|---|---|---|---|
| Ring phase | 0.9 | 25000 | 409.836 |
| | 1 | 26428.6 | 334.539 |
| | 1.1 | 27857.1 | 273.109 |

**Table 4** Impact of larger air hole diameter ($d_3$)

| | Diameter ($d_3$) (μm) | Wavelength sensitivity (nm/RIU) | FOM ($RIU^{-1}$) |
|---|---|---|---|
| Ring phase | 1.5 | 26428.6 | 254.129 |
| | 1.6 | 26428.6 | 334.539 |
| | 1.7 | 27142.9 | 376.984 |

### Impact of Varying Large Air Hole Diameter ($d_3$)

Figure 7 displays the confinement loss versus wavelength for varying large air hole diameter $d_3$. The balanced symmetrical air holes situated on the top and bottom layers assist in minimizing the extent of loss and confine the light into the core. The desired optimal diameter range is taken as 1.5 to 1.7 μm in steps of 0.1 μm in magnitude. The maximum WS in the ring phase is 27,142.9 nm/RIU at 1.7 μm, while at 1.6 μm and 1.5 μm, it is 26,428.6 nm/RIU. Although 1.7 μm of $d_3$ shows a slight increase in sensitivity, in another aspect, it also shows higher loss compared to the others. Additionally, on selection of 1.7 μm as $d_3$ poses fabrication challenges due to the reduced spacing between holes. Thus, 1.6 μm is considered as the optimized parameter. Table 4 includes the sensitivity variation and FOM for ring-phase RBC with the variation of diameter $d_3$.

### Impact of Varying Gold Layer Thickness ($t_g$)

The final optimizing parameter is gold layer thickness ($t_g$). It plays a crucial role in the performance of the sensor. Figure 8 depicts the variation of confinement loss with respect to wavelength at different gold layer thickness. The range of gold layer thickness for optimization is from 38 to 40 nm with a step size of 2 nm. The respective graph shows that increasing the gold layer thickness causes a blue shift in the resonance wavelength and ultimately broadens the loss curve. This broadening is due to a decrease in coupling efficiency between the SPP mode and the core mode. The maximum wavelength sensitivity of 26,428.6 nm/RIU is achieved at a 40-nm thickness, along with an FOM of

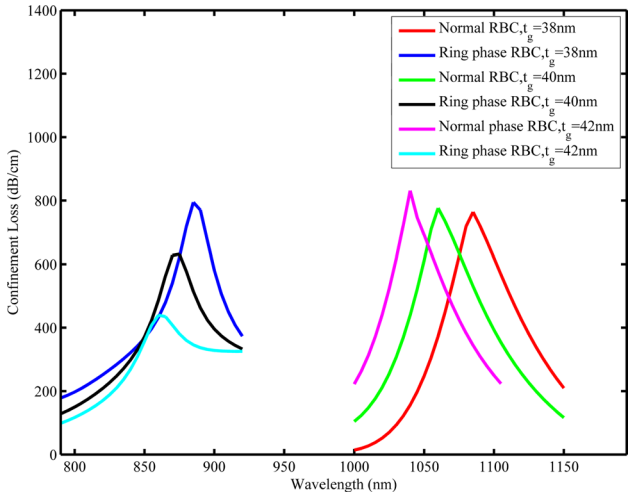



**Fig. 8** Confinement loss spectrum for varying gold layer thickness ($t_g$) for normal and ring-phase RBCs

**Table 5** Impact of gold layer thickness ($t_g$)

| | Thickness ($t_g$) (nm) | Wavelength sensitivity (nm/RIU) | FOM ($RIU^{-1}$) |
|---|---|---|---|
| Ring phase | 38 | 25714 | 252.101 |
| | 40 | 26428.6 | 334.539 |
| | 42 | 23517.4 | 380.184 |

**Table 6** Optimized parameters of the sensor

| Parameter | Optimized value |
|---|---|
| Larger diameter ($d_3$) | 1.6 µm |
| Smaller diameter ($d_2$) | 1 µm |
| Central diameter ($d_1$) | 0.6 µm |
| Gold thickness ($t_g$) | 40 nm |

334.54 $RIU^{-1}$. Although the FOM value slightly increases for a 42-nm thickness, the sensitivity is lower compared to 40 nm. Additionally, there will be high damping loss due to increased gold layer thickness. Therefore, the penetration depth of evanescent field reduces which causes less plasmon generation on the gold layer. Therefore, a 40-nm thickness of gold layer is chosen to achieve high sensitivity along with a moderate FOM value. The tabular summary of the variation of sensitivity and FOM value is given as follows (Table 5).

Through the systematic investigation of the effects of structural parameters on the sensing performance and peak loss of the proposed sensor, the optimal parameters were identified. These carefully fine-tuned parameters, listed in Table 6, are associated with the proposed SPR PCF sensor for malaria detection. Table 7 summarizes the performance with these optimized parameters. Figure 9 represents the loss spectrum obtained by these optimized parameters for different RI of normal and infected RBCs. By examining the figure, it is clear that the normal RBC peak loss is of 776.18 dB/cm at the resonance wavelength 1060 nm. For the ring-phase RBC, the resonance wavelength is 875 nm and corresponding loss peak is 631.63 dB/cm; here, the wavelength shift is about 185 nm. For trophozoite phase, the resonance wavelength is 755 nm and the corresponding peak loss is 910.38 dB/cm; now, the wavelength shift from the normal RBC is 305 nm and for schizont phase, the resonance wavelength and the peak loss is 690 nm and 1052.5 dB/cm, respectively; the corresponding wavelength shift with the ring phase is 365 nm. The pronounced shift in the wavelength is the key parameter for the measurement of wavelength sensitivity. The wavelength sensitivity corresponding to ring phase, trophozoite phase, and schizont phase is 26,428.57 nm/RIU,16,052.63 nm/RIU, and 12,586.21 nm/RIU, respectively. Apart from wavelength sensitivity, our proposed sensor shows maximum amplitude sensitivity of − 318.82 $RIU^{-1}$, and the same is displayed in Fig. 10.

Table 8 provides a performance comparison of the proposed sensor against other PCF SPR designed for malaria detection. The data clearly demonstrates that the proposed sensor outperforms both circular and elliptical structures. This enhanced performance is a viable alternative not only to existing circular-based PCF sensors for malaria detection but also to certain dual-side polished and dual-core SPR PCF RI biosensors. The findings suggest that the proposed sensor could significantly improve detection capabilities, potentially leading to more accurate and efficient diagnostic methods.

**Table 7** Performance summary of the proposed sensor

| Parameter | Peak loss (dB/cm) | Resonance wavelength (nm) | Wavelength sensitivity (nm/ RIU) | Amplitude sensitivity ($RIU^{-1}$) | Average wavelength sensitivity (nm/RIU) | Average amplitude sensitivity ($RIU^{-1}$) | FOM ($RIU^{-1}$) |
|---|---|---|---|---|---|---|---|
| Normal RBC | 776.18 | 1060 | | | 18,355.4 | − 172.96 | |
| Ring phase RBC | 631.63 | 875 | 26,428.6 | − 318.82 | | | 334.53 |
| Trophozoite RBC | 910.38 | 755 | 16,052.6 | − 122.52 | | | 321.05 |
| Schizont RBC | 1052.5 | 695 | 12,586.2 | − 77.55 | | | 381.4 |

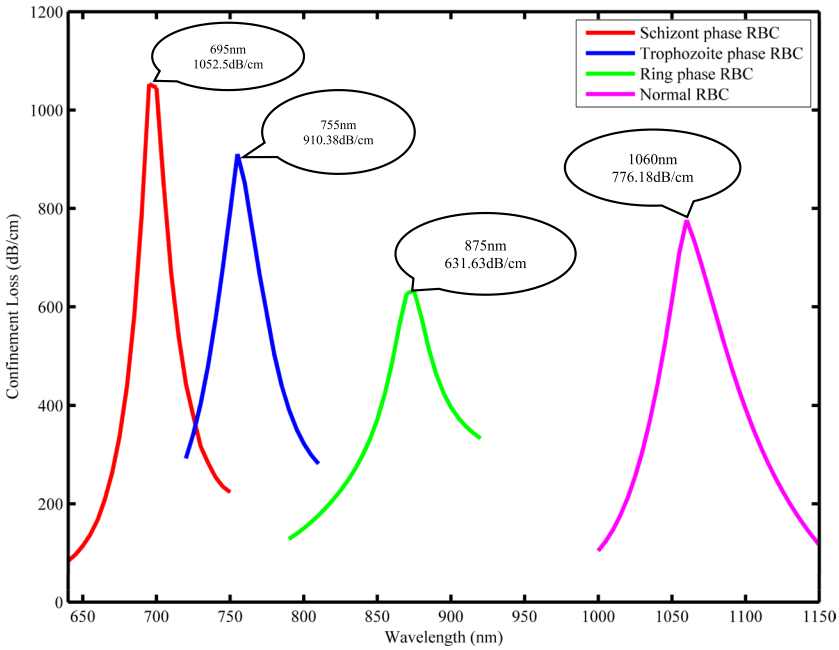



**Fig. 9** Wavelength sensitivity for different RBCs with optimized parameters

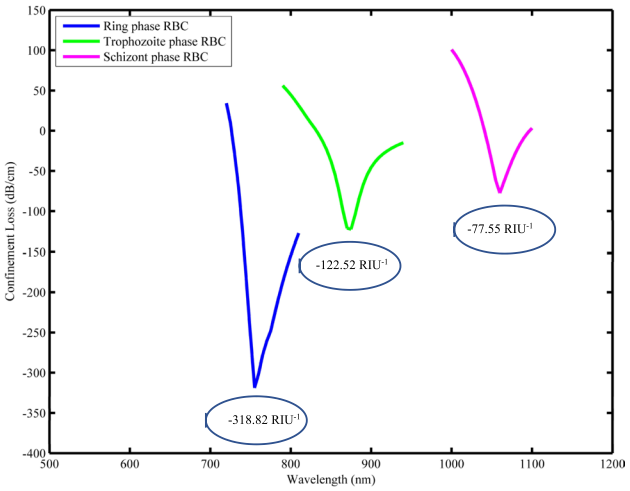



**Fig. 10** Amplitude sensitivity for different RBCs with optimized parameters

## Tolerance Investigation of Structural Parameters

To analyze the fabrication tolerance of the geometrical parameters of the proposed Malaria sensor, we have regulated the parameters by ± 3% and examined the variations in the effective results. The first step of overall investigation consists of variation of thickness of gold by ± 3% which concludes that the peak loss at RI 1.402 is influenced by 2.75%, and the important factor wavelength sensitivity is only suffered a ± 2.7% change. The parameter $d_3$ has not much influence on the core confinement; therefore, negligible variations are observed with the ± 3% variation. The third study covers the variation of small air holes $d_2$ by ± 3%; the extracted results show the peak loss is only varied by ± 3.5% whereas wavelength sensitivity suffered a ± 2.7% change only. The final parameter $d_1$ is assorted by ± 3% in the next study, and the respective study shows that the peak loss in affected by ± 7%, but the wavelength sensitivity is only by ± 2.7%. The comprehensive investigation shows that on variation of parameters by ± 3%, the sensitivity is only affected by ± 2.7%. The results define that the proposed sensor is majorly tolerable to the fabrication faults and other experimental imperfections.

## Fabrication Viability of the Proposed Sensor

Nowadays, fabrication of PCF sensors faces several limitations and challenges. The dual side–polishing technique makes fabrication easier than methods with complex internal coatings, making it better suited for large-scale production. The use of external sensing methods also reduces complications associated with analyte injection and extraction. There are many fabrication techniques such as stack and draw, sol–gel casting, and injection molding. Due to lower cost and flexibility, the stack and draw approach is chosen [40, 41]. The silica capillaries used for PCF fabrication offer a negligible transmission loss of 0.18 dB/km at 1.55 µm. Since we are working on below 1.55 µm wavelength range, the loss can be neglected. Figure 11 shows the sequential fabrication steps involved in the stack and draw method. The thin and thick capillaries and solid rods are stacked still, the PCF

**Table 8** Performance comparison with other malaria detection PCF sensor

| Reference | PCF configuration | Materials used | $\lambda_{range}$ in μm | Wavelength sensitivity (nm/RIU) |
|---|---|---|---|---|
| 29 | Circular | Gold coated and $TiO_2$ | 0.7–1.1 | 17,857.14 (ring phase)<br>10,210.53(trophozoite phase)<br>8758.62 (schizont phase) |
| 30 | Circular | Gold | 0.7–1.0 | 14,285.71 (ring phase)<br>10,000 (trophozoite phase)<br>8206.9 (schizont phase) |
| 31 | Circular | Gold coated and Mxene | 0.7–1.05 | 12,142 (ring phase)<br>9736 (trophozoite phase)<br>8241 (schizont phase) |
| 32 | Elliptical | Silica | 1.3–2 | 11,428.57 (ring phase)<br>9473.68 (trophozoite phase)<br>9655.15 (schizont phase) |
| 38 | Dual-square groove | Gold and $MoS_2$ | 0.7–0.95 | 12,248 (ring phase)<br>9210 (trophozoite phase)<br>7758 (schizont phase) |
| 39 | Circular | Gold and $Al_2O_3$ | 0.6–1.0 | 5714.28 (ring phase)<br>5263.15 (trophozoite phase)<br>5931.03 (schizont phase) |
| **This work** | **Side polished** | **Gold** | **0.65–1.15** | 26,428.57 (ring phase) 16,052.6 (trophozoite phase)<br>12,586.2 (schizont phase) |

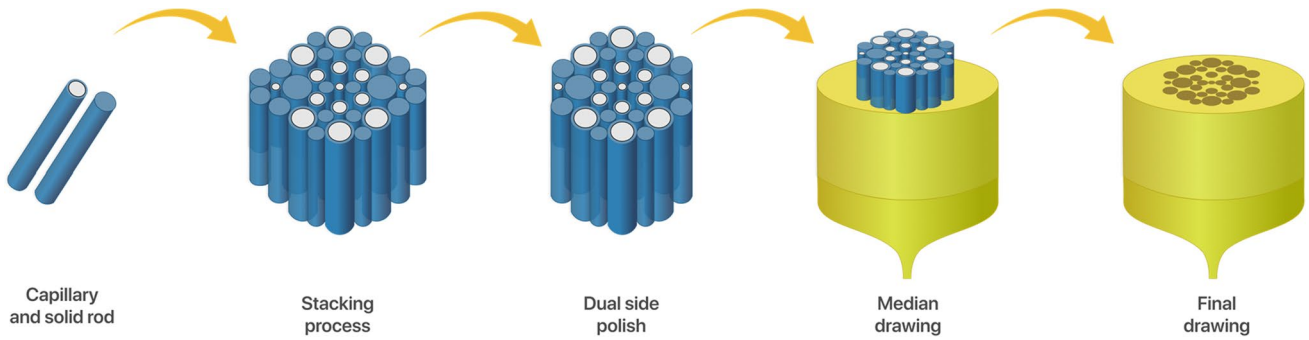



**Fig. 11** Proposed sequential fabrication steps involved in the stack and draw method

is formed. Once the PCF is formed, two sides are polished symmetrically by different methods such as diamond polishing [42]. A thin layer of gold is coated above these polished surfaces. Normal metal coating methods such as RF sputtering, thermal evaporation methods, electro plating, and wet-chemistry deposition create massive surface roughness in coating the circular surface. The chemical vapor deposition technique (CVD) is the most suitable one, which provides uniform coating in the nanometer thickness range. The usage of a surfactant layer of oxidized copper or chromium on Au and Ag films can avoid percolation tendencies of gold [43]. At the final step, the RBC sample is introduced above this layer for the malaria detection [17].

## Conclusion

Malaria is a life-threatening disease that requires prompt attention, as early detection is essential for effective treatment, reducing complications, and preventing further transmission. In this paper, we have theoretically proposed a highly sensitive twin-core double side–polished SPR PCF sensor for the detection of malaria during various intraerythrocytic cycle stages such as ring phase, trophozoite phase, and schizont phase. The symmetrically polished surfaces provide high sensitivity by enhancing the coupling between the SPP mode and core mode. The hourglass-shaped arrangement of air holes enhances the performance of the sensor by confining the light into the

twin cores. The structural parameters have been optimized to achieve high sensitivity and robustness in the sensor. The research results demonstrate that the proposed sensor design exhibits high wavelength sensitivities of 26,428.57 nm/RIU, 16,052.6 nm/RIU, and 12,586.2 nm/RIU for the ring, trophozoite, and schizont phases, respectively. Additionally, it achieves high figure of merit (FOM) values of 334.54 $RIU^{-1}$, 321.05 $RIU^{-1}$, and 381.40 $RIU^{-1}$ for the ring, trophozoite, and schizont phases, respectively. A major challenge in the current scenario is the shortage of skilled technicians and delays in malaria detection mainly in rural distant location. The straightforward detection process, simple design, compatibility, and portability provided by the proposed SPR PCF biosensor make it an ideal solution for malaria detection.

**Acknowledgements** Devi Kangan B S, student of Masters of Physics in the Department of Applied Physics at Delhi Technological University, is acknowledged for providing excellent research facilities for conducting research.

**Author Contributions** Devi K B S: Scientific computation, writing original draft, methodology, validation. A. Khamaru: Review and editing draft, validation, A. Kumar.: Supervision, conceptualization, visualization, editing draft.

**Data Availability** No datasets were generated or analysed during the current study.

## Declarations

**Competing Interests** The authors declare no competing interests.

## References

1. Nureye D, Assefa S (2020) Old and recent advances in life cycle, pathogenesis, diagnosis, prevention, and treatment of malaria including perspectives in Ethiopia. Sci World J 2020(1):1295381
2. Liu PY et al (2016) Cell refractive index for cell biology and disease diagnosis: past, present and future. Lab Chip 16(4):634–644
3. Cowman AF, Healer J, Marapana D, Marsh K (2016) Malaria: biology and disease. Cell 167(3):610–624
4. Molina-Franky J et al (2020) Plasmodium falciparum pre-erythrocytic stage vaccine development. Malar J 19:1–18
5. World Health Organization (2021) World Malaria Report 2021. Available at; https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2022
6. Zhang M et al (2018) Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis. Science 4 360(6388):eaap7847
7. Weber MW et al (1997) Evaluation of an algorithm for the integrated management of childhood illness in an area with seasonal malaria in the Gambia. Bull World Health Organization 75(Suppl 1):25
8. Jager MM, Murk JL, Piqué RD, Hekker TA, VandenbrouckeGrauls CM (2011) Five-minute Giemsa stain for rapid detection of malaria parasites in blood smears. Trop Doct 41(1):33–35
9. Parija SC, Dhodapkar R, Elangovan S, Chaya DR (2009) A comparative study of blood smear, QBC and antigen detection for diagnosis of malaria. Indian J Pathol Microbiol 52(2):200–202
10. Ranadive N et al (2017) Limitations of rapid diagnostic testing in patients with suspected malaria: a diagnostic accuracy evaluation from Swaziland, a low-endemicity country aiming for malaria elimination. Clin Infect Dis 64(9):1221–1227
11. Makler MT, Palmer CJ, Ager AL (1998) A review of practical techniques for the diagnosis of malaria. Ann Trop Med Parasitol 92(4):419–434
12. Hänscheid T, Grobusch MP (2002) How useful is PCR in the diagnosis of malaria. Trends Parasitol 18(9):395–398
13. Sodeman WA Jr, Jeffery GM (1966) Indirect fluorescent antibody test for malaria antibody. Public Health Rep 81(11):1037
14. Poon LL et al (2006) Sensitive and inexpensive molecular test for falciparum malaria: detecting Plasmodium falciparum DNA directly from heat-treated blood by loop-mediated isothermal amplification. Clin Chem 52(2):303–306
15. Wongchotigul V et al (2004) The use of flow cytometry as a diagnostic test for malaria parasites. Southeast Asian J Trop Med Public Health 35:552–559
16. De Langen AJ, Van Dillen J, Witte PD, Mucheto S, Nagelkerke N, Kager P (2006) Automated detection of malaria pigment: feasibility for malaria diagnosing in an area with seasonal malaria in northern Namibia. Tropical Med Int Health 11(6):809–816
17. Rifat AA et al (2017) Photonic crystal fiber based plasmonic sensors. Sens Actuators B Chem 243:311–325
18. Singh G, Sharma S, Kumar A (2024) Design and simulation of solid-core octagonal photonic crystal fiber for terahertz wave propagation. Microw Opt Technol Lett 66(5):e34173
19. Singh N, Khamaru A, Kumar A (2024) Design and analysis of a rectangular core refractive index-based PCF sensor for biosensing application. Opt Quant Electron 56(7):1133
20. Singh J, Khamaru A, Kumar A (2024) Spiral shaped highly sensitive rectangular PCF-based cancer cells detector in terahertz regime. Phys Scr 99(11):115546
21. Khamaru A, Kumar A (2024) $As_{38}Se_{62}$ based segmented cladgraded index photonic crystal fiber for supercontinuum generation covering 3–9.5 μm with moderate peak power. Opt Quantum Electron 56(7):1246
22. Khamaru A, Kumar A (2024) Ge-Se-Te based penrose photonic quasi-crystal fiber for SCG covering 2–21 μm MIR regime. Opt Mater 155:115849
23. Das S, Singh VK (2022) Highly sensitive PCF based plasmonic biosensor for hemoglobin concentration detection. Photonics Nanostructures Fundamentals Appl 51:101040
24. Jain S, Choudhary K, Kumar S (2022) Photonic crystal fiber-based SPR sensor for broad range of refractive index sensing applications. Opt Fiber Technol 73:103030
25. Helmerhorst E, Chandler DJ, Nussio M, Mamotte CD (2012) Real-time and label-free bio-sensing of molecular interactions by surface plasmon resonance: a laboratory medicine perspective. Clin Biochemist Rev 33(4):161
26. Robelek R (2009) Surface plasmon resonance sensors in cell biology: basics and application. Bioanal Rev 1:57–72
27. Boruah R, Mohanta D, Choudhury A, Nath P, Ahmed GA (2015) Surface plasmon resonance-based protein bio-sensing using a Kretschmann configured double prism arrangement. IEEE Sens J 15(12):6791–6796
28. Liu C, Lü J, Liu W, Wang F, Chu PK (2021) Overview of refractive index sensors comprising photonic crystal fibers based on the surface plasmon resonance effect. Chin Opt Lett 19(10):102202
29. Das S, Sen R (2024) Design and numerical analysis of a PCF-SPR sensor for early-stage malaria detection. Plasmonics 1–16

30. Chaudhary VS, Kumar D, Kumar S (2021) Gold-immobilized photonic crystal fiber-based SPR biosensor for detection of malaria disease in human body. IEEE Sens J 21(16):17800–17807
31. Kumar A, Verma P, Jindal P (2023) Machine learning approach to surface plasmon resonance sensor based on MXene coated PCF for malaria disease detection in RBCs. Optik 274:170549
32. Shafkat A, Rashed ANZ, El-Hageen HM, Alatwi AM (2021) Design and analysis of a single elliptical channel photonic crystal fiber sensor for potential malaria detection. J Sol-Gel Sci Technol 98:202–211
33. Khamaru A, Kumar A (2024) Highly sensitive graded index PCF-based refractometric SPR sensor offering enhanced FOM. Plasmonics 1–13
34. Lv J et al (2022) Surface plasmon resonance sensor based on the dual core D-shape photonic crystal fiber for refractive index detection in liquids. Opt Eng 61(8):086111–086111
35. Sarfraz N, Khan I (2021) Plasmonic gold nanoparticles (AuNPs): properties, synthesis and their advanced energy, environmental and biomedical applications. Chem An Asian J 16(7):720–742
36. Singh S, Prajapati YK (2020) $TiO_2$/gold-graphene hybrid solid core SPR based PCF RI sensor for sensitivity enhancement. Optik 224:165525
37. Praveena S, Senthilnathan K (2024) Performance enhancement of gold coated D-shaped PCF sensor using Monolayer $MoS_2$. Plasmonics 1–12
38. Praveena S, Senthilnathan K (2024) Ultra low loss dual square groove incorporated SPR-based PCF sensor for malaria diagnosis. Plasmonics 1–19
39. Yadav SM (2024) Au-$Al_2O_3$ coated highly sensitive broad range refractive index sensor for detecting malaria disease in human blood. IEEE Sens J
40. Pysz D et al (2014) Stack and draw fabrication of soft glass microstructured fiber optics. Bulletin of the polish academy of sciences. Techn Sci 62(4):667–682
41. AmouzadMahdiraji G et al (2014) Challenges and solutions in fabrication of silica-based photonic crystal fibers: an experimental study. Fiber Integr Opt 33(1–2):85–104
42. Schuelke T, Grotjohn TA (2013) Diamond polishing. Diam Relat Mater 32:17–26
43. Luhmann N et al (2020) Ultrathin 2 nm gold as impedance-matched absorber for infrared light. Nat Commun 11(1):2161

**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.

ISSN: 2584-0495 Vol. 2, Issue 7, pp. 1044-1049

International Journal of Microsystems and IoT

ISSN: (Online) Journal homepage: https://www.ijmit.org

# Design and Implementation of an IoT-Based Patient Monitoring System for Comprehensive Vital Tracking and Secure Data Access

**Vihaan Ravishankar, Mohammad Ibrahim, Anjali Jain and Neelam Verma**

**Cite as:** Ravishankar, V., Ibrahim, M., Jain, A., & Verma, N. (2024). Design and Implementation of an IoT-Based Patient Monitoring System for Comprehensive Vital Tracking and Secure Data Access. International Journal of Microsystems and IoT, 2(7), 1044-1049. https://doi.org/10.5281/zenodo.13332142

© 2024 The Author(s). Published by Indian Society for VLSI Education, Ranchi, India

Published online: 17 July 2024

Submit your article to this journal:

Article views:

View related articles:

View Crossmark data:

**DOI:** https://doi.org/10.5281/zenodo.13332142

Full Terms & Conditions of access and use can be found at https://ijmit.org/mission.php

International Journal of Microsystems and IoT
ISSN: 2584-0495; Vol. 2, Issue 7, pp.1044-1049; DOI: https://doi.org/10.5281/zenodo.13332142

Check for updates

# Design and Implementation of an IoT-Based Patient Monitoring System for Comprehensive Vital Tracking and Secure Data Access

Vihaan Ravishankar, Mohammad Ibrahim, Anjali Jain and Neelam Verma

Department of Electrical and Electronics Engineering, Amity University Uttar Pradesh

**ABSTRACT**
The healthcare industry is undergoing a transformation driven by technological advancements, and the integration of Internet of Things (IoT) devices into patient care has emerged as a promising solution. This report presents the design and implementation of an IoT-based Patient Monitoring System (PMS) that offers comprehensive vital tracking and ensures secure data access for healthcare providers and patients. The IoT-based PMS comprises a network of wearable and non-invasive sensors that continuously collect real-time data on patients' vital signs, including heart rate, blood pressure, body temperature, and oxygen saturation. This data is transmitted securely to a centralized cloud platform via a robust and encrypted communication protocol. The cloud platform acts as the central hub for data storage and processing. One of the key features of the system is its comprehensive vital tracking capabilities, which enable healthcare providers to monitor patients' health conditions with a high level of granularity. Secure data access is ensured through strong authentication mechanisms and data encryption, adhering to strict privacy regulations and standards. Robust access control mechanisms and encryption techniques are implemented to safeguard sensitive patient information. This report also addresses the technical aspects of system implementation, including hardware and software components, communication protocols, and database design. The system's scalability and reliability are discussed to accommodate varying healthcare settings and patient populations. The IoT-based Patient Monitoring System presented in this report represents a significant step toward enhancing patient care by providing comprehensive vital tracking and secure data access.

Keywords
Healthcare industry, technological advancements, Internet of Things (IoT), patient care, Patient Monitoring System (PMS), vital tracking, wearable sensors, non-invasive sensors, database design, scalability, reliability, patient care enhancement.

## 1. INTRODUCTION

In an era defined by technological advancements, the field of healthcare has been witnessing a transformative shift towards patient-centric care. As medical technology continues to evolve, there is a growing need for innovative solutions that facilitate real-time, comprehensive health monitoring for patients, especially those with chronic conditions. This project is centered around the design and implementation of an IoT-based patient monitoring system, which aims to address this pressing need. The system's primary objectives are to provide comprehensive vital tracking and ensure secure data access, thus fostering a proactive, efficient, and patient-friendly healthcare environment [1].
The healthcare industry is undergoing a profound evolution characterized by a shift from the traditional episodic, reactive care model to a more continuous, proactive, and patient-centric approach. As part of this evolution, healthcare providers and researchers are increasingly recognizing the importance of continuous monitoring of patients' vital signs, both in clinical and home settings. The capability to monitor patients' health in real-time enables early detection of health anomalies, personalized treatment plans, and enhanced disease management.

The Internet of Things (IoT) has emerged as a transformative force in healthcare, offering a plethora of applications aimed at improving patient outcomes and healthcare efficiency. IoT technology, through the interconnectivity of devices, sensors, and data, can revolutionize patient monitoring, enabling healthcare professionals to remotely track vital signs, providing timely interventions, and empowering patients with health insights.

Accurate and continuous monitoring of vital signs, such as heart rate, blood pressure, temperature, and oxygen saturation, is fundamental for assessing a patient's health status. By designing a system that can seamlessly collect, transmit, and analyze these vital parameters, healthcare providers can detect early warning signs, track disease progression, and make data-driven decisions that improve patient outcomes [14].

This project is centered around the design and implementation of an IoT-based patient monitoring system, which aims to address this pressing need. The system's primary objectives are to provide comprehensive vital tracking, including heart rate, blood pressure, body temperature, and oxygen saturation, and ensure secure data access, thus fostering a proactive, efficient, and patient-friendly healthcare environment.

### Unique Selling Proposition

The unique selling proposition (USP) of an IoT-based Patient Monitoring System lies in its ability to revolutionize the healthcare industry by providing real-time and remote patient monitoring, thus enhancing the quality of care, improving patient outcomes, and increasing overall operational efficiency. The USP of this system can be articulated through several key points, as outlined below:

1. **Real-time Remote Monitoring:**

The system allows healthcare providers to remotely monitor

patients in real-time, enabling them to keep track of vital health parameters without the need for constant physical presence [7]. This feature is especially crucial for patients requiring continuous monitoring, such as those with chronic illnesses or those recovering from surgeries.

2. **Enhanced Patient Care and Safety**:

By continuously monitoring vital signs and health parameters, the system can swiftly detect any abnormalities or critical events, such as flatlining, allowing healthcare providers to intervene promptly [8]. This proactive approach significantly enhances patient safety and reduces the risk of adverse events.

3. **Data-Driven Decision Making:**

The system collects and processes a wealth of patient data, providing healthcare professionals with comprehensive insights into a patient's health status over time. By leveraging advanced analytics and data visualization tools, healthcare providers can make informed, data-driven decisions regarding treatment plans, medication adjustments, and overall care management [4].

4. **Efficient Workflow and Resource Optimization:**

By automating the monitoring process and providing streamlined access to patient data, the system optimizes healthcare workflows, allowing medical professionals to allocate their time and resources more efficiently. This optimization can lead to improved patient care, reduced operational costs, and enhanced overall hospital management.

5. **Customized and Adaptive Care Plans:**

Through the analysis of historical patient data and the implementation of machine learning algorithms, the system can identify patterns and trends in a patient's health, enabling the customization of individualized care plans [5]. This tailored approach ensures that patients receive personalized care that is specifically tailored to their unique medical needs and conditions.

6. **Scalability and Interoperability:**

The system is designed with scalability and interoperability in mind, allowing for seamless integration with existing healthcare infrastructures and the potential for expansion to accommodate a growing number of patients and healthcare facilities. This adaptability ensures that the system can evolve alongside technological advancements and the changing needs of the healthcare industry [12].

7. **Secure and Compliant Data Management:**

Data security and patient privacy are paramount in healthcare systems. The IoT-based Patient Monitoring System ensures the highest standards of data security and compliance with regulatory requirements, safeguarding patient information from unauthorized access or breaches. By adhering to strict data protection protocols, the system maintains patient trust and ensures the confidentiality of sensitive medical data [20].

## 2. SYSTEM DESIGN

In the realm of patient monitoring systems, sensors and wearable devices play an integral role. These sophisticated technologies are designed to collect essential health data from individuals in real-time, offering a holistic view of their well-being.

At the heart of the system lies a diverse range of sensors, each with as specific role in patient health monitoring. These sensors serve as the system's sensory network, capturing a wealth of critical health data [13].

Following Components are used:

### 1. MAX30102 Heart Rate and Pulse Oximeter

The MAX30102 is a highly integrated optical sensor that combines two LEDs, a photodetector, optimized optics, and low-noise analog signal processing to detect pulse oximetry and heart rate signals. It operates based on the principle of photoplethysmography (PPG), which involves measuring the volumetric variation of blood circulation within tissue. This process utilizes the LEDs to emit light into the skin and the photodetector to measure the intensity of light transmitted or reflected by the blood.
In pulse oximetry, the MAX30102 detects the oxygen saturation level in the blood by analyzing the ratio of oxygenated hemoglobin to the total concentration of hemoglobin in the blood. The advanced signal processing algorithms integrated within the MAX30102 enable it to filter out ambient light noise and motion artifacts, ensuring reliable and precise measurements. Its compact size and low power consumption make it suitable for various wearable health monitoring devices and fitness trackers, enhancing the accessibility of real-time health data for users. Additionally, its compatibility with a wide range of microcontrollers and development platforms facilitates seamless integration into diverse healthcare applications.

### 2. DHT 11 Temperature and Humidity Sensor:

The DHT11 sensor is a cost-effective digital temperature and humidity sensor that utilizes a thermistor to measure ambient temperature accurately. Operating on the principle of resistance change with temperature, the sensor converts the temperature variations into digital signals, providing reliable temperature data for various applications. Its simplistic design and easy interfacing make it an ideal choice for temperature monitoring in environments where precision requirements are moderate.

Equipped with a capacitive humidity sensor, the DHT11 can also measure relative humidity, providing valuable insights into the moisture content of the surrounding air. Its low cost, low power consumption, and high reliability have made it popular in applications such as climate control systems, weather stations, and indoor environmental monitoring devices [18].

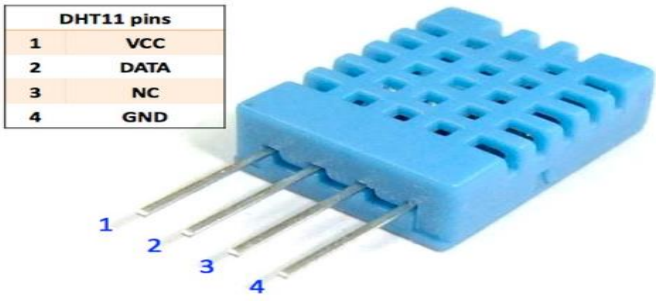



Figure 1: DHT11 sensor[2].

### 3. Arduino Mega 2560 R3:

The Arduino Mega 2560 R3 is an advanced microcontroller board designed for complex and demanding projects that require a larger number of digital input/output pins and more memory space for data storage and processing.

Featuring 54 digital input/output pins (of which 15 can be used as PWM outputs), 16 analog inputs, and a larger program memory space of 256 KB, the Arduino Mega 2560 R3 provides enhanced flexibility and scalability for developers and enthusiasts working on projects that demand extensive connectivity and computational power. Its multiple communication interfaces, including UART, SPI, and I2C, enable seamless integration with a wide range of sensors, displays, and other peripheral devices, facilitating the creation of sophisticated embedded systems and automation solutions.

**Communication Protocols:**

1. Wi-Fi
   - Employ a compatible Wi-Fi module (like ESP8266 or ESP32) to establish a connection to your Wi-Fi network.
   - Transmit the sensor data to a remote server or cloud platform using HTTP, MQTT, or another communication protocol [11].

2. $I^2C$ Protocol
   a. Developed by Philips Semiconductor, now NXP Semiconductors, in the 1980s.
   b. Enables communication between integrated circuits on a circuit board.
   c. Master-slave protocol with one master and multiple slaves.
   d. Two communication lines: SDA (Serial Data Line) and SCL (Serial Clock Line).
   e. Communication initiated by the master, which sends the start condition.

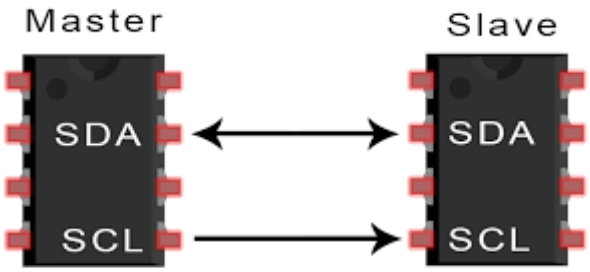



Figure 2 : Basics of I2C Communication [6].

   f. Slaves have unique addresses for communication with the master.
   g. Data transfer in 8-bit bytes with acknowledgment bits.

Following Software are used:

**Arduino IDE**: Arduino Integrated Development Environment (IDE) is a comprehensive software platform that serves as the primary interface for programming Arduino boards. It provides a user-friendly environment for writing, compiling, and uploading code to Arduino microcontrollers, enabling developers, makers, and hobbyists to easily create and deploy a wide range of interactive projects and prototypes.

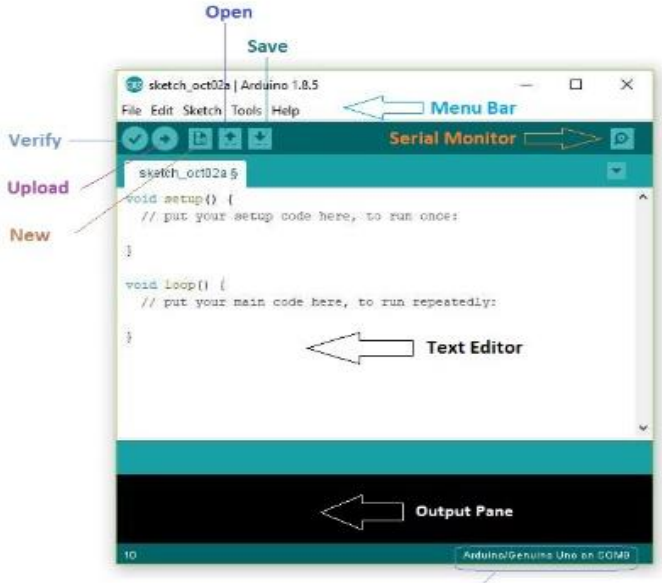



Figure 3: Sample interface of an Arduino IDE software [3].

## 3. SYSTEM BENEFITS

Certainly, here is a list of benefits specific to the IoT-based Patient Monitoring System utilizing an Arduino Mega with DHT11 and MAX30102 sensors, integrated with a Blynk server:

   i. Real-Time Patient Monitoring: Enables continuous, real-time monitoring of vital health parameters,

providing healthcare professionals with instant access to critical patient data.

ii. Data Visualization and Analysis: Presents vital health metrics such as temperature, heart rate, and SpO2 concentration through intuitive and visually informative gauges, aiding in the quick and comprehensive assessment of patient health status [10].

iii. Early Anomaly Detection: Facilitates the early detection of anomalies or irregularities in patient vital signs, enabling timely medical interventions and potentially preventing adverse health events [16].

iv. Remote Alarming and Notifications: Enables the setup of remote alarms and notifications, alerting healthcare providers in real-time in the event of critical changes in patient health parameters, ensuring prompt medical attention [19].

v. Cost-Effective and Scalable Solution: Offers a cost-effective and scalable solution for patient monitoring, eliminating the need for constant physical presence and enabling the expansion of remote monitoring capabilities as needed.[15]

vi. Integration with Existing Systems: Provides the flexibility to integrate with existing healthcare systems, allowing for the seamless incorporation of patient data into electronic health records and other medical information systems.

## 4. IMPLEMENTATION

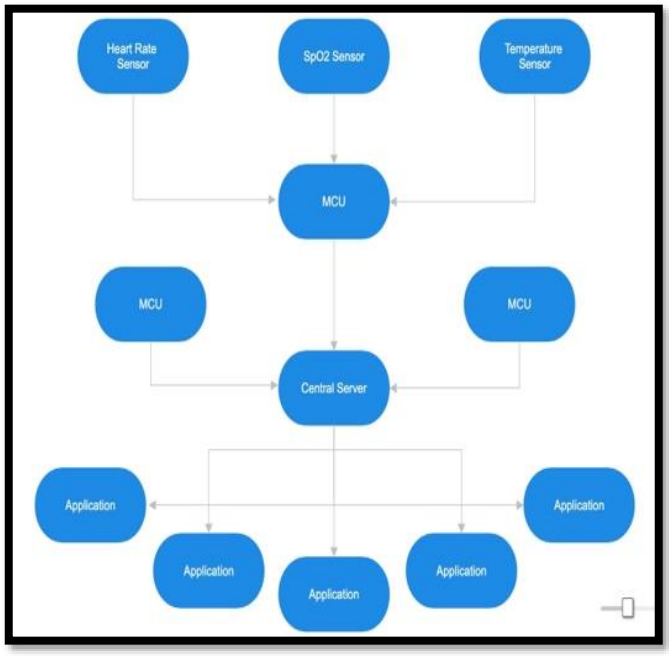



Figure 4: Step by step process of circuit methodology

### Algorithm Design:

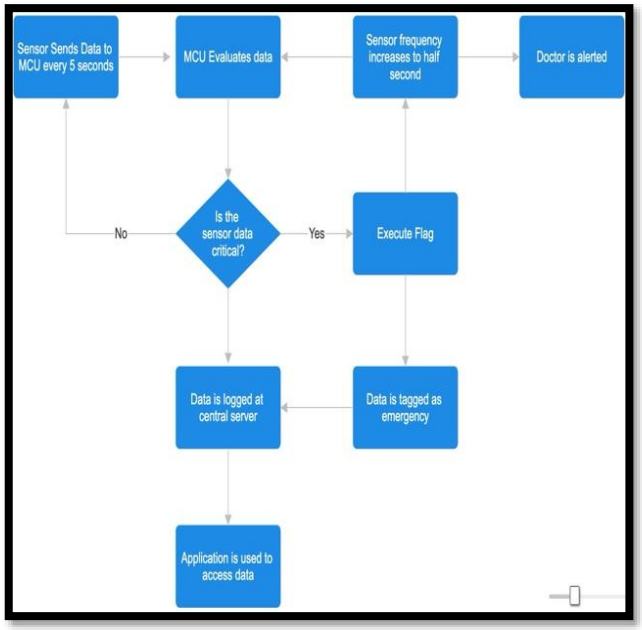



Figure 5 Step by step flow chat

To create the patient monitoring system using an Arduino Mega, a DHT11 sensor, and a MAX30102 sensor, we carefully connected all the parts together. The Arduino Mega acted as the brain, helping us manage and understand the information from the sensors. We used its special pins to connect the sensors and make sure they were powered properly. The DHT11 sensor helped us check how warm and humid it was around the patient. It was like a small weather station that told us about the room's temperature and humidity. The MAX30102 sensor was our special tool for checking the patient's heart rate and how much oxygen was in their blood. It gave us important information about the patient's health in real-time.

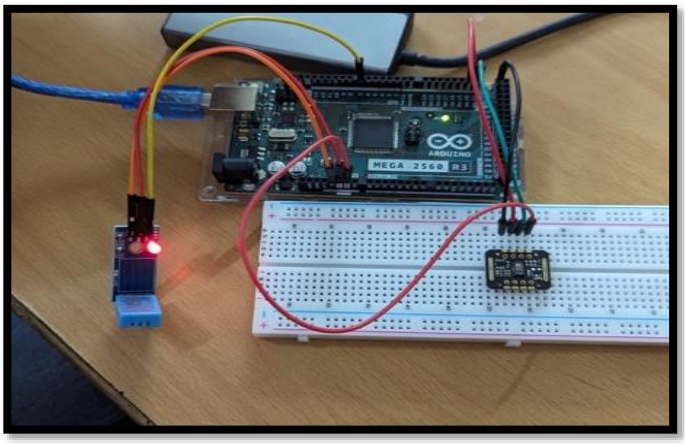


Figure 6: Sample Circuit of project

## 5. DIFFERENT TYPE OF TECHNIQUE USED BY PROGRAMMABLE INSTRUMENTATION AMPLIFIER(PG-INA)

The experiment demonstrated the successful integration of an Arduino Mega with DHT11 and MAX30102 sensors, enabling real-time data collection and transmission to a Blynk server. The system effectively monitored and displayed temperature, heart rate, and SpO2 concentration through dedicated gauges on the Blynk platform.

## 6. CONCLUSION

In summation, our project's successful execution, centered on deploying an IoT-driven patient monitoring system, has unveiled a paradigm shift in the healthcare landscape. The incorporation of cutting-edge technology has given rise to a comprehensive solution that not only ensures the continuous surveillance of a patient's vital parameters but also guarantees unparalleled security and seamless accessibility to their medical records [9].

As we venture into an era where remote healthcare and telemedicine assume pivotal roles, our IoT-based patient monitoring system seamlessly aligns with the contemporary demands of healthcare provision [8]. This endeavor represents only the tip of the iceberg; it is a precursor to continued research and refinement, promising innovative and resilient solutions that will benefit both patients and healthcare providers. The road ahead teems with possibilities for IoT technology in healthcare, and our project signifies a pivotal step in this transformative odyssey. The horizon shimmers with the promise of elevated patient care, streamlined healthcare processes, and the evolution of medical research, driven by the conscientious application of advanced technology in the service of human health [17].

## REFERENCES.


1. Carter, E. (2021). Enhancing Patient Monitoring Through IoT: A Comprehensive Analysis," IEEE Access. 16(1), 3681-13700. Doi: https://www.ijsdcs.com/index.php/TLHS/article/view/520

2. Abdulmalek, S., Nasir, A., Jabbar, W. A., Almuhaya, M. A. M., Bairagi, A.K. (2020). IoT-Based Smart Healthcare Monitoring System. 20(1), 5771-5780. Doi: 10.3390/healthcare10101993

3. Patelkana, K., Devapatla, C., & Ramesh, R. (2021). IoT-Based Smart Health Monitoring System Using Cloud Services. 8(1), 575–581. Doi: https://link.springer.com/chapter/10.1007/978-981-19-8338-2_48

4. Taştan, M. (2018). IoT Based Wearable Smart Health Monitoring System. IEEE Transactions on Consumer Electronics, 14(1), 343-350. Doi: 10.18466/cbayarfbe.451076

5. Hamim, M., Paul, S., Hoque, S. I., Rahman, M. N. (2019). IoT-Based Patient Monitoring System for Remote Healthcare. International Journal of Engineering and Technology. 7(1),162-167. Doi:10.1109/ICREST.2019.8644514

6. Moghadas, E., Rezazadeh, J. (2020). An IoT patient monitoring based on fog computing and data mining: Cardiac arrhythmia usecase. 11. 100251Doi: https://doi.org/10.1016/j.iot.2020.100251

7. Gómez, J., Oviedo, B., Zhuma, E.(2016). Patient Monitoring System Based on Internet of Things. IEEE Transactions on Biomedical Circuits and Systems. 83, 90-97. Doi: https://doi.org/10.1016/j.procs.2016.04.103

8. Vasan, C.R.S., Charan, G., Babu, C. (2020). An IoT based SMART patient health monitoring system.18(3), 1657~1664.Doi: 10.11591/ijeecs.v18.i3.pp1657-1664

9. Kumar, A., Singh, J. (2017). IoT Based Elderly Healthcare Monitoring System. 6(1), 753-758. Doi: https://doi.org/10.22214/ijraset.2022.44645

10. Malik S. A., Bhuiyan, M. Z. H. (2019). IoT-Based Patient Monitoring System for Public Health. IEEE Transactions on Emerging Topics in Computing.7, 114-125. Doi: https://doi.org/10.3390%2Fhealthcare10101993

11. Vahdati, M., Hamlabadi, K.G., Saghiri, A. M. (2020). IoT-Based Patient Monitoring System Using Blockchain Technology. Sensors, 19(15),141-170. Doi:https://doi.org/10.1016/bs.adcom.2018.12.002.

12. Bhardwaj, V., Joshi, R., and Gaur, A. M.(2021). IoT-Based Patient Monitoring System for COVID-19 Patients. 8(7), 5274-5285. Doi: https://doi.org/10.1007%2Fs42979-022-01015-1

13. Shaik, T., Tao, X., Higgins, N., Li, L., Gururajan, R., Zhou, X., Acharya, U. R. (2023). Remote patient monitoring using artificial intelligence: Current state, applications, and challenges. 24(5),1513-1523. Doi: https://doi.org/10.1002/widm.1485

14. Yogaraj, A., Ezilarasan, M.R., Anuroop, R.V., Sivanthiram, C.S., Thakur, S. K. (2017). IoT Based Smart Healthcare Monitoring System for Rural/Isolated Areas.114(12), 679-688, Doi: https://www.ijpam.eu/

15. Garbhapu, V.V., Gopalan, S. (2021). IoT Based Low Cost Single Sensor Node Remote Health Monitoring System. IEEE Access. 9(1), 116930-116944. Doi: https://doi.org/10.101_6/j.procs.2017.08.357

16. Waleed, M., Kamal, T., Um, T.W., Hafeez, A. (2020). Unlocking Insights in IoT-Based Patient Monitoring:

Methods for Encompassing Large-Data Challenges. 23(15), 6750 Doi: https://doi.org/10.3390%2Fs23156760

17. Gupta, N.(2023). IoT-Based Patient Monitoring System for Sports Medicine. 67(1), 1-10. Doi: http://dx.doi.org/10.13140/RG.2.2.13004.18568

18. Gowthami, D., Aakash, K., Mohammed, A., Santhosh, A.V. (2023) Smart Animal Health Monitoring System using IoT. Animals. 05. 247-259. Doi: https://www.irjweb.com/Smart%20Animal%20Health%20M

19. Malathi, M., Muniappan, A., Rajagopal, B. R., Misra, P. K. (2021). A smart healthcare monitoring system for patients using IoT and cloud computing. 57(2), 1175-1186. Doi: http://dx.doi.org/10.1063/5.0126275

20. Joo, M., Kang,D.Y., Kang,M.S., Kim,H.C.(2022). IoT-Based Patient Monitoring System Using Edge Computing. 9(10), 8179-8189. Doi: http://dx.doi.org/10.3390/app14010414

## AUTHORS



**Vihaan Ravishankar** is a student currently pursuing his Bachelor of Technology (B. Tech) degree in Electrical and Electronics Engineering. Ravishankar's main fields of study are electronics and electrical engineering, with interests in the domains of Internet of Things (IoT) and embedded systems. He has skills in programming languages like C, C++, Python and familiarity with platforms such as Arduino and Raspberry Pi. His senior year project involves designing and prototyping a IoT based HVAC System.

Corresponding Author E-mail: vihaan.ravishankar@gmail.com



**Mohammad Ibrahim** is a passionate student pursuing his degree in Electronics and Electrical Engineering. Mohammad's expertise spans programming languages like C++ and Python, the latter of which he has applied specifically to data science applications. He is also proficient in AutoCAD software for computer-aided design and worked on an intelligent HVAC system project.

E-mail: mohammad.ibrahim9956@gmail.com



**Dr. Anjali Jain** is currently working as Associate Professor in Amity University Uttar Pradesh, Noida. She completed her Ph. D. in Electrical Engineering from AUUP in 2022. She received her M. Tech. degree in Electrical Engineering from YMCA Institute of Engineering (presently J.C. Bose University of Science and Technology, YMCA, Faridabad) in 2006 and B.E. in Electrical Engineering from MDU, Rohtak, Haryana, India in 2002. Her teaching interests include Analytics & Computing Methodologies in Power Systems, HVDC and EHVAC Transmission, Utilization of Electrical Energy, Electric Power Generation, Transmission and Distribution,. She has published several research papers in reputed international journal and conference proceedings.



**Neelam Verma** is currently working as Assistant Professor in Amity University Uttar Pradesh, Noida. She is pursuing her PhD from Department of Electrical Engineering, Delhi Technological University, Delhi, India. Her major area of research includes Nonlinear Control system, Controller designing for underactuated systems, Industrial Automation, Optimization Techniques.

Multimedia Tools and Applications
https://doi.org/10.1007/s11042-024-20585-1

# Effi-CNN: real-time vision-based system for interpretation of sign language using CNN and transfer learning

**Pranav**[1] · **Rahul Katarya**[1]

Received: 5 September 2022 / Revised: 13 May 2023 / Accepted: 26 December 2024
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2025

**Abstract**
The deaf and mute population has difficulty conveying their thoughts and ideas to others. Sign language is their most expressive mode of communication, but the general public is callow of sign language; therefore, the mute and deaf have difficulty communicating with others. A system that can correctly translate sign language motions to speech and vice versa in real time is required to overcome this communication barrier. Effi-CNN, a vision-based Sign Language Recognition (SLR) system employing transfer learning with EfficientNetB2, is proposed in this paper. We have also developed a system covering sign gestures to text in real time. Our approach was evaluated on eight publicly available datasets, including the Massey University gesture dataset, the ArSL2018 dataset, the MNIST-ASL dataset, and others. When comparing our results to state-of-the-art algorithms, the experimental findings showed that our proposed work is more successful than existing models. Our proposed model delivered an accuracy of 99.90 on MNIST Dataset. The results show that our Effi-CNN surpasses most currently existing solutions and can categorize many gestures with a low error rate.

**Keywords** Convolutional Neural Network (CNN) · Deep Learning (DL) · Sign Language (SL) · Transfer Learning

## 1 Introduction

Not everyone can hear or talk; to connect with society, such people have adopted unique communication processes rather than their voices. People with hearing and speech difficulties use hand gestures and accompanying motions to express their intended ideas [1]. Sign Language (SL) is a set of actions organized according to syntax, grammar, semantics, pragmatics, and morphology. For persons who are deaf or mute, SL happens to be

✉ Rahul Katarya
rahuldtu@gmail.com

Pranav
prrnnv@gmail.com

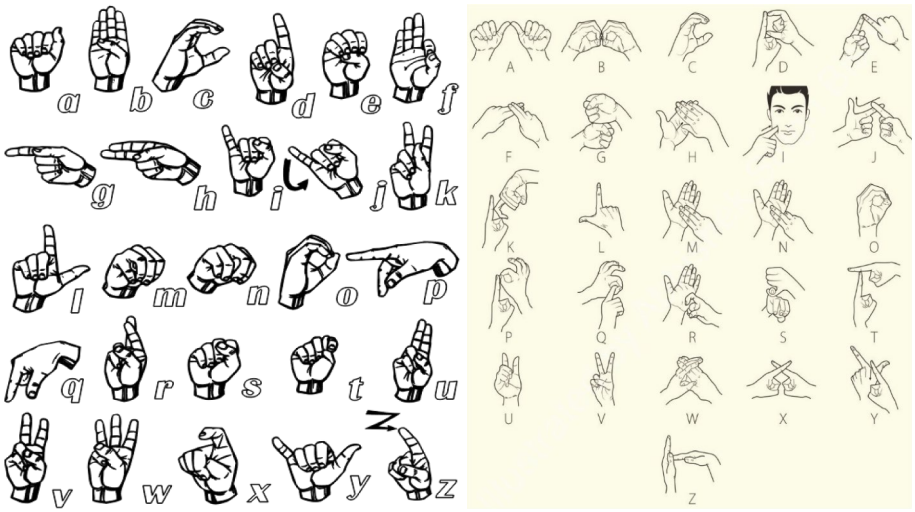
[1] Big Data Analytics and Web Intelligence Laboratory, Computer Science & Engineering Department, Delhi Technological University, New Delhi, India

the natural way of communication [2]. There is no universal Sign Language (SL), and SLs have evolved spontaneously as diverse groups of people engage with one another. Different nations and areas utilize other SLs, and practically every country has an official SL [3]; examples include Indian SL (ISL), Arabic SL (ArSL), American SL (ASL), etc. Even in countries with the same spoken language, the sign languages used are different, e.g. British SL and ASL. In Fig. 1, it can be seen that the alphabet 'Q' of ASL [4] looks similar to the alphabet 'U' of Indian SL (ISL) [5]. Just like spoken languages, SLs differ in terms of most linguistic aspects, like grammar structure [6]. One specific SL may differ from another in the essential alphabetical representation, as shown in Fig. 1, including the shape of the sign and the way the sign is performed [3, 7, 8]. In this work, datasets from three different SLs are chosen for study.

Literature prominently cites 2 types of techniques for SLR systems: vision-based techniques (VBT) and sensor-based gesture techniques (SBT); other than these, some other systems employing RADAR waves [9], speakers [10], EEG sensors[11–14], Wi-Fi modules [15–18], RF waves [19], etc. have also been used for the same.

The VBT is an artificial intelligence-based solution to sign gesture identification challenges that use cameras and image processing techniques. Many researchers have used a single camera to record indications as the default method. Edge-based active contours [20], frame subtraction [21], wavelet transform [22], and other approaches are used to segment the signer's hand from gesture images. Some researchers have employed RGB-D sensors to gather depth data of the signer's hand and body, such as Microsoft Kinect [23–26]. Another device mentioned in the literature [27, 28] that records hand motions on a 3D axis is the leap motion sensor. The detailed photos produce 3D data of the hand and body of the signer. Many critical metrics may be calculated using this data, including hand direction, wrist orientation, and joint angles.

In SBT, wearable devices consisting of sensors such as Flex Sensors, Motion Processing units, and pressure/contact sensors have been utilized for SLR in real-world scenarios, such




(a) (b)

**Fig. 1** Signs for fingerspelling of English alphabets (**a**) in the ASL [4], (**b**) in the ISL [5]

as harsh lighting and intricate backgrounds. These economical, lightweight sensors and the possibility to process the data on a low-power device such as an Arduino microcontroller make them viable solutions. However, as the number of gestures in a system increase, the chances of the being similar to another gesture also increase, which may affect the system's recognition accuracy.

Gestures are categorized based on whether they consist of movement of arms by the signer, called static and dynamic signs. Our paper deals with the recognition of static gesture signs. A significant problem with SL gestures is that the shape of one sign may be similar to that of another sign in terms of finger bend and wrist position; hence, the two signs may seem identical when done. This ambiguity might lead to incorrect categorization and low accuracy. Another problem is signer reliance. Because an individual cannot maintain all of their hands/fingers, the same differences in the capture of values can be seen by repeating the same sign many times.

Hence, the major contributions of this paper are as follows:

- Transfer learning-based vision-based system Effi-CNN proposed for recognition of SL gestures.
- A framework is proposed for SL gesture recognition to assist the deaf and hard-of-hearing people
- Experimental results indicate improved performance on five different SL datasets as compared to the existing SLR systems.

The remainder of this work is divided into the following sections. Section 2 introduces some related work in the field; Sect. 3 describes the system's components. Section 4 provides the exact implementation details of our model; Sect. 5 covers the datasets used and a comparative analysis of the proposed model's performance to those reported in the literature. Section 6 covers some discussion about the practical aspects of the SLR process—finally, Sect. 7 details the conclusions and future work.

## 2 Related work

Some of the existing work in literature in the domain of SLR using vision-based techniques are shown in Table 1.

The research gaps and learnings from the above Table 1 can be summarised into the following points:

1. Transfer learning is an effective method for SL classification problems since it provides superior performance and decreased training time since these models are already trained for superior performance on much larger classification problems of a similar category.
2. If the weights of pre-trained models are frozen, the models tend to be less accurate on the specific task since the models are more generalized.
3. Using RoI image segmentation results in improved classification accuracy due to removing unwanted features from the data and reducing training time.
4. Data augmentation has proven to be a highly effective method for making the model more robust to noise and improving overall performance.
5. Deep learning models are more accurate than machine learning models for image classification tasks.

**Table 1** Existing vision based techniques for Sign Language Recognition

| Ref | Methodology | Classifier | Conclusion | Dataset |
|---|---|---|---|---|
| [29] | In this paper, authors have performed data augmentation by flipping, rotating (0°−30°), and brightness changes on the images. The image features were extracted using pre-trained CNN models ResNet50, MobileNet-V2 architectures, and a combination of the two | ResNet50, MobileNet-V2 | Trained for less epochs, leading to reduced performance. Accuracy increase was observed by using a combination of both models | ArSL 2018 [38, 39] |
| [30] | To overcome issues with classes containing different sample sizes, they used Synthetic Minority Oversampling Technique, after which their model's accuracy increased. After normalization and standard scaling, custom CNN model with 7 convolution layers was used for classification | CNN | Synthetic oversampling works better than minority oversampling, and minority under-sampling, showcasing the advantages of using data augmentation | ArSL 2018 |
| [31] | After image pre-processing and data normalization, the data was passed to a CNN model with 4 convolution layers. Compared to system trained on reduced size database, using the complete database, and training for more epochs, the testing accuracy increased to 97.6%, however, this significantly increased the training time by almost 9 times | CNN | Increasing the number of layers in the CNN model increased significantly; however, the number of images used for the training process didn't greatly impact the model's performance | ArSL 2018 |
| [32] | After resizing images to 32*32, data augmentation using horizontal flipping of the images, data normalization, and removing noisy images, 20,227 and 12,480 images were used for training and testing, respectively, of two custom CNN models with different layers | CNN | In their implementation, they only used 20,227 for the training set, had they used about 40,000 images from the overall set, their overall accuracy might have been better than 96.6% | ArSL 2018 |
| [33] | To overcome the imbalanced number of images in classes, they used only 1000 randomly selected images from each class before performing data augmentation and normalization. The data was then classified using an EfficientNetB4 based CNN classifier trained for 30 epochs | EfficientNetB4 | Their system took 10 h to train the model, which is considerably high; still, they attained an accuracy of 95%, which is low compared to other models in the study | ArSL 2018 |

**Table 1** (continued)

| Ref | Methodology | Classifier | Conclusion | Dataset |
|---|---|---|---|---|
| [34] | They used a pre-trained VGG16 model for feature extraction from the images before passing it to fully connected dense layers of the CNN. GridSearchCV was used to find optimal values for hyperparameters tuning | VGG16 | This dataset achieved a considerably low accuracy of 94.33% compared to other models | ArSL 2018 |
| [35] | They used Sobel operator to find the 'absolute gradient magnitude' of the images for segmentation of the input images before passing to a 3 different transfer learning-based models, out of which VGGNet achieves the highest classification accuracy | VGGNet, GoogleNet, ALexNet, | Freezing weights of the pre-trained models led to the model not being trained for the actual dataset; otherwise, the accuracy of their system might have been better than 97% | ArSL 2018 |
| [36] | Initially, they resized the images to 28*28 and then applied the Gabor filter with 4 different sizes and 4 orientations to obtain spectral representations, which were then passed to a CNN classifier for recognition | CNN | The usage of Gabor filters increased the dataset by a factor of 16, which led to a high recognition accuracy of 99.05 on this dataset | ArSL 2018 |
| [37] | After under-sampling to equalize the number of images per class, they performed data augmentation before passing it to pre-trained VGG16, and ResNet152 models and trained for 100 epochs with the original weights frozen | VGG16, ResNet152 | Even though it attained a validation accuracy of 99.6%, it increased by only 0.1 percent from epoch 60 to 100 | ArSL 2018 |
| [40] | After image pre-processing, data augmentation, and standard scaling, the data was passed to a VGG-Net16 model for training model. The authors only trained their model for very few epochs because the pre-trained model's converged very rapidly | VGG16 | Their model underperforms for '0', 'N', 'W' gestures; if the model were trained for more epochs, it might have improved the classification accuracy | Massey [50, 51] |
| [41] | In this paper, a pre-trained Google Net architecture trained on the ILSVRC2012 dataset has been used for sign gesture classification | GoogleNet | The model's scalability remains a concern as their model provides validation accuracy of 98% with 5 letters and 75% with 10 letters | Massey |

**Table 1** (continued)

| Ref | Methodology | Classifier | Conclusion | Dataset |
|---|---|---|---|---|
| [42] | The input data was passed to 3 models: A custom CNN model, pre-trained VggNet16 model, and Inception-V3 model. It was found that the pre-trained models outperform the custom CNN model by ~ 4%. Further, by using data augmentation, accuracy increased by ~ 3.5% | VGG16, Inception-V3 | Since the images were passed to 2 different models, and the fused feature set was passed to PCA for feature extraction, the selected features from the fused set resulted in high recognition accuracy of 99.63% | Massey |
| [43] | Single Shot Multibox Detector was used to detect and crop hand region before background subtraction on cropped image. 10 different sets of random background images were added to make the classifier more robust. The pixel values were then rescaled to be in 0–1 range, and passed to a Wide Residual Network (WRN) classifier | WRN-CNN | The performance increase of pre-trained WRN over a baseline CNN model further signifies the merits of pre-trained models trained over large datasets | Massey |
| [44] | This paper uses a multi-channel CNN-based feature fusion technique as a classifier. The input image is passed to one pipeline of the CNN as it is, and a Gabor filter is applied to the image before passing it to the second CNN pipeline for feature extraction | CNN | Ensemble model for feature extraction removes the high frequency spatial dependencies usually found in models trained with RGB images only and improves overall performance | Massey |
| [45] | They used Media-pipe hands API to extract coordinates of 21 joint points on the hand of the signer. Then distances between the sets of these joint points were calculated for a total of 190 distance vectors; similarly, 210 angle vectors were calculated using the relative distance between the joint points. Also, angle-based vectors were calculated to fetch the hand's orientation. All these feature vectors were then passed to a SVM and a light Gradient Boosting Machine (GBM) for training and classification. Testing accuracies of 99.39% and 97.80% were obtained for SVM, and GBM | SVM, GBM | Their system achieves a per sample recognition time of only 14 ms, which is highly adequate for a real-time SLR system given that their system also achieves high recognition accuracy of 99.39% on this dataset and, respectively, high accuracies on 2 more ASL datasets | Massey |

**Table 1** (continued)

| Ref | Methodology | Classifier | Conclusion | Dataset |
|---|---|---|---|---|
| [46] | They used a RCNN-based model to detect hand regions in the image, from which 5 different regions were cropped out. After this, Gaussian and Salt-and-pepper noise were added to the images in 4 different varying proportions. The original image, cropped images, and noise-added cropped images were passed to 3 different Restricted Boltzmann Machines, whose outputs were fused for final output generation | RCNN | Multi-channel data fusion and multiple transformations of each image results in high accuracy of 99.31%. The drawback with their system is that it was run for 10,000 epochs and the high design complexity | Massey |
| [47] | YCbCr based hand region segmentation and Convex Hull Algorithm based hand shape detection for image cropping was used before passing to a CNN classifier | CNN | Hand location-based segmentation reduces useless features and training time, as well as increases accuracy | Massey |
| [28] | VGG16 for classification, GridSearchCV for optimal hyperparameters tuning | VGG16 | Model's lower performance w.r.t. to others models can be due to the absence of specified feature extraction steps | Massey |
| [48] | After cropping and resizing images based on skin segmentation, they used Local-Binary-Pattern (LBP), Histogram-of-Oriented-Gradients (HOG) as feature extractors, and CNN, SVM as a classifier. They ran multiple combinations of these classifiers and feature extractors | CNN, SVM | The HOG-LBP-SVM model provides better accuracy than the CNN-SVM and other models, however, increased implementation complexity remains a concern | Massey |
| [49] | The authors propose a SLR system based on CNN models built using transfer learning from image multiple classifier systems. The input images will be resized, and background subtraction and convex hull algorithm for image processing will be used | CaffeNet, GoogleNet, AlexNet, VGGNet | Only a model is proposed for future implementation and does not include any implemented model | Massey |

**Table 1** (continued)

| Ref | Methodology | Classifier | Conclusion | Dataset |
|---|---|---|---|---|
| [52] | In this paper, sign images were acquired from the video feed, and after pre-processing to extract the hand region and convert the RGB image to HSV, the CNN model was used for training and classification | CNN | Their model converges in only 5 epochs with a high learning rate of 0.1; it can be assumed that model's accuracy of 98.55% can increase further with an improved learning rate | ASL [52, 54] |
| [53] | In this paper, the authors have used the same technique as in [52]; however, they achieve better test accuracy on the dataset by using a different configuration of the CNN layers. They further evaluated their model's performance on 2 more datasets, as shown | CNN | The increasing number of epochs to 100 allowed them to get increased accuracy of 99.41% on the dataset [52], 99.48% on the dataset [55], and 99.38% on their own synthetic dataset [56] | ASL [52]<br>ASL [55]<br>ASL [56] |
| [57] | Human skin colour based hand portion segmentation using HSV values was done before passing the images to a SVM classifier for training and classification. The system works on mobile phones, making it available to the masses for real-time SLR | SVM | However, the system's accuracy is only 80–90 percent for different alphabets, which makes it inadequate compared to other systems in our study | ASL [57] |
| [58] | A Region Proposal Network was used for ROIs in the image then pooling was used to bring all those segments to the same size before passing it to the CNN classifier | CNN | Using an attention-based model to detect the hand region in the image reduces the processing load for the classification part and increases accuracy by 5% | ASL [64] |
| [59] | After converting the RGB images to grayscale and using the Canny edge detector for edge detection [60], the image was passed to multiple feature extractors, including HOG, LBP, PCA, and ORB [61] feature detector, which produces a 32-dimension feature vector. After converting these features to a bag-of-words [62] model, multiple machine learning-based classifiers were used to train and classify the gestures | SVM, logistic regression, Naïve Bayes, KNN, Random forest, MLP, | A combination of PCA with a multi-layer-perceptron (MLP) classifier achieved highest classification accuracy | ASL [64] |

**Table 1** (continued)

| Ref | Methodology | Classifier | Conclusion | Dataset |
|---|---|---|---|---|
| [60] | They used multiple CNN based classifiers using transfer learning methods. It was found that the pre-trained models offer much better classification accuracy than custom CNN models, with the ResNet50 model achieving highest accuracy | VGG16, InceptionNet, and ResNet50 | This paper found that the performance of InceptionNet was lower than baseline CNN model, and Resnet50 outperformed VGG16 | ASL [64] |
| [61] | After resizing, the image was broken down into smaller patches. These patches were passed to the Vision Transformer (ViT) encoder consisting of the attention layer and MLP. The output of this layer was then passed to another MLP for final classification | MLP | Due to the low accuracy of 80.6%, this model cannot be considered an adequate solution for the problem under consideration | ASL [64] |
| [62] | This paper used image resizing to 50*50 pixels, data augmentation using Gaussian noise adding and image rotation. The images were then passed to a CNN model with 4 convolution layers for training and classification | CNN | Their model provides excellent testing accuracy of 99.89% but takes 0.236 s per image for recognition, which makes it unsuitable for a real-time system | ASL [64] |
| [63] | This paper tested the transfer learning model for gesture classification using VGG16 model with and without Bimodal Distribution Removal (BDR). It was found that the use of BDR leads to significant improvement in recognition accuracy | VGG16 | Only 9% of images from the dataset were used for training, and they got 68% testing accuracy. Hence the results obtained cannot be extrapolated to the entire dataset | ASL [64] |
| [42] | A combination of VggNet, and Inception-V3 were used for this dataset too | VggNet, Inception-V3 | 99.02% accuracy on dataset [64] | ASL [64] |
| [65] | Image data was passed to two pre-defined models with and without augmentation. It was observed that LeNet obtained lowest accuracy; accuracy increased by 6.74% in the CapsNet model fed with the original dataset, while accuracy increased by another 6.62% using augmented data | LeNet, CapsNet | The results in this paper further strengthen the case that data augmentation used during training can increase the system's overall performance | MNIST [71] |

**Table 1** (continued)

| Ref | Methodology | Classifier | Conclusion | Dataset |
|---|---|---|---|---|
| [36] | Gabor filters with CNN | CNN | 99.9 accuracy on MNIST dataset signifies the advantage of using Gabor filter for vector creation for image classification | MNIST |
| [66] | A custom CNN model was proposed with 3 convolution layers and ran for 10 epochs over the training set | CNN | The system was evaluated using only 10 images, which cannot be extrapolated for complete dataset | MNIST |
| [67] | In this paper, numeric data from xls files of the dataset were converted to PNG files and then passed to two models: MobileNet and InceptionV3. These models were trained, and then the trained model was used for gesture classification on a smartphone | MobileNet, InceptionV3 | The average recognition time for a gesture was about 2.42 s, which implies that the model was not sufficiently optimized to run on low-performance hardware | MNIST |
| [68] | In this paper, Particle Swarm Optimisation (PSO) was used to optimize the parameters affecting the classification accuracy of a CNN-based image classifier. The use of PSO results in an accuracy of 99.53% | CNN | The time required for the system to run for all the combinations of parameters is quite long and not scalable for large datasets | MNIST |
| [69] | This model uses a CNN classifier incorporating 7 convolution layers and 2 dense layers for image classification | CNN | The model's accuracy can be considered to be low | MNIST |
| [70] | In this paper, discrete wavelet transform was used to extract features from the images, to be classified by the CNN module | CNN | Using wavelet transform to extract low-level features from images results in high accuracy | MNIST |

6. The classification accuracy improves by increasing the number of images per class to train the models.
7. After a certain high level of accuracy has been achieved, increasing the complexity of the model to suit the task better provides diminishing improvements.
8. In the cases where the number of epochs for which the model is trained is deficient (less than 20), the model usually did not learn enough features from the data to classify the test images accurately.

## 3 Preliminary

In order to recognize Sign language gestures [72], transfer learning [73] using modified pre-trained EfficientNetB2 [74] based CNN [75, 76] models have been used. The architecture of these elements has been discussed in the following sections. Also, detailed information used in our proposed implementation, like OpenCV [77], Keras [78], TensorFlow [79] etc., have also been discussed below.

a) Sign Language: It is the way of communication by using hand gestures, body language, and facial expressions. It is the main form of contact for the Deaf and Hard-of-Hearing community. SL gestures are categorized based on the movement type as static gestures that remain invariant to motion for the period of the gesture and dynamic gestures that are centred on movement. Signs can be performed using one hand or both hands, static or dynamic, and other categorizations of signs are possible, as shown in Fig. 2. Figure 2, the type 0 sign indicates a two-handed dynamic sign in which both hands are active. Type 1 is a two-handed dynamic sign with only the primary hand performing the motion.

b) Convolution neural networks are one of the most widely used deep learning approaches for visual imagery analysis. In comparison to other image classification techniques, CNN requires less pre-processing. CNN is a feed-forward artificial neural network comprising many neural network layers, each having multiple neurons. The network learns feature extraction filters that are hand-modelled in other paradigms. Convolution, pooling, flattening, and fully linked layers are the four types of processes in a CNN. The convolutional layer frequently captures low-level characteristics like colour, edges, and gradient direction. The pooling layer reduces the spatial dimension of the convolved features. This process minimizes the needed computing time for working with the data through dimensionality alleviation. It also has the benefit of keeping dominating characteristics throughout model training, which are invariant in position and rotation. Higher-level characteristics can be categorised after the input picture has been processed. As a result, the picture is flattened into a one-dimensional vector. The flattened output is sent into a fully connected layer in CNN, which classifies the features retrieved by the convolutional/pooling layer into labels. The model may offer the probability of predicting items in the picture after training using SoftMax [81] classification.

c) Transfer Learning. Transfer learning [73] is a procedure that influences pre-trained models that are trained on an enormous and miscellaneous dataset to solve novel, related tasks with partial data. Transfer learning has assisted attain achievement in deep learning, with fine-tuning a pre-trained model providing an unlimited way for knowledge transfer. The process of transfer learning involves leveraging the knowledge gained from a source domain and applying it to a target domain. This is done by

reusing the learned features, such as weights and biases, from the pre-trained model and fine-tuning them to fit the new task. This approach reduces the need for large amounts of labeled data and decreases training time, while still attaining high performance. By utilizing pre-trained models, transfer learning can effectively handle domain adaptation, allowing the model to perform well on the target task.
d) EfficientNet. CNNs are frequently built at a fixed resource cost and then scaled up when more resources become available to improve accuracy. However, traditional model scaling strategies are extremely unpredictable. Some models are scaled in depth, while others are scaled in width. Some models simply use higher-resolution photos to achieve better outcomes. This method of arbitrarily scaling models necessitates manual adjustment and many man-hours, with little or no performance increase. EfficientNet scales up models using a simple yet effective compound coefficient approach. Compound scaling consistently increases each dimension with a predetermined set of scaling coefficients rather than randomly scaling up width, depth, or resolution. The authors created seven models [B0 to B7] of varied dimensions, surpassing most CNNs' performance and efficiency [82, 83].

## 4 Proposed system

Our proposed system consists of a pre-trained model, EfficientNetB2. Several design considerations were taken to get to the final architecture. Decisions on the pre-trained model, optimization techniques, and hyper-parameters to test and assess are more specific. After the base model, we added two dense layers at the end for classification. Weights of the base model pre-trained on the Imagenet dataset have been chosen in the model; also, the weights of the base model were not frozen and the model was fine-tuned on the dataset for domain adaptation. The optimizer used is Adamax with loss as categorical cross-entropy, a learning rate of 0.001 with a learning rate reduction factor of 0.5, and a momentum of 0.99. Dwell is set as true so that if the monitored metric does not improve on the current epoch, then set model_weights back to weights of the previous epoch. The include_top parameter was set as false to fetch outputs activation from the last pooling or convolution layer of the pre-trained model and feed them to a dense layer with 256 neurons, followed by a dropout layer with a dropout rate of 0.45, followed by another dense layer with softmax activation and several neurons equal to the number of classes in the input dataset. In total, the final model contains 344 layers and 8,144,414 Trainable parameters.

In this study, we also developed an application for real-time SLR. Images from webcam footage were captured using the OpenCV library. During runtime, three windows appear when the application is in real-time recognition mode. The first window is the picture capture window, which has a 192*192 pixel predetermined rectangle; the user must execute the signs inside the given area, as shown in Fig. 3 (a).

After capturing the picture from the predefined region, the cropped image is converted to HSV and displayed in the preview window (shown overlaid on the preview window) using user-defined parameters during runtime by using a third window. Then, the application supplies it to the pre-trained CNN model for classification. This study's model was trained using augmented data from multiple ASL datasets. Finally, the anticipated character appears in the image capture window's specified region. A third control window as shown in Fig. 3 (b) with track bars, is used for controlling the HSV threshold

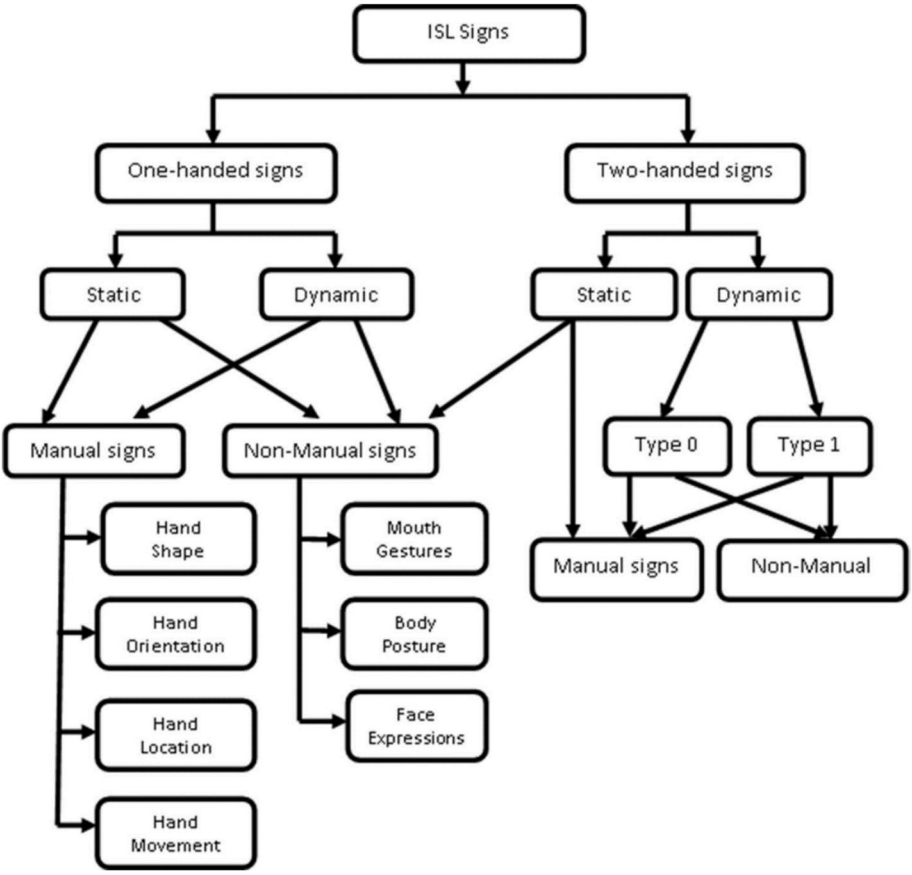



**Fig. 2** Systematic representation of the Hierarchy of sign language gestures [80]

for skin colour based feature extraction during runtime; this is needed since, in different lighting conditions, skin colour appears different and appears to be of a different hue. Figure 4 has the following overall steps:

i. Collect and pre-process the data.
ii. Split the dataset into training and validation sets.
iii. Transfer Learning model: Initialize the EfficientNetB2 model with pre-trained weights on the image dataset. Then, freeze all the layers except the last few layers to prevent overfitting on the small dataset and replace the last layer with a new fully connected layer with the appropriate number of output classes.
iv. Fine-tuning: Train the model on the training set using an appropriate optimizer and loss function. Monitor the model's performance on the validation set and adjust hyper-parameters such as learning rate and batch size as needed.
v. Train and evaluate the model.

The data flow diagram depicting the training process of the system is shown in Fig. 5, while the data flow of the real-time SLR system is shown in Fig. 4.

## 5 Experiment analysis

This section describes the SLR datasets in our implementation and implementation details.

### 5.1 i. Datasets

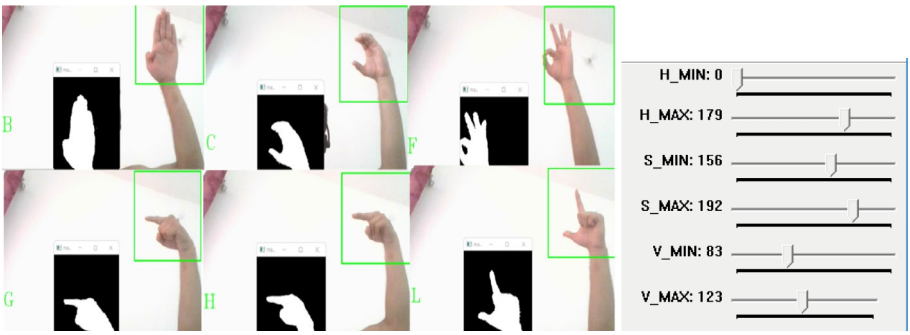
Our experiments are carried out on eight different Sign language datasets, including six ASL datasets and one Arabic and Pakistani SL dataset. Details about these datasets are shown in Table 2. If the train and test sets were not already in segregated form for all the datasets, then the data was split in an 80–20 ratio for training-test sets, respectively. After this, for all the datasets, the training set was split into a 95–5 ratio for training and validation data, respectively.

### 5.2 ii. Dataset limitations

1) In most cases, the whole datasets were recorded with the same background conditions at the same place.
2) There is insufficient orientation, posture, noise, lightning, and zooming variations in the dataset's classes.
3) The number of signers involved in dataset creation was usually very low.
4) In some datasets, the number of images differs from one class to another, e.g. in [38, 71].

### 5.3 iii. Classification prediction results and performance comparison with other models

Eight separate SL datasets were used in the evaluation process of our proposed model. The details of these datasets are presented in the methodology section. If the train data




(a) (b)

**Fig. 3** **a** First two windows of the application, (**b**) HSV parameters control trackbars window

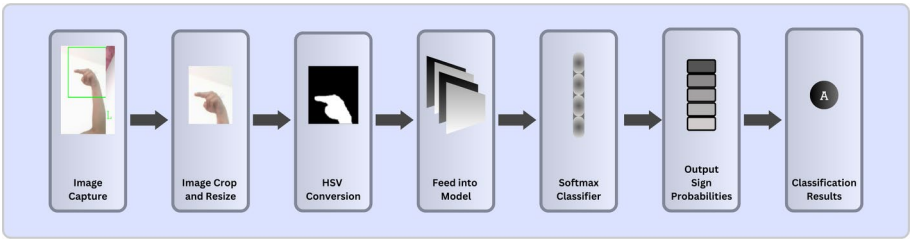



**Fig. 4** Data flow model of the proposed system for real-time testing of the system

set is extensive, we have chosen to trim the training dataset to 1000 samples per class and the validation dataset to 50 samples per class to reduce training time. All the images were resized to 50*50 pixels before feeding to the model. According to the results of the experiments, our proposed Effi-CNN model outperformed almost other existing models in test accuracy. The performance of our model on the datasets, as mentioned earlier, is shown below in Figs. 6, Fig. 7, Fig. 8, and Fig. 9. A comparison of Effi-CNN with the best approach in the literature and training time taken is presented in Table 3.

Compared to other systems, our system takes much less time for the training process because the model's weights did not have to be trained from scratch, resulting in a faster convergence rate.

The training was halted during our training process when validation accuracy increased beyond 99.5% to avoid overfitting.

## 6 Discussion

I. Processing Time: Our model was run on a Google Colab online execution platform for training and testing the datasets. The time required for training our system for all the datasets is specified in Table 3. In practical applications, the processing time taken for gesture recognition is essential for a better user experience. In our system, we have developed a real-time application to capture gesture input using a webcam and perform some pre-processing on the RAW image before sending it to pre-trained CNN model for classification. The processing delays were found to be within a tolerable range of 0.01 s on an 'AMD A8-7410 APU Quad Core 2.2 GHz system, with 4 GB DDR-3 RAM running at 1200 MHz', which is adequate for a real-time system. It is worth noting that only one CPU core is activated during the recognition process, with about 25 percent load on the core.

II. Skin Tone Impact: In our application, in the pre-processing stage, the RGB image acquired from the camera is converted using HSV filter to extract only human skin coloured parts from the image. However, people have a wide variety of skin tones, and various skin tones absorb light differently, affecting the accuracy of gesture detection. As a result, different HSV values were required for successful segmentation in various illumination scenarios.

III. Impact of Intense Body Movements: During the real-time gesture recognition process, the recognition rate of the system is limited by the frame rate of the camera in consideration, which is usually 24–30 fps on a laptop webcam. If abrupt motions are made, captured images might be blurred due to motion artefacts, which leads to distorted image, which is rendered useless for the recognition process.

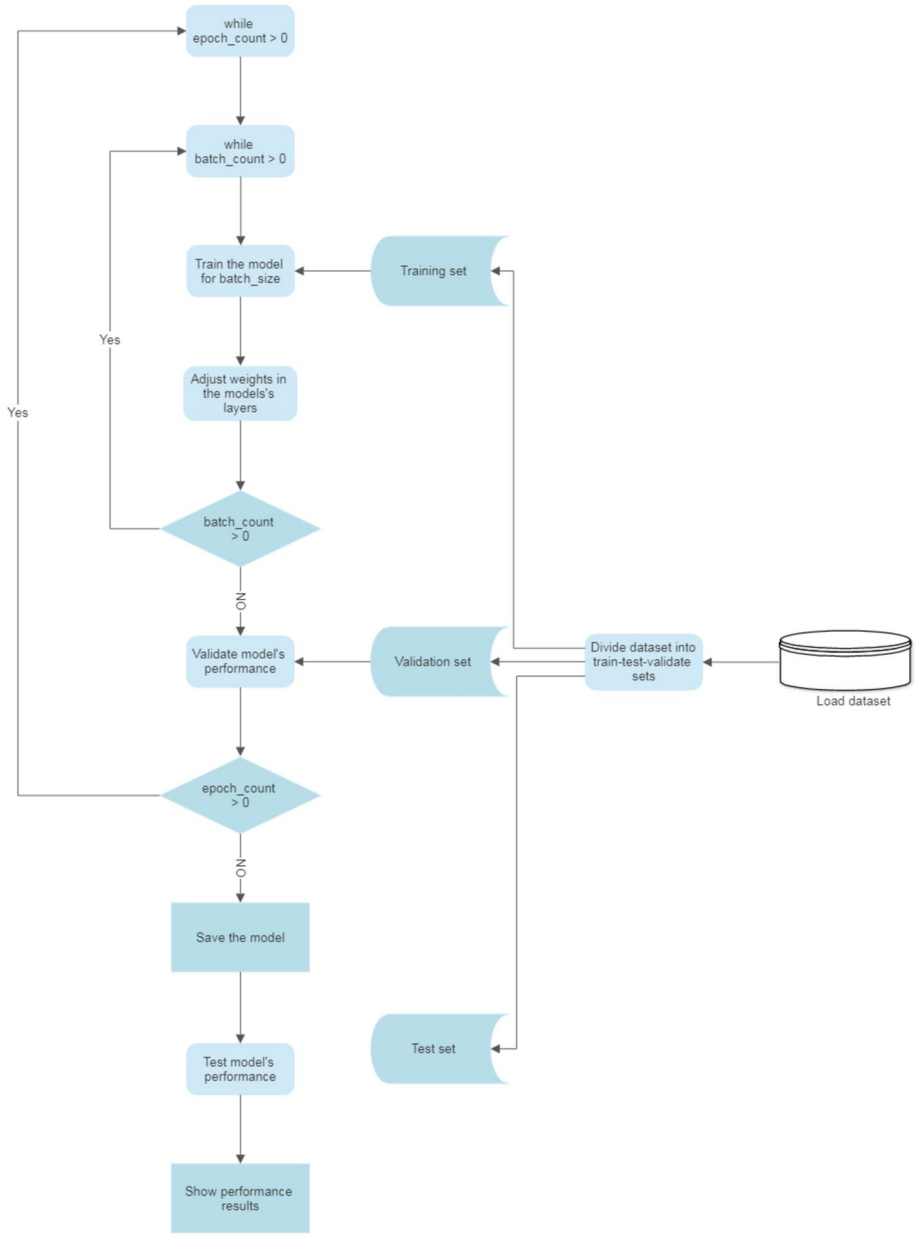



**Fig. 5** Data flow model of the proposed system for training and testing accuracy of the system

IV. Background and lighting conditions: Both background colour and lighting conditions play a crucial role in image classification tasks. If the background is uniformly coloured and does not match the signer's skin colour, extracting the hand segment from the image is more accessible. If any skin-color objects are present in the ROI, the system may fail to segment the hand area from the image.

**Table 2** Details of datasets included in our study

| Dataset | Sign Language | Type of dataset | Vocabulary size | Number of Signers | Total number of images |
|---|---|---|---|---|---|
| ArSL [38, 39] | Arabic SL | Grayscale Images | 32 | 40 | 54,049 |
| Massey [50, 51] | American SL | RGB Images | 36 | 5 | 2,425 |
| [52, 54] | American SL | HSV-BW Images | 26 | 1 | 52,000 |
| [55] | American SL | HSV-BW Images | 44 | 1 | 105,600 |
| [56] | American SL | HSV-BW Images | 26 | 1 | 104,000 |
| PSL [57] | Pakistani SL | RGB Images | 37 | NA | 1,509 |
| [64] | American SL | RGB Images | 29 | NA | 87,000 |
| MNIST [71] | American SL | CSV file | 24 | NA | 27,455 |

*NA* data not provided

In dim light situations, the captured images contain more noise than those captured in well-lit conditions. This leads to poor edge detection and error prone segmentation, which leads to poorer recognition accuracy.

## 7 Conclusions and future work

A hand gesture recognition approach for vision-based sign language recognition is provided in this study. A deep learning-based Effi-CNN model based on transfer learning using EfficientNetB2 as the base model is presented for vision-based static gesture

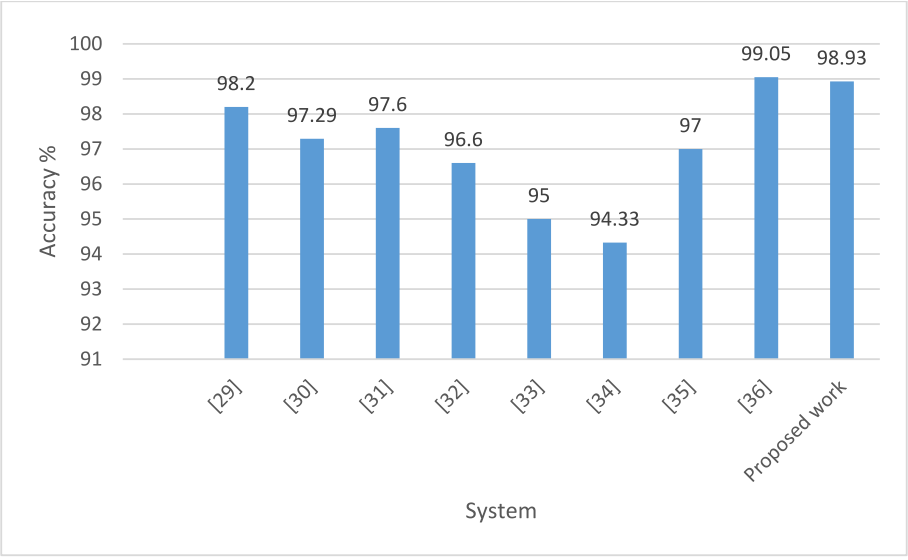



**Fig. 6** Performance of different models on ArSL dataset

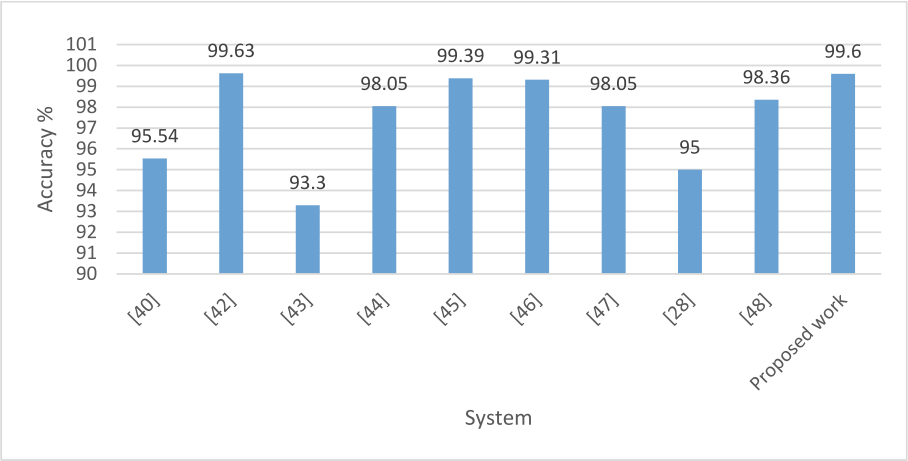



**Fig. 7** Performance of different models on Massey University dataset

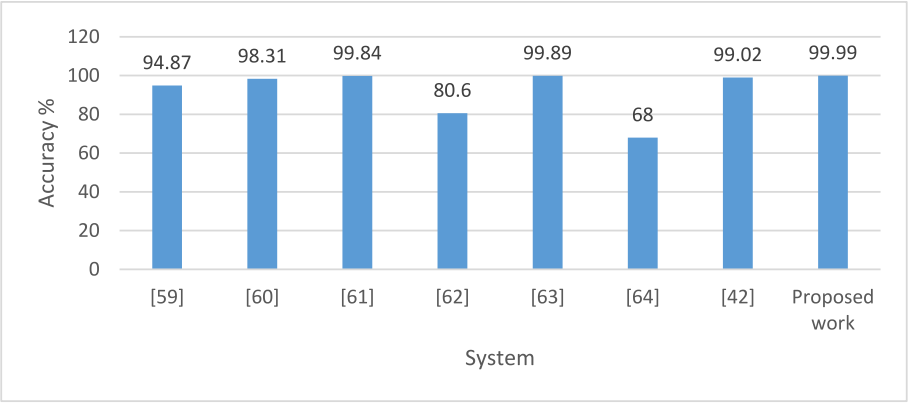



**Fig. 8** Performance of different models on dataset [64]

identification. External gear such as Kinect sensors is not required for the proposed vision-based paradigm, making it more feasible. The capacity of this research to recognize the indicators of distinct SL with good recognition results over state-of-the-art techniques is a noteworthy addition. The performance of this work is assessed using 8 publically available SL datasets. According to rigorous experimental assessment, the Effi-CNN model either outperforms most existing systems or produces competitive results compared to other state-of-the-art approaches.

The model attempts to classify the input image based on relevant trained patterns learned during training by calculating the probability of a class being most likely. To correctly translate hand-sign-spelled words and sentences into written language, we must continue to create future work that includes all punctuation marks and all types of joint letter representation. This study, on the other hand, will serve as a point for researchers working in the field of SLR. The system may be used as a translator for sign language and non-sign language communication, as well as for human–computer/machine interface and robot

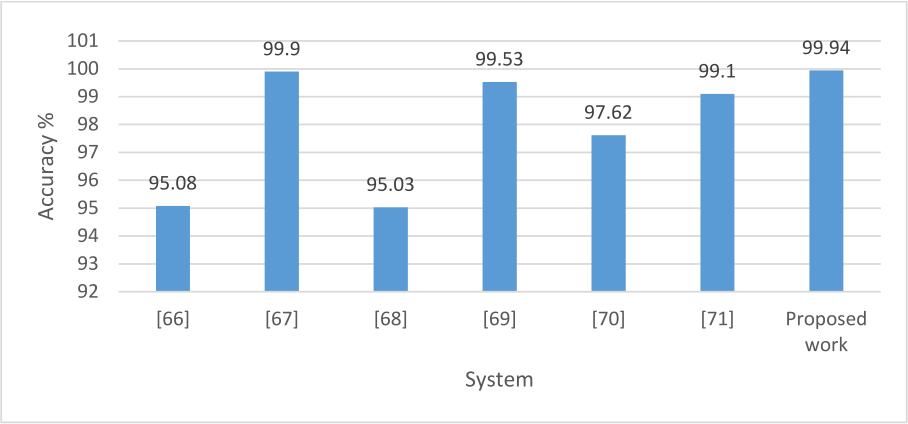



**Fig. 9** Performance of different models on the MNIST dataset

**Table 3** Training details of Effi-CNN on the datasets and comparison with the best approach present in literature

| Dataset | Accuracy of our system | Best accuracy in literature | Training time |
|---|---|---|---|
| MNIST [71] | 99.88—99.94 | 99.90 | 12 min, 24 s |
| PSL [57] | 99.34 – 99.60 | 80–90 | 26 min, 48 s |
| ArSL2018 [38] | 98.93 | 99.05 | 44 min, 33 s |
| Massey [50] | 99.45 – 99.60 | 99.63 | 24 min, 8 s |
| [64] | 99.92 | 99.89 | 59 min, 11 s |
| [52, 54] | 99.49 – 99.61 | 99.41 | 33 min, 12 s |
| [55] | 99.62 | 99.48 | 59 min, 54 s |
| [56] | 99.29 – 99.31 | 99.38 | 36 min, 31 s |

control. Currently, our work focuses on static gesture recognition, and we also intend to recognize dynamic sign language gestures through video frames, which is a challenging task. Also, our present method is based on a definite zone of interest that has been designated. There would be no need for a predefined ROI if object detection was used to find the hand region, which is a must for practical applications. If the text output of the system can be translated to voice using a text-to-speech application such as Google's Cloud Text-To-Speech for a better experience.

**Data availability** Not applicable to this article as no datasets were generated or analyzed during the current study'.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

## References


1. National Institute on Aging. n.d. Hearing Loss: A Common Problem for Older Adults.[online] Available at: <https://www.nia.nih.gov/health/hearing-loss-common-problem-older-adults> [Accessed 24 May 2022].
2. Disabled World. (2022, April 7). Deaf Communication: Sign Language and Assistive Hearing Devices. Disabled World. Retrieved May 24, 2022 from www.disabled-world.com/disability/types/hearing/communication/
3. Ai-Media creating accessibility, one word at a time. 2021. Sign Language Alphabets From Around The World - ASL - Ai-Media. [online] Available at: <https://www.ai-media.tv/ai-media-blog/sign-language-alphabets-from-around-the-world/> [Accessed 24 May 2022].
4. En.wikipedia.org. Fingerspelling - Wikipedia. [online] Available at: <https://en.wikipedia.org/wiki/Fingerspelling> [Accessed 28 April 2022].
5. Islrtc.nic.in. Poster of the Manual Alphabet in ISL | Indian Sign Language Research and Training Center (ISLRTC), Government of India. [online] Available at: <http://www.islrtc.nic.in/poster-manual-alphabet-isl> [Accessed 26 April 2022].
6. Lifeprint.com. n.d. American Sign Language (ASL). [online] Available at: <https://www.lifeprint.com/asl101/pages-layout/grammar.htm> [Accessed 24 May 2022].
7. Ai-Media creating accessibility, one word at a time. 2022. Sign Language Alphabets From Around The World - ASL - Ai-Media. [online] Available at: <https://www.ai-media.tv/ai-media-blog/sign-language-alphabets-from-around-the-world/> [Accessed 27 May 2022].
8. Lackner A (2017) Comprehending the Complexities of Sign Language. [online] The Wire. Available at: <https://thewire.in/culture/amazing-complexity-sign-language> [Accessed 24 May 2022].
9. Li B, Yang J, Yang Y, Li C, Zhang Y (2021) Sign Language/Gesture Recognition Based on Cumulative Distribution Density Features Using UWB Radar. IEEE Trans Instrum Meas 70(1–13):2511113. https://doi.org/10.1109/TIM.2021.3092072
10. Yang Y, Li J, Li B et al (2022) MDHandNet: a lightweight deep neural network for hand gesture/sign language recognition based on micro-doppler images. World Wide Web. https://doi.org/10.1007/s11280-021-00985-1
11. Wang P, Zhou Y, Li Z, Huang S, Zhang D (2021) Neural Decoding of Chinese Sign Language With Machine Learning for Brain-Computer Interfaces. IEEE Trans Neural Syst Rehabil Eng 29:2721–2732. https://doi.org/10.1109/TNSRE.2021.3137340
12. Malaia EA, Borneman SC, Krebs J, Wilbur RB (2021) Low-Frequency Entrainment to Visual Motion Underlies Sign Language Comprehension. IEEE Trans Neural Syst Rehabil Eng 29:2456–2463. https://doi.org/10.1109/TNSRE.2021.3127724
13. AlQattan D, Sepulveda F (2017) Towards sign language recognition using EEG-based motor imagery brain computer interface. 2017 5th International Winter Conference on Brain-Computer Interface (BCI), 5–8, https://doi.org/10.1109/IWW-BCI.2017.7858143.
14 Kubicek E, Quandt LC (2019) Sensorimotor system engagement during ASL sign perception: An EEG study in deaf signers and hearing non-signers. Cortex 119:457–469. https://doi.org/10.1016/j.cortex.2019.07.016
15 Zhang L, Zhang Y, Zheng X (2020) WiSign: Ubiquitous American Sign Language Recognition Using Commercial Wi-Fi Devices. ACM Trans Intell Syst Technol 11(3):24. https://doi.org/10.1145/3377553
16. Shang J, Wu J (2017) A Robust Sign Language Recognition System with Multiple Wi-Fi Devices. In Proceedings of the Workshop on Mobility in the Evolving Internet Architecture (MobiArch '17). Association for Computing Machinery, New York, NY, USA, 19–24. https://doi.org/10.1145/3097620.3097624
17. Shang J, Wu J (2017) A Robust Sign Language Recognition System with Sparsely Labeled Instances Using Wi-Fi Signals. 2017 IEEE 14th International Conference on Mobile Ad Hoc and Sensor Systems (MASS), 99–107, https://doi.org/10.1109/MASS.2017.41.
18 Ahmed HFT, Ahmad H, Narasingamurthi K, Harkat H, Phang SK (2020) DF-WiSLR: Device-Free Wi-Fi-based Sign Language Recognition. Pervasive Mob Comput 69:101289. https://doi.org/10.1016/j.pmcj.2020.101289
19 Gurbuz SZ et al (2021) American Sign Language Recognition Using RF Sensing. IEEE Sensors J 21(3):3763–3775. https://doi.org/10.1109/JSEN.2020.3022376
20. Amrutha K, Prabu P (2021) ML Based Sign Language Recognition System. Int Conf Innov Trends Inform Technol (ICITIIT) 2021:1–6. https://doi.org/10.1109/ICITIIT51526.2021.9399594
21. Zhang LG, Chen Y, Fang G, Chen X, Gao W (2004) A vision-based sign language recognition system using tied-mixture density HMM. In Proceedings of the 6th international conference on Multimodal

interfaces (ICMI '04). Association for Computing Machinery, New York, NY, USA, 198–204. https://doi.org/10.1145/1027933.1027967
22. Karami A, Zanj B (2011) Azadeh Kiani Sarkaleh, Persian sign language (PSL) recognition using wavelet transform and neural networks. Expert Syst Appl 38(3):2661–2667. https://doi.org/10.1016/j.eswa.2010.08.056
23. Wang CC et al (2021) Real-Time Block-Based Embedded CNN for Gesture Classification on an FPGA. IEEE Trans Circuits Syst I Regul Pap 68(10):4182–4193. https://doi.org/10.1109/TCSI.2021.3100109
24. Sincan OM, Keles HY (2020) AUTSL: A Large Scale Multi-Modal Turkish Sign Language Dataset and Baseline Methods. IEEE Access 8:181340–181355. https://doi.org/10.1109/ACCESS.2020.3028072
25. Plouffe G, Cretu A (2016) Static and Dynamic Hand Gesture Recognition in Depth Data Using Dynamic Time Warping. IEEE Trans Instrum Meas 65(2):305–316. https://doi.org/10.1109/TIM.2015.2498560
26. Sun C, Zhang T, Bao B, Xu C, Mei T (2013) Discriminative Exemplar Coding for Sign Language Recognition With Kinect. IEEE Transactions on Cybernetics 43(5):1418–1428. https://doi.org/10.1109/TCYB.2013.2265337
27. Avola D, Bernardi M, Cinque L, Foresti GL, Massaroni C (2019) Exploiting Recurrent Neural Networks and Leap Motion Controller for the Recognition of Sign Language and Semaphoric Hand Gestures. IEEE Trans Multimedia 21(1):234–245. https://doi.org/10.1109/TMM.2018.2856094
28. Rho E, Chan K, Varoy EJ, Giacaman N (2020) An Experiential Learning Approach to Learning Manual Communication Through a Virtual Reality Environment. IEEE Trans Learn Technol 13(3):477–490. https://doi.org/10.1109/TLT.2020.2988523
29. Alnuaim Abeer, Zakariah Mohammed (2022) Human-Computer Interaction with Hand Gesture Recognition Using ResNet and MobileNet. Comput Intell Neurosci 8777355(16):2022. https://doi.org/10.1155/2022/8777355
30. Alani, AA, Cosma G (2021): ArSL-CNN: a convolutional neural network for Arabic sign language gesture recognition. Loughborough University. Journal contribution. https://hdl.handle.net/2134/16878787.v1 https://doi.org/10.11591/ijeecs.v22.i2.pp1096-1107
31 Latif G, Mohammad N, AlKhalaf R, AlKhalaf R, Alghazo J, Khan M (2020) An automatic Arabic sign language recognition system based on deep CNN: an assistive system for the deaf and hard of hearing. Int J Comput Digital Syst 9(4):715. https://doi.org/10.12785/ijcds/09041824
32. Hayani S, Benaddy M, El Meslouhi O, Kardouchi M (2019) Arab sign language recognition with convolutional neural networks. In: 2019 International Conference of Computer Science and Renewable Energies (ICCSRE). Agadir, Morocco, pp 1–4. https://doi.org/10.1109/ICCSRE.2019.8807586
33 Zakariah M, Alotaibi YA, Koundal D, Guo Y, Mamun Elahi M (2022) (2022) Sign Language Recognition for Arabic Alphabets Using Transfer Learning Technique. Comput Intell Neurosci 4567989:15. https://doi.org/10.1155/2022/4567989
34. Nurnoby MF, El-Alfy ESM, Luqman H (2020) Evaluation of CNN Models with Transfer Learning for Recognition of Sign Language Alphabets with Complex Background. 2020 International Conference on Innovation and Intelligence for Informatics, Computing and Technologies (3ICT), 1–6, https://doi.org/10.1109/3ICT51146.2020.9311989.
35. Duwairi RM, Halloush ZA (2022) Automatic recognition of Arabic alphabets sign language using deep learning. International J Electr Comput Eng (2088–8708) 12(3)
36. Luqman H, El-Alfy ESM, BinMakhashen GM (2021) Joint space representation and recognition of sign language fingerspelling using Gabor filter and convolutional neural network. Multimed Tools Appl 80:10213–10234. https://doi.org/10.1007/s11042-020-09994-0
37. Saleh Y, Issa GF (2020) Arabic Sign Language Recognition through Deep Neural Networks Fine-Tuning. Int J Online Biomed Eng (iJOE) 16(05):71–83. https://doi.org/10.3991/ijoe.v16i05.13087
38. Latif G, Alghazo J, Mohammad N, AlKhalaf R, AlKhalaf R (2018) Arabic Alphabets Sign Language Dataset (ArASL). [online] Mendeley Data. Available at: <https://data.mendeley.com/datasets/y7pckrw6z2/1> [Accessed 24 May 2022].
39 Latif G, Mohammad N, Alghazo J, AlKhalaf R, AlKhalaf R (2019) ArASL: Arabic Alphabets Sign Language Dataset. Data in Brief 23:103777. https://doi.org/10.1016/j.dib.2019.103777
40. Masood S, Thuwal H, Srivastava A (2018). American Sign Language Character Recognition Using Convolution Neural Network. https://doi.org/10.1007/978-981-10-5547-8_42.
41. Garcia B, Viesca SA (2016) Real-time American sign language recognition with convolutional neural networks. Convolutional Neural Netw Visual Recognition 2:225–232
42. Rajan RG (2021) Transfer-learning analysis for sign language classification models. Turkish J Comput Math Educ (TURCOMAT) 12(9):1423–1433

43. Kania K, Markowska-Kaczmar U (2018). American Sign Language Fingerspelling Recognition Using Wide Residual Networks. https://doi.org/10.1007/978-3-319-91253-0_10.
44 Chevtchenko SF, Vale RF, Macario V, Cordeiro FR (2018) A convolutional neural network with feature fusion for real-time hand posture recognition. Appl Soft Comput 73:748–766. https://doi.org/10.1016/j.asoc.2018.09.010
45. Shin J, Matsuoka A, Hasan MAM, Srizon AY (2021) American Sign Language Alphabet Recognition by Extracting Feature from Hand Pose Estimation. Sensors 21(17):5856. https://doi.org/10.3390/s21175856
46. Rastgoo R, Kiani K, Escalera S (2018) Multi-Modal Deep Hand Sign Language Recognition in Still Images Using Restricted Boltzmann Machine. Entropy 20:809. https://doi.org/10.3390/e20110809
47. Taskiran M, Killioglu M, Kahraman N (2018) A Real-Time System for Recognition of American Sign Language by using Deep Learning. 2018 41st International Conference on Telecommunications and Signal Processing (TSP), 1–5, https://doi.org/10.1109/TSP.2018.8441304.
48. Nguyen HBD, Do HN (2019) Deep Learning for American Sign Language Fingerspelling Recognition System. 2019 26th International Conference on Telecommunications (ICT), 314–318, https://doi.org/10.1109/ICT.2019.8798856.
49. Salian S, Dokare I, Serai D, Suresh A, Ganorkar P (2017). Proposed system for sign language recognition. 058–062. https://doi.org/10.1109/ICCPEIC.2017.8290339
50. Massey.ac.nz (2012) Massey University. [online] Available at: <https://www.massey.ac.nz/~albarcza/gesture_dataset2012.html> [Accessed 24 May 2022].
51. Barczak ALC, Reyes NH, Abastillas M, Piccio A, Susnjak T (2011) A new 2D static hand gesture colour image dataset for ASL gestures. Res Lett Inform Math Sci 15:12–20 (http://iims.massey.ac.nz/research/letters/)
52. Jeny JRV, Anjana A, Monica K, Sumanth T, Mamatha A (2021) Hand Gesture Recognition for Sign Language Using Convolutional Neural Network. 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI), 1713–1721, https://doi.org/10.1109/ICOEI51242.2021.9453072.1713-1721). IEEE.
53. Kasapbaşi A, Elbushra AE, Omar AH, Yilmaz A (2022) DeepASLR: A CNN based human computer interface for American Sign Language recognition for hearing-impaired individuals. Comput Methods Programs Biomed Update 2:100048. https://doi.org/10.1016/j.cmpbup.2021.100048.2022
54. GitHub (2018) GitHub - rrupeshh/Simple-Sign-Language-Detector: Simple Sign Language Detector. [online] Available at: <https://github.com/rrupeshh/Simple-Sign-Language-Detector> [Accessed 24 May 2022].
55. GitHub (2018) GitHub - EvilPort2/Sign-Language: A very simple CNN project.. [online] Available at: <https://github.com/EvilPort2/Sign-Language> [Accessed 24 May 2022].
56. GitHub (2020) GitHub - Ahmed-KASAPBASI/Success_Team_ASL: Project of Deep Learning (ASL). [online] Available at: <https://github.com/Ahmed-KASAPBASI/Success_Team_ASL> [Accessed 24 May 2022].
57. Imran A, Razzaq A, Baig IA, Hussain A, Shahid S, Rehman TU (2021) Dataset of Pakistan Sign Language and Automatic Recognition of Hand Configuration of Urdu Alphabet through Machine Learning. Data in Brief 36:107021. https://doi.org/10.1016/j.dib.2021.107021
58. Zhang S, Zhang Q (2021) Sign language recognition based on global-local attention. J Vis Commun Image Represent 80:103280. https://doi.org/10.1016/j.jvcir.2021.103280
59. Sharma A, Mittal A, Singh S, Awatramani V (2020) Hand Gesture Recognition using Image Processing and Feature Extraction Techniques. Procedia Comput Sci 173:181–190. https://doi.org/10.1016/j.procs.2020.06.022
60. Bantupalli K, Xie Y (2018) American sign language recognition using deep learning and computer vision. In: 2018 IEEE International Conference on Big Data (Big Data). Seattle, WA, USA, pp 4896–4899. https://doi.org/10.1109/BigData.2018.8622141
61. Reddy Karna SN, Kode JS, Nadipalli S, Yadav S (2021) American Sign Language Static Gesture Recognition using Deep Learning and Computer Vision. 2nd International Conference on Smart Electronics and Communication (ICOSEC), 1432–1437, https://doi.org/10.1109/ICOSEC51865.2021.9591845
62. Elsayed N, ElSayed Z, Maida AS (2022) Vision-Based American Sign Language Classification Approach via Deep Learning. arXiv preprint arXiv:2204.04235.
63. Rakshit A. Smart learners: Choosing What to Learn Using Bimodal Distribution Removal. Availabale at: http://users.cecs.anu.edu.au/~Tom.Gedeon/conf/ABCs2018/paper/ABCs2018_paper_10.pdf
64. Kaggle.com (2018) ASL Alphabet. [online] Available at: <https://www.kaggle.com/grassknoted/asl-alphabet> [Accessed 24 May 2022].

65. Bilgin M, Mutludoğan K (2019) American Sign Language Character Recognition with Capsule Networks. 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), 1–6, https://doi.org/10.1109/ISMSIT.2019.8932829.
66. Makhashen GMB, Luqman HA, El-Alfy EM (2019) Using Gabor filter bank with downsampling and SVM for visual sign language alphabet recognition. 2nd Smart Cities Symposium (SCS 2019), 1–6, https://doi.org/10.1049/cp.2019.0188.
67. Rathi D (2018) Optimization of transfer learning for sign language recognition targeting mobile platform. arXiv preprint arXiv:1805.06618.
68 Fregoso J, Gonzalez CI, Martinez GE (2021) Optimization of Convolutional Neural Networks Architectures Using PSO for Sign Language Recognition. Axioms 10(3):139. https://doi.org/10.3390/axioms10030139
69. Hasan MM, Srizon AY, Sayeed A, Hasan MAM (2020) Classification of Sign Language Characters by Applying a Deep Convolutional Neural Network. 2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT), pp. 434–438, https://doi.org/10.1109/ICAICT51780.2020.9333456.
70. Beniwal R, Nag B, Saraswat A, Gulati P (2022) Static Hand Sign Recognition Using Wavelet Transform and Convolutional Neural Network. In: Luhach AK, Poonia RC, Gao XZ, Singh Jat D (eds) Second International Conference on Sustainable Technologies for Computational Intelligence. Advances in Intelligent Systems and Computing, vol 1235. Springer, Singapore. https://doi.org/10.1007/978-981-16-4641-6_13
71. Kaggle.com (2020) Sign Language MNIST | Kaggle. [online] Available at: <https://www.kaggle.com/c/sign-language-mnist> [Accessed 24 May 2022].
72. En.wikipedia.org (2022) Sign language - Wikipedia. [online] Available at: <https://en.wikipedia.org/wiki/Sign_language> [Accessed 24 May 2022].
73. TensorFlow (n.d.) Transfer learning and fine-tuning | TensorFlow Core. [online] Available at: <https://www.tensorflow.org/tutorials/images/transfer_learning> [Accessed 24 May 2022].
74. Tan M, Le Q (2019) Efficientnet: rethinking model scaling for convolutional neural networks. In International conference on machine learning. PMLR, pp 6105–6114
75. Scince Direct (n.d.) Deep convolutional neural network–based image classification. [online] Available at:<https://www.sciencedirect.com/topics/engineering/convolutional-neural-network> [Accessed 24 May 2022].
76. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient- based learning applied to document recognition. Proc IEEE 86(11):2278–2324
77. OpenCV (n.d.) About - OpenCV. [online] Available at: <https://opencv.org/about/> [Accessed 24 May 2022].
78. Chollet F others (2015) Keras. Available at: https://github.com/fchollet/keras. Software available from https://keras.io
79. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M, Ghemawat S (2016) Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467. Software available from https://tensorflow.org.
80 Wadhawan A, Kumar P (2019) Sign language recognition systems: a decade systematic literature review. Arch Comp Methods Eng 28:785–813
81. Google Developers (2020) Multi-Class Neural Networks: Softmax | Machine Learning Crash Course | Google Developers. [online] Available at: <https://developers.google.com/machine-learning/crash-course/multi-class-neural-networks/softmax> [Accessed 24 May 2022].
82. Agarwal V (2020) Complete Architectural Details of all EfficientNet Models. [online] Medium. Available at: <https://towardsdatascience.com/complete-architectural-details-of-all-efficientnet-models-5fd5b736142> [Accessed 24 May 2022].
83. Sarkar, D.: a-comprehensive-hands-on-guide-to-transfer-learning-with-real-world-applications-in-deep-learning-212bf3b2f27a @ towardsdatascience.com, https://towardsdatascience.com/a-comprehensive-hands-on-guide-to-transfer-learning-with-real-world-applications-in-deep-learning-212bf3b2f27a


**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.

# Empirical Analysis of Nature-Inspired Algorithms for Autism Spectrum Disorder Detection Using 3D Video Dataset

Aneesh Panchal*[1], Kainat Khan†[2], and Rahul Katarya‡[2]

[1]Department of Computational and Data Sciences, Indian Institute of Science, Bangalore, Karnataka−560012, INDIA
[2]Big Data Analytics and Web Intelligence Laboratory, Department of Computer Science & Engineering, Delhi Technological University, Delhi−110042, INDIA

arXiv:2501.01202v1 [cs.LG] 2 Jan 2025

**Abstract**

Autism Spectrum Disorder (ASD) is a chronic neurodevelopmental disorder symptoms of which includes repetitive behaviour and lack of social and communication skills. Even though these symptoms can be seen very clearly in social but a large number of individuals with ASD remains undiagnosed. In this paper, we worked on a methodology for detection of ASD from 3-dimensional walking video dataset, utilizing supervised machine learning (ML) classification algorithms and using nature-inspired optimization algorithms for feature extraction from dataset. The proposed methodology involve the classification of ASD using supervised ML classification algorithm and extracting important and relevant features from the dataset using nature-inspired optimization algorithms. We also included the ranking coefficients for finding the initial leading particle. This selection of particle significantly reduces the computation time and hence, improving the total efficiency and accuracy for the ASD detection. To evaluate the efficiency of the proposed methodology, we deployed various combination of classification algorithm and nature-inspired algorithms resulting in an outstanding classification accuracy of 100% using random forest classification algorithm and gravitational search algorithm for feature selection. The application of the proposed methodology with different datasets would enhance the robustness and generalizability of the proposed methodology. Due to high accuracy and less total computation time, the proposed methodology will offer a significant contribution to the medical and academia fields, providing a foundation for future research and advancements in ASD diagnosis.

***Keywords*** - Autism spectrum disorder; Machine learning; Nature-inspired algorithm; Ranking coefficient; 3D walking human video dataset.

## 1 Introduction

In today's world, the growing awareness and understanding of various neurodevelopmental disorders serves a basis of diagnosis and early treatment of these disorders. Advances in technology and research in genetics, neuroimaging and cognitive and brain sciences provides useful insights related to these neurodevelopmental disorders. These advances helps in the development of new and updated treatment process for these disorders. These disorders not only impact the individuals but their family and surroundings as well. These disorders require parental and social

*Email: aneeshpanchal840850@gmail.com
†Email: khankainat388@gmail.com
‡Email: rahuldtu@gmail.com

support for the individual to tackle these disorders. The support may consist of academic councilling, behavioral and social communication councilling. The healthcare sector and academia sector collaboratively contribute to the minimization of these cases through early detection and proper treatment. Along with proper treatment, emotional support is also required from the family so that the individual do not felt left out from the society. Hence, early diagnosis and treatment of neurodevelopmental disorders are required.

### 1.1 Motivation

The cases neurodevelopmental disorders are increasing day by day. According to World Health Orgainization (WHO) [2], there are approximately 30 crore individuals having depression, 6 crore individuals having dementia, 4.75 crore individuals having schizophrenia and 2.1 crore individuals having bipolar disorder around the world. These neurodevelopmental disorders lead to mental, emotional and social disabilities.

ASD is one of the most common neurodevelopmental disorder which affects approximately 1 in 68 children in United States and 1 in 500 chilren in India. ASD cases are increasing day by day, most recent research indicates that approximately 1.5% of the global population suffering from autism spectrum. Due to it's widespread presence, many cases of ASD remains untreated in initial stages which worsen their conditions. Hence, accurate and timely treatment of ASD is very crucial in early stages. The availability demand of larger dataset including behaviour observations, medical records, cognitive assessments and genetic information, is very much required and will play a crucial role in understanding ASD. By analyzing these datasets deeply, researchers can develop much more sufficient algorithms required for ASD detection and treatment which can be data-driven approaches such as ML algorithms. These algorithms may increase accuracy and efficiency in treatment of ASD.

Attempts for early and accurate detection of ASD and it's treatment is going on in academia and research domain. ASD detection using advanced mathematical technologies such as ML and optimization shows promising results in ASD detection. It is well known that ML algorithms are very useful in detecting patterns in large datasets and hence increasing prediction accuracy. Using power of these ML algorithms, we will create a new methodology to detect ASD using 3-dimensional walking video dataset which will help in early detection and treatment for the ASD affected individuals.

### 1.2 Major contributions

The main objective of this study is to detect ASD in children using supervised ML classification algorithms and nature-inspired optimization algorithm. Supervised ML algorithms are used for classification whether individual is ASD patient or not. Nature-inspired optimization algorithms used for feature extraction from the dataset. These classification and nature-inspired algorithms have proven their efficiency in predictions using high-dimensional and complex datasets. The main objective of this study can also be seen as detection of ASD from 3-dimensional walking video dataset of children. As walking is one of the fundamental skill developed in human being and also holds potential to indicate the the presence of ASD in an individual and hence can be used for ASD detection. Hence, the paper focuses on utilizing the dataset extracted from walking videos of the children to detect ASD. In addition to ML algorithms which are used for classification, we also used nature-inspired optimization algorithm for the feature selection from the large dataset. This paper have some of the modifications which are not done in any previous research on ASD detection which includes,

- Integration of nature-inspired optimization algorithms and the ranking of features using various ranking coefficients.
- Use of 3D walking videos of children as the dataset for ASD detection.

- Comparative study of different ML algorithms, ranking coefficients, and various nature-inspired algorithms.

By including these modifications, the proposed method aims to increase the efficiency and accuracy of the existing ASD detection models. Proposed methodology also aims to reduce the total computational time required during ASD detection. Hence, this research aims to contribute to the field of ASD detection and highlights the potential of ML classification algorithms and nature-inspired optimization algorithms for feature selection. These contributions not only make a base for further research in early detection of ASD but also contributes to the treatment of individuals of ASD.

### 1.3 Problem statement

Through a novel methodology that combines the ML classification algorithm with nature-inspired optimization algorithm for feature extraction we aim to contribute to the field of ASD detection. The nature-inspired optimization algorithms gained attention in recent times for feature extraction due to their efficiency in selecting important and relevant features along with reduction of total computation time significantly. Adding the ranking coefficients in the methodology to find the initial leading particles for the nature-inspired optimization algorithm increases the efficiency and reduces the total computation time of the method significantly.

### 1.4 Organization of the paper

The remaining paper is organized as follows, section 2 provides a comprehensive literature review of the concepts related to ASD detection using ML algorithms. Section 3 gives the preliminaries required for understanding the subsequent sections of the research paper. Section 4 presents the algorithm and methodology employed in this research paper. Section 5 describes the dataset specifications, data processing techniques and feature extraction methods utilizing nature-inspired optimization algorithms. Section 5 also presents the performance evaluation metrics and results obtained for each algorithm combination. Finally, section 6 concludes the paper through summarizing the key findings and suggesting potential areas for future research.

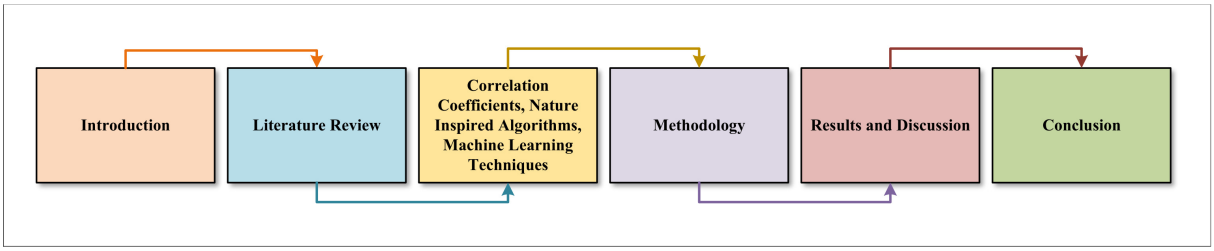



Fig. 1: Organization of the paper

## 2 Literature review

Early detection of ASD is very important, so that the individual gets the appropriate treatment for the disorder. Accuracy in detection of ASD still remains a big concern in field of medical treatment and requires high level assessment and expertise in this field. Some of the advancements in these are as follows,

- ASD detection instruments like Autism Diagnostic Observation Schedule (ADOS), Childhood Autism Rating Scale (CARS) and Autism Diagnostic Interview (ADI-R) have been developed in recent time for the detection of ASD in individuals [3, 5, 6, 7].

- Computerized advancements such as robotics, virtual reality and computer gaming shows promising results in developing and improving social communication skills in ASD detected individuals.

- Mathematical advancements in technology and ML optimization techniques such as Natural Language Processing (NLP), Long Short-Term Memory (LSTM), autoencoder and deep learning models prove their efficiency in detection of ASD from a provided dataset with signifiantly high accuracy as compared to traditional methodologies [3, 4, 7].

- The use of biological and physical indicators such as heartbeat data and eye gazing data, shows promising results in recent times in detecting ASD and also these can be easily implemented in practical scenarios. One of the system that is presently used for ASD detection is known as the Autonomous Social Orienting Training System (ASOTS) with Response to Name (RTN) mechanism [5, 6].

Table 1 shows some of the most recent advancements in ASD detection using machine learning. As can be seen from the table, computerized technologies have gained importance in the detection and treatment of ASD. Recent research proves the efficiency of nature-inspired optimization algorithms to select the important features which will reduce total computation time and helps in detection of ASD.

The main focus of this research paper is to find solutions to some of the problems exists in recent advancements in detection of ASD which are handling high-dimensional dataset, feature extraction technique and reduction in total computation time. These problems with existing methodologies can be explained as follows,

- Handling high-dimensional dataset is one of the major problem as most of the latest ASD dataset contains videos and image data which are of very high-dimensions. That is, these datasets contains large number of features. Hence, these dataset increases computational complexity and decreases the performance and accuracy. Therefore, this research paper aims to develop a method using which we can easily analyze and use the high-dimensional datasets which leads to improvement in ASD accuracy.

- Feature extraction is also one of the major problem as choosing relevant and important features or, information from the dataset plays a crucial role in increasing accuracy and decreasing the computational time. Hence, feature selection algorithm must be good enough to select important features out of very large features which finally increase the accuracy and reduce computational time.

- Computational time problem is one of the most crucial problem due to the demand of faster results in medical field, which will be tackled in this research paper. Using proper feature selection algorithm and correlation coefficients used here to select the initial leading particle for selected features. Hence, reducing the total computational time significantly.

# 3 Preliminary

This section provides essential details and background concepts required for understanding the subsequent sections of the research paper. This section gives details about various correlation coefficients, nature-inspired algorithms and supervised classification machine learning algorithms.

## 3.1 Correlation coefficients

In statistics, correlations gives a medium to find the relationship between 2 different variables which can be random or deterministic. Correlation refers to the change in 1 variable due to the change in other variable. Correlation can be seen as representation of the monotonic association

Table 1: Literature Review of recently published research papers in related fields.

| Year | Author | Aim | Dataset used | Methodology | Results | Limitations |
|---|---|---|---|---|---|---|
| 2022 | Malathi and Kannan [9] | Adaptive Whale Optimization-based Support Vector Machine for Prediction of ASD | Child, Adult, and Adolescent ASD Dataset from UCI Machine Learning Website. | Adaptive Whale Optimization based SVM is divided into 3 phases Exploitation, Exploration, and Classification. | Accuracy for the dataset, Child - 90.753%; Adolescent - 90.385%; Adult - 89.347%; | Better algorithms already exist. Example, Modified Grasshopper Optimization algorithm. |
| 2022 | Sahu and Verma [11] | Classification of Autistic Spectrum Disorder Using Deep Neural Network with Particle Swarm Optimization | Child and Adolescent ASD Dataset from UCI Machine Learning Website. | Particle Swarm Algorithm (PSO) for feature selection and Multi-Layer Perceptron (MLP) and Deep Neural Network (DNN) for classification | Accuracy for the PSO-DNN model, Adolescent - 98.077%; Child - 100%; | Only can be applied to textual data It is prone to different types of datasets such as video dataset. |
| 2022 | Malathi and Kannan [12] | Robust Artificial Bee Colony Optimization-Based Classifier for Prediction of ASD | Child, Adult, and Adolescent ASD Dataset from UCI Machine Learning Website. | Robust Artificial Bee Colony Optimization (RABC) for feature selection and Decision Tree based Gradient Boosting (DTGB) for classification | Accuracy for the dataset, Child - 82.877%; Adolescent - 82.692%; Adult - 84.233%; | Better algorithms already exist. Example, Modified Grasshopper Optimization algorithm. |
| 2021 | Saranya and Anandan [8] | Autism Spectrum Prognosis using Worm Optimized Extreme Learning Machine (WOEM) Technique | Kaggle Dataset on Autism Spectrum Disorder. | Glow Worm Algorithm with Extreme Learning Machine (Single Hidden Layer) | Accuracy - 99%; Better results and faster than particle swarm Optimization algorithm | Only can be applied to textual data It is prone to different types of datasets such as video datasets. |
| 2021 | Elshoky et al. [13] | Machine Learning Techniques based on Feature Selection for Improving ASD | Child and Adult ASD Dataset from UCI Machine Learning Website. | Binary Firefly Feature Selection Algorithm then applied various ML Algorithm for classification. | Accuracy - 100%; Binary Firefly for feature selection and Logistic Regression. | Only can be applied to textual data It is prone to different types of datasets such as video dataset. |
| 2021 | Balakrishnan et al. [14] | Detecting Autism Spectrum Disorder with Sailfish Optimization | Child and Adult ASD Dataset from UCI Machine Learning Website. | Random Opposition based Learning Sailfish Optimization (ROBL-SFO) with 10-fold cross validation and SVM for classification | Accuracy for the dataset, Child - 97.30%; Adult - 94.20%; Efficiency increases by 3.07%. | Not validated for Adolescence dataset. No surety of good results in video/audio dataset. |
| 2020 | Jayakumaran and Sweetlin [10] | Optimized Neural Network with Inertia Weight Variation of PSO for the detection of Autism | Kaggle Dataset on Autism Spectrum Disorder for Toddlers. | Artificial Neural Network and Particle Swarm Optimization with Inertia Weight updation | Accuracy - 97.05%; Better results than standard feed forward neural network | Not validated against multiple high-dimensional datasets Recurrent Neural Network. |
| 2020 | Zheng et al. [3] | Detect Autism Spectrum Disorder in Saccadic Eye Movement | Dataset consists of 59 children from French schools (Restricted Data) | NLP Based eye movement data and predictions using CNN and LSTM. | Accuracy - 84%; CNN and LSTM based model. | Dataset consists of only 59 participants which is very less. |
| 2020 | Lee et al. [4] | Detection of Infants with Autism Spectrum Disorder Using AutoEncoder Feature Representation | Seoul National University Bundang Hospital ASD Data (Restricted Data) | Autoencoder based feature extraction and joint optimization method for automated detection. | Accuracy - 68.18%; Bidirectional long short-term memory (BLSTM) model. | Comparatively very low accuracy in comparison to other methods proposed till now. |
| 2020 | Goel et al. [15] | Modified Grasshopper Optimization Algorithm for detection of ASD | Child, Adult and Adolescent ASD Dataset from UCI Machine Learning Website. | Modified Grasshopper Optimization Algorithm for feature selection. Used 2-3 principal components and train using Random Forest for classification. | Accuracy for the dataset, Child - 100%; Adolescent - 100%; Adult - 99.29%; | Convergence speed is slow for Grasshopper Optimization Algorithm, hence computational time is increased. |
| 2020 | Farrell et al. [16] | Combined Firefly Algorithm Random Forest (FARF) to Classify Autistic Spectrum Disorders | Child ASD Dataset from UCI Machine Learning Website. | Firefly algorithm for feature selection and Random Forest (RF) for classification | Accuracy - 94.32%; Better than RF with 90.78% accuracy. | Optimal number of Decision trees can't be found. Better algorithm already exists for ASD classification |
| 2020 | Alice Z. Guo [17] | Automated Autism Detection based on Characterizing Observable Patterns from Photos | Photo dataset by Wang et el. 'Revealing the world of autism through the lens of a camera'. | Repetitive Histogram Method for detection of ASD cases. | Accuracy - 85.8% | Deep Learning models and advanced feature extraction methods can be used for more accuracy. |
| 2020 | Li et al. [7] | Classifying ASD children with LSTM based on raw videos | Dataset consists of 272 children of primary and special ed. schools (Restricted Data) | Three layer Long Short-Term Memory Network is used for classification. | Accuracy - 93.4% | Feature extraction may reduce computation time and give satisfactory results. |
| 2019 | Abitha and Vennila [18] | Swarm based Symmetrical Uncertainity Feature Selection method for ASD | Child ASD Dataset from UCI Machine Learning Website. | Combination of Symmetrical Uncertainty and PSO used for feature selection and Naïve Bayes used for classification | Naïve Bayes, Accuracy - 86.75%; Artificial Neural Networks, Accuracy - 83.53%; | Not validated for Adult and Adolescence dataset. |
| 2018 | Samad et al. [19] | Autism behavioural Markers in Spontaneous Facial, Visual, and Hand Movement Response Data | Dataset consists of volunteering basis humans (Restricted Data) | Movement of cursor and Eye gazing on the cursor (Response time analysis) for detection of ASD | Poor correlation between eye gaze and hand movement in ones who suffers ASD | Theoretical based paper which can be further improved by using Machine Learning approaches. |

between 2 variables. In positive correlation, the changes in both variables are in same direction while in negative correlation, the changes are in opposite direction [20].

### 3.1.1 Pearson correlation coefficient

Pearson Correlation Coefficient (PCC), denoted by $\rho$, measures the correlation between two different variables within range $[-1, 1]$, where positive and negative value represents the respective correlation and zero indicates no correlation between variables. PCC is used to find the strength and direction of correlation between the variables and hence, can be used for selection of relevant features and selection of most useful observation from the dataset [21].

PCC between variables $x$ and $y$ can be calculated as follows,

$$\rho_{x,y} = \frac{cov(x, y)}{\sigma_x \sigma_y} \tag{1}$$

where, $cov(x, y)$ is the covariance between variables $x$ and $y$, and $\sigma_i$ is the standard deviation for the $i^{th}$ variable.

### 3.1.2 Spearman correlation coefficient

Spearman Correlation Coefficient (SCC), denoted by $r_S$, also measure the correlation between two different variables. SCC also enables to find out the degree of correlation between the two variables. SCC is less sensitive to the extreme values or, outliers. Hence, it is better to choose SCC than choosing PCC in case when outliers may be present in the dataset. SCC is also useful when dealing with ordinal values as it doesn't assume the linear relation between variables as assumed by PCC [21].

SCC between variables $x$ and $y$ can be calculated as follows,

$$r_S = \rho_{R_x, R_y} \tag{2}$$

where, $R_x$ and $R_y$ are the converted ranks.

### 3.1.3 Relief ranking coefficient

Relief ranking algorithm is an iterative algorithm that measure the relevance and power of the feature in the dataset. During every iteration, algorithm chooses the instances that are correlated with respect to feature space and hence, update their weight values accordingly. We call these weights as relief ranking coefficients. These weights are useful in finding the most relevant and informative attribute and hence, rank them accordingly using these weights. Relief ranking algorithm is very useful when we are dealing with high-dimensional datasets [22].

The weight vector for relief algorithm can be calculated as follows,

$$W_i = W_i - (x_i - \text{nearHit}_i)^2 + (x_i - \text{nearMiss}_i)^2 \tag{3}$$

where, $x$ is the feature vector, $W_i$ is the weight vector for $i^{th}$ attribute.

## 3.2 Nature-inspired algorithms

Nature-inspired algorithms are optimization algorithms used for various tasks including optimization, feature selection and many more. These are also known as bio-inspired algorithms and derived from various natural systems and phenomenon. These algorithms copies the dynamical behaviour of various natural phenomenon such as evolutionary process, swarm behaviour, genetic inheritance and others. Using these nature-inspired algorithms, we can solve more complex and challenging optimization and other useful tasks [23, 24, 25].

### 3.2.1 Gravitational search algorithm

Gravitational Search Algorithm (GSA) is a nature-inspired optimization algorithm based on Newton's law of gravitation. In GSA, features are taken as objects and hence, copying the nature of gravitational force comes into action where all the objects go towards the heaviest mass object due to force. Using this law, GSA explores whole solution space and find the optimal solution within the solution space. Hence, GSA proved to be very efficient and powerful optimization technique and can be used to solve complex optimization problems [26]. Newton's law of gravitation is given by,

$$F = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} G \frac{M_i M_j}{R_{i,j}^2} \qquad \text{and} \qquad a_i = \frac{F}{M_i} \tag{4}$$

where, $n$ is the total number of objects, $G$ is the gravitational constant, $F$ is total gravitational force, $R_{i,j}$ is distance between $i^{th}$ and $j^{th}$ objects, $M_i$ is mass of $i^{th}$ object, $a_i$ is the acceleration of $i^{th}$ object.

### 3.2.2 Binary bat search algorithm

Bat Algorithm (BA) is a nature-inspired optimization algorithm based on the concept of echolocation used by bats to navigate through in night and finding prey. Bats sends a ultrasonic ranged high frequency sound which reflects back after striking the prey and surrounding. The reflected sound is analyzed by bats and hence predict the surroundings in dark nights as well. Binary Bat Algorithm (BBA) is just a binary version of original BA using the same concept of echolocation [27]. Transfer function and position function used in BBA is given by,

$$S(v_i^k(t)) = \frac{1}{1 + e^{-v_i^k(t)}} \tag{5}$$

$$x_i^k(t+1) = \begin{cases} 0, & \text{if } rand < S(v_i^k(t+1)) \\ 1, & \text{if } rand \geq S(v_i^k(t+1)) \end{cases} \tag{6}$$

where, $x_i^k(t)$ and $v_i^k(t)$ represents the position vector and velocity vector of $i^{th}$ particle at $t^{th}$ iteration in $k^{th}$ dimension respectively.

### 3.2.3 Cuckoo search algorithm

Cuckoo Search Algorithm (CS) is a nature-inspired optimization algorithm based on breeding behaviour of cuckoo birds in the nests of other birds. CS incorporates the levy flights which means random walk of step size following heavy-tailed probability distribution which is helpful in escaping the local optima. Hence, CS is more suitable and efficient than other swarm algorithms [28]. The nonlinear system in CS can be represented as follows,

$$x_i^{t+1} = x_i^t + \alpha L(s, \lambda) \tag{7}$$

$$L(s, \lambda) \sim \frac{\lambda \Gamma(\lambda) \sin(\frac{\pi \lambda}{2})}{\pi} \frac{1}{s^{1+\lambda}}, \ (s > 0) \tag{8}$$

where, $\alpha > 0$ is the step size scaling factor which is user defined parameter.

### 3.2.4 Genetic algorithm

Genetic Algorithm (GA) is a nature-inspired evolutionary optimization algorithm based on the Darwin's theory of evolution, survival of the fittest principle. GA is an iterative algorithm, where best offspring are chosen at each iteration for the creation of next offspring. So, generally saying iteration means population generation. GA consists of 3 stages as follows,

1. ***Selection:*** The dominant parents out of whole population are selected from the given set of features. Probability and expectation of the survival concepts are used to determine the fittest features which will be used to find the offspring.

2. ***Crossover:*** Crossover is the stage which involves the production of the next generation offspring from the selected parents. Crossover process and methods depends on the requirements of the user.

3. ***Repeat the Process:*** The above stages of selection and crossover are repeated iteratively to improve the population over multiple generations. However, if the population doesn't improve over multiple generations then we proceed to mutation of the species.

4. ***Mutation:*** This stage rarely required because mutation is required when evolution does not happen even after huge iterations or, we can say after a considerable time.

GA is one of the basic nature-inspired optimization algorithm. Selection and crossover steps of GA can be seen as natural selection and genetic recombination for better offspring and increasing the sustainability of the population and hence, increasing the accuracy of the algorithm. Through selection of the fittest population and combining their features for future offspring, GA extensively search the whole solution space and hence, find the optimal solution required.

### 3.2.5 Grey wolf optimization algorithm

Grey Wolf Optimization (GWO) is a nature-inspired optimization algorithm based on the hunting behaviour of grey wolf species. Grey wolves lives and hunt in 4 hierarchical groups $\alpha$, $\beta$, $\delta$, $\omega$ where, $\alpha$ is the dominant group and $\omega$ is lowest level group [29]. GWO is however difficult to implement in comparison with other similar nature-inspired algorithms but the it produces somewhat better result due to variety of parameter existence in this algorithm. Mathematical model of GWO can be represented as follows,

$$D = |CX_p(t) - X(t)|\omega \tag{9}$$

$$X(t+1) = X_p(t) - AD \tag{10}$$

$$A = 2ar_1 - a \tag{11}$$

$$C = 2r_2 \tag{12}$$

$$a = 2 - \frac{2t}{T_{max}} \tag{13}$$

$$\begin{cases} D_\alpha = |CX_\alpha - X| \\ D_\beta = |CX_\beta - X| \\ D_\delta = |CX_\delta - X| \end{cases} \tag{14}$$

$$\begin{cases} X_1 = |X_\alpha - A_1 D_\alpha| \\ X_2 = |X_\beta - A_2 D_\beta| \\ X_3 = |X_\delta - A_3 D_\delta| \end{cases} \tag{15}$$

$$X(t+1) = \frac{X_1 + X_2 + X_3}{3} \tag{16}$$

where, $X_p(t)$ and $X(t)$ are prey location vector and grey wolf location vector at $t^{th}$ iteration, $r_1, r_2$ are random vectors, $T_{max}$ is maximum number of iterations, $D_i$ is distance between wolf unit $w$ and $i^{th}$ wolf, and $(X_1, X_2, X_3)$ represents location vector of new generation wolves.

### 3.2.6 Particle swarm optimization algorithm

Particle Swarm Optimization (PSO) is a nature-inspired metaheuristic algorithm based on the swarming behaviour of fishes and birds. In PSO, each bird or fish (generally called particle) moves through solution space with adjustments of their velocity thrust vectors according to the global nature of the swarm and personal experience. Global nature creates a thrust towards global optima, personal experience creates a thrust towards personal optima which particle has already reached in history and current velocity direction. Vector addition of all of these velocity thrust components will result in final movement of particle at that particular iteration. By iterative updation of velocity vector and position vector, the whole swarm finally converges to the global optima which makes PSO one of the most useful optimization algorithm widely used in many fields. In PSO, the velocity vector and position vector updation is done as follows,

$$v_i^{t+1} = v_i^t + \alpha\epsilon_1(g^* - x_i^t) + \beta\epsilon_2(x_i^* - x_i^t) \tag{17}$$

$$x_i^{t+1} = x_i^t + v_i^{t+1} \tag{18}$$

where, $x_i$ is position vector, $v_i$ is velocity vector of $i^{th}$ swarm respectively. $n$ is the population size and $\epsilon_1, \epsilon_2$ are uniformly distributed random vectors in $[0, 1]$. $\alpha$ and $\beta$ are user defined thrusts parameters in the directions of global optimum and personal optimum respectively.

### 3.2.7 Whale optimization algorithm

The Whale Optimization Algorithm (WOA) is a nature-inspired optimization algorithm based on the predatory behaviour of the whales. Whales uses bubble-net forage to concentrate prey in a smaller area and finally capturing them. Similarly, WOA concentrates the whole solution space in smaller and smaller area as the iteration continues. Hence, WOA provides a new and unique way for solving optimization problems [30]. Mathematical model of WOA can be represented as follows,

$$D = |CX^*(t) - X(t)| \tag{19}$$

$$X(t+1) = X^*(t) - AD \tag{20}$$

$$A = 2a(\text{rand}) - a \tag{21}$$

$$C = 2(\text{rand}) \tag{22}$$

$$a = 2 - \frac{2t}{T_{max}} \tag{23}$$

$$X(t+1) = D'e^{bl}\cos(2\pi l) + X^*(t) \tag{24}$$

$$D' = |X^*(t) - X(t)| \tag{25}$$

where, $t$ is the current number of iterations, $X^*(t)$ is the prey position vector which represents current optimal solution, $X(t)$ is the prey position vector, $AD$ is the surrounding step size, $T_{max}$ is maximum number of iterations, $D'$ is the distance between the $i^{th}$ whale and the current optimal position $X^*(t)$, $b$ is a constant coefficient used to define the logarithmic spiral form, $l$ takes the random number of $[-1, 1]$ and $|A| \geq 1$

## 3.3 Supervised ML algorithms

Supervised ML algorithms are those algorithms which work on labelled input and output data. The supervised ML algorithms find a mapping function or, transformation function which maps the input variable $x$ with the corresponding output variable $y$. Supervised ML algorithms are data-driven that is, they rely on the data provided to them and maps using that data only. So we have to consider that the data provided is accurate. The mapping function then helps us to predict the output value for the unknown input, which will be based on the data provided. Supervised learning learns mathematically using mapping function and hence, make more accurate predictions than unsupervised ML algorithms [31, 32].

### 3.3.1 Support vector machine

Support Vector Machines (SVM) is a supervised ML algorithm which can be used for classification as well as regression problems. SVM can be used for classification as it creates a hyperplane which divides different classes. The dimension of hyperplane depends on the number of features present in the dataset. SVM is known for its adaptability as it prevents the problem of overfitting. Hence, SVM is one of the best model that can be used for classification problems and can be applied in various fields including optimization.

### 3.3.2 K-nearest neighbour

K-Nearest Neighbors (KNN) is also a supervised ML algorithm and is one of the simplest algorithm that can be implemented in the domain of classification. Euclidean distance is used to determine the K nearest neighbour to a particular individual which optimizes the predictions through distance minimization. K nearest neighbours are used to make predictions for classification using majority voting algorithm. Hence, KNN proved to be a good algorithm for classification and already applied to various fields including optimization.

### 3.3.3 Random forest

Random Forest (RF) is also a supervised ML algorithm which uses ensemble learning approach. RF consists of large number of decision trees either either sequenced parallelly or simultaneously. Large number of trees create forest like structure hence, RF gets its name from here. RF can be easily implemented as it is just collection of decision trees. As there are large number of decision trees, there are large number of trainable parameters as well which can be optimized for better predictions. Due to existence of large number of trainable parameter, it can handle complex optimization problems and hence, RF can be used in various field including optimization.

These preliminary concepts will be used in the subsequent sections of the research paper and will be used in the proposed method. These concepts are combined to give a new method which will reduce the total computational time significantly along with an increase in accuracy.

# 4 Methodology

For achieving, high accuracy classification of ASD, we followed a methodology consisting of various steps including data pre-processing which includes correlation analysis, feature extraction using nature-inspired algorithms and ML algorithm selection for proper classification. This methodology adopted by us for ASD predictions is briefly explained in this section.

Fig. 2, shows the general flowchart of the proposed methodology for predicting ASD using nature-inspired algorithms and ML algorithms.

## 4.1 Data pre-processing

Data pre-processing step in proposed methodology is the first and most important step which is required to ensure the quality of the dataset. Data pre-processing consists of several minor steps including, data cleaning and null value removal. Data cleaning involved the removal of unnecessary data in the dataset, for example, variance and standard deviation represent one and same values. After data cleaning, dimension of dataset is reduced by some extent. There is no null values present in our dataset, hence removal of null value step is not required. Finally, after data pre-processing step our dataset left with 860 feature columns out of 1259 feature columns.

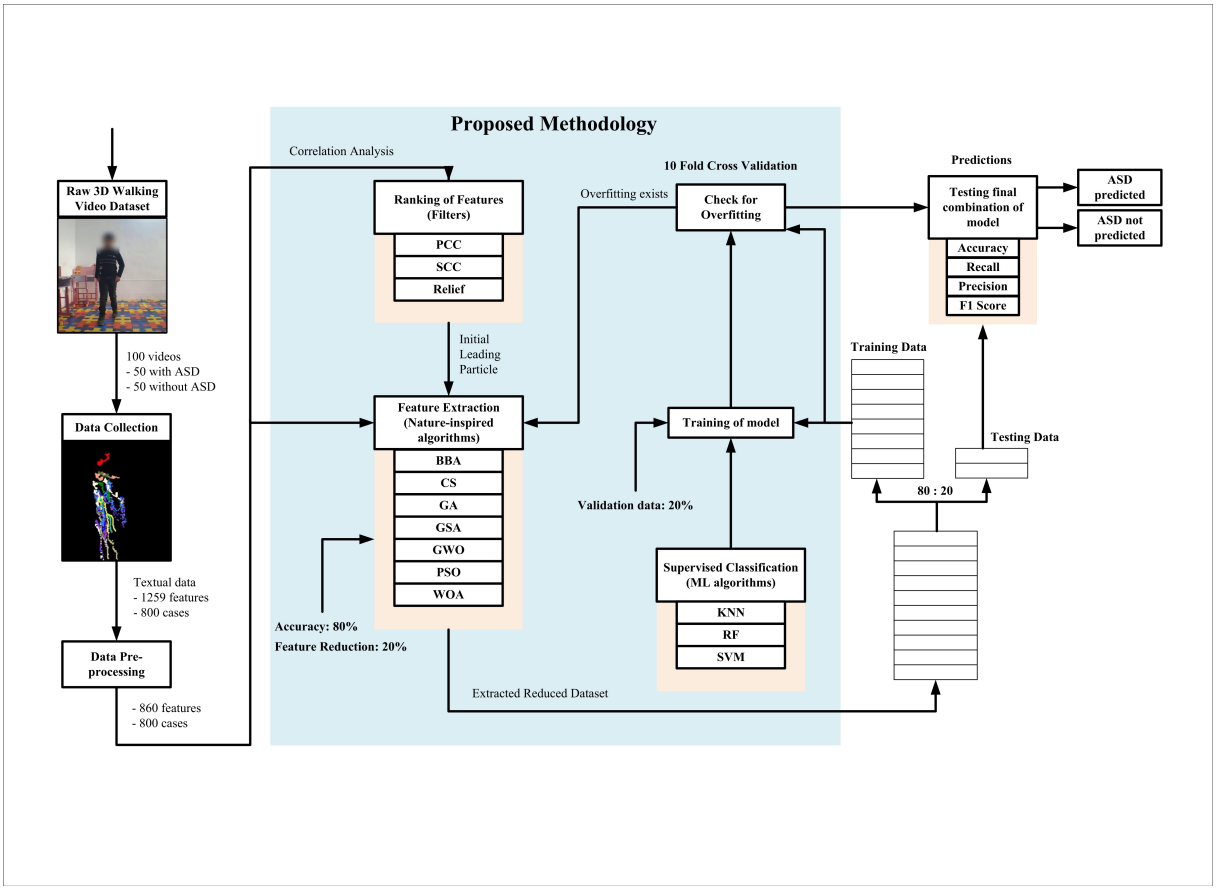



Fig. 2: Methodology adopted

## 4.2 Ranking of features

Ranking of features is second step in our proposed methodology which is required to find out correlation between different features. For this step, we used PCC, SCC and Relief ranking coefficient for correlation analysis. These coefficients helps us to find which feature is most important in the dataset and hence, we can find initial leading particle which will be used in nature-inspired feature extraction step. Finally, after ranking of feature step we found initial leading particle using correlation coefficients.

## 4.3 Feature extraction

Feature extraction is next step in our proposed methodology. From previous step we got initial leading particle for the chosen nature-inspired algorithm and hence, reducing the computational time. To achieve optimized feature selection, 80% weightage is provided to accuracy and 20% weightage is provided to feature selection. This weightage maintains the balance between accuracy and feature reduction percentages with emphasis on accuracy. Finally out of total data, 20% data is taken for validation purpose to check for proper accuracy. For comparison study various nature-inspired algorithms are applied for feature selection which includes, BBA, CS, GA, GSA, GWO, PSO and WOA. Finally, after feature selection dimension of the dataset is hugely reduced and total computational time is also reduced by a great extent.

## 4.4 Classification

Classification is next step in our proposed methodology. In this step we got the reduced dataset from previous step and we evaluate the accuracy and make predictions using different ML

classification algorithms. The classification algorithms includes KNN, RF and SVM. These ML algorithms are used to classify whether the individual have ASD or not. Finally, after classification step we have some ML algorithms which we are going to use in further steps.

### 4.5 Training and testing of ML models

Training and testing of ML models is next step in our proposed methodology. We divided the dataset obtained from feature extraction set into $80-20$ split that is, 80% is used for training purpose and 20% is used for testing purpose. After train-test split we use the ML classification algorithms obtained from previous step for classification purpose. To check for overfitting, we used 10-fold cross validation. 10-fold cross validation includes partitioning the dataset into 10 equal subsets. Training will be done on 9 subsets and the remaining one will be used for testing purpose. This process repeated for 10 times with various combinations of the subsets. Finally, 10-fold cross validation on the dataset using classification algorithm shows that our method do not suffer from overfitting.

### 4.6 Evaluation of models

In this research paper, we analysed various combinations of correlation ranking coefficients, nature-inspired optimization algorithms for feature extraction and ML classification algorithms. We evaluate the various combination models using confusion matrices. We computed accuracy, recall, precision and F1 score. From our finding, we come to know that the combination relief ranking coefficient, GSA and RF achieved highest accuracy of 100% with 380 features. On other hand, the combination relief ranking coefficient, CS and RF achieved accuracy of 96.875% with only 4 features. Hence, we finally got 2 combinations, one with highest accuracy and other with highest dimensional reduction.

### 4.7 Algorithm analysis

The algorithm 1 can handle textual data, making it suitable for processing video, image and textual data (section 5.1). It is suitable because of conversion of video and image data into textual data through pixel analysis. In our research work, we used initial video whose still is given in Fig. 3a, followed by selection of some specific points on the body of individual for analysis which are given in Fig. 3b. Finally, the paths traced by these points given in Fig. 3b throughout the video are used for analysis and creation of dataset. One such path is given in Fig. 3c. Then through feature extraction from obtained textual data, irrelevant features are removed. Finally, we received the dataset which is used for proposed methodology and finding best algorithms combination.

For example, based on Table 7, the combination of relief correlation coefficient, CS and RF yields the optimal results with only 4 features and an accuracy of 96.875%. These features are as follows,

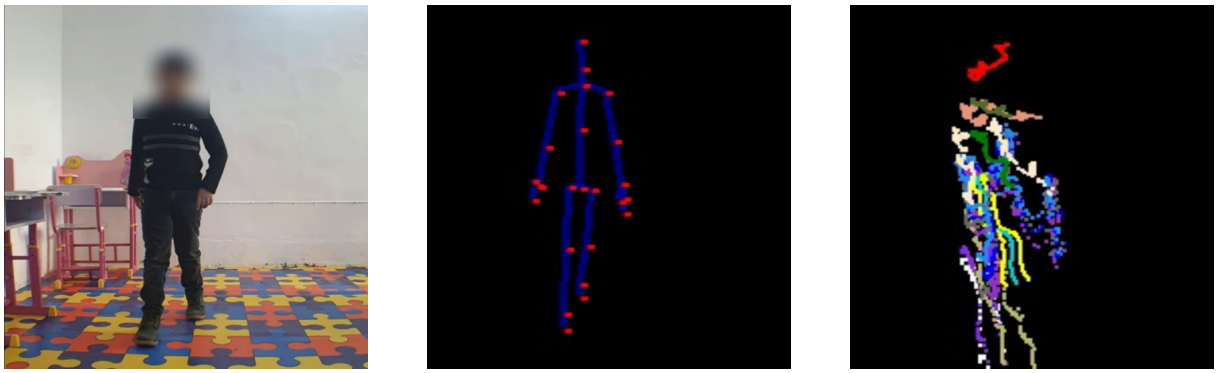
1. **mean-x-FootRight:** The mean position of the right foot along the x-axis (horizontal).
2. **std FoRTThR:** The correlation between the right foot and right thumb.
3. **std HaRTKeR:** The correlation between the right hand and right knee.
4. **std KeLTSb:** The standard deviation of the left knee.

The observation reveals that 3 out of the 4 features exhibits deviations from the mean path, which aligns with fact that ASD is a neurodevelopmental disorder. To achieve an accuracy of approximately 96%, the combination of relief correlation coefficient, CS and RF obtained from the proposed methodology can be used. Hence, these 3 extracted features can be used for predicting ASD in an unknown individual.

**Algorithm 1:** Algorithm for the Proposed Model

**Data:** ASD dataset with 1259 features and 800 test cases
**Result:** Model with highest accuracy with minimum number of features

S1 : CorrCoeff $\leftarrow \{PCC, SCC, Relief\}$;
S2 : NatureAlgo $\leftarrow \{BBA, CS, GA, GSA, GWO, PSO, WOA\}$;
S3 : MLAlgo $\leftarrow \{KNN, RF, SVM\}$;
accuracy $\leftarrow \langle int \rangle$ ;
Models $\leftarrow set\{\{\{s1, s2, s3\}, Result, accuracy\}\}$;
**while** *any combination (S1, S2, S3) exists* **do**
    % Combination selected {s1, s2, s3} (say);
    RankedFeatures $\leftarrow s1(features)$;
    Lead $\leftarrow Sort(RankedFeatures, Ranks)(1)$;
    Result $\leftarrow 1$;
    **while** *Result* **do**
        % Feature Extraction
        InitialLeading $\leftarrow$ *Lead*;
        totalfeatures $\leftarrow$ 1259;
        weight $\leftarrow$ 0.8;
        MaxItertions $\leftarrow$ 100;
        NumAgents $\leftarrow$ 30;
        fitness $\leftarrow$ weight$\times$Accuracy + (1 $-$ weight) $\times \frac{\text{Total Feature}-\text{Selected Features}}{\text{Total Features}}$;
        FEModel $\leftarrow \{dataset, InitialLeading, MaxItertions, NumAgents, fitness\}$;
        ExtractedFeatures $\leftarrow s2(FEModel)$;
        newDataset $\leftarrow Dataset(ExtractedFeatures)$;
        % Classification using ML Model
        train $\leftarrow$ 0.8, test $\leftarrow$ 0.2, validatevalue $\leftarrow$ 0.5;
        validate $\leftarrow$ test $\times$ validatevalue;
        target $\leftarrow$ target column from newDataset;
        data $\leftarrow$ drop target from newDataset;
        traindata, validatedata, testdata $\leftarrow$ split(data, target, train, validate, 1$-$ validate);
        result $\leftarrow$ evaluate(traindata, testdata, validatedata, s3);
        accuracy $\leftarrow$ result.accuracy;
        **if** *satisfiable accuracy* **then**
            % K fold cross validation for check overfitting
            $k \leftarrow 10$;
            score $\leftarrow$ Mean score for 10 fold cross validation
            **if** *satisfiable score* **then**
                Models $\leftarrow Models \cup$ {$\{s1, s2, s3\}$, result, accuracy};
                Result $\leftarrow$ 0;
            **else**
                continue;
        **else**
            continue;
%Selecting Model with highest accuracy
RequiredModel $\leftarrow sort(Models, accuracy)(1)$;



(a) Still from video (b) Points taken for study (c) Path followed by points

Fig. 3: 3D Videos of children walking towards camera

# 5 Result analysis and visualization

The result analysis section of this research paper cover various analysis of combinations. Initially, a study of dataset is done which provides indepth knowledge of data which is used in our methodology. After this, performance of various combinations of algorithms is done using confusion matrices. Finally, we provide full explaination of the outcomes we got in previous steps. Adding up this, these evaluations gives us final performance and significance of our methodology required for future reserach and applications.

## 5.1 Dataset description

In order to utilize video dataset for classification, we must have to transform the dataset into textual data. This is done through analysis of joint motions along all the 3 axes and determining standard deviation as well as variance. Angular deviation of some of the joints and other physical parameters are also extracted as textual data. These extracted features comes out to be very useful in determining ASD in an individual as it determine walking movement patterns and ASD affects it very severely.

The dataset used in our research is '*Three-Dimensional Dataset Combining Gait and Full Body Movement of Children with Autism Spectrum Disorders Collected by Kinect v2 Camera Dataset*' [1], which is a publicly available dataset.

Total of 100 Children are studied in video dataset, 50 of which are with typical development and the remaining 50 with ASD detected. Total of 1259 features are initially present in the extracted textual dataset.

Table 2 shows concise data in the given dataset which includes number of children tested for creation of the dataset, number of data cases present in the given dataset and the total number of missing values (nil/NA values).

Table 3 shows all of the features extracted from the 3D walking video dataset present with their physical units and total number of features belongs to that particular set of attributes. Physical units can be used for data cleaning process in the data pre-processing step in our proposed methodology.

Table 2: General Dataset Description

| **Data Cases/ Children** | **Typical Development** | **Autism Spectrum Disorder** | **Total** |
|---|---|---|---|
| Number of Children | 50 | 50 | 100 |
| Number of Data Cases | 400 | 400 | 800 |
| Missing Values | Nil | Nil | Nil |

Table 3: Features included in dataset (.xlsx extracted from .avi) [1]

| Abbreviation | Description | Type | Unit | Features |
|---|---|---|---|---|
| mean-coordination-Joint | Mean of x, y, and z coordination of 25 body joints | Floating-point | m | 75 |
| variance-coordination-Joint | Variance of x, y, and z coordination of 25 body joints | Floating-point | m2 | 75 |
| std-coordination-Joint | Standard deviation of x, y, and z coordination of 25 body joints | Floating-point | m | 75 |
| mean Angle | Mean of 18 angles between body joints | Floating-point | deg | 18 |
| variance Angle | Variance of 18 angles between body joints | Floating-point | deg2 | 18 |
| std Angle | Standard deviation of 18 angles between body joints | Floating-point | deg | 18 |
| mean joint1 T Joint2 | Mean of distance between body joints | Floating-point | m | 24 |
| variance joint1 T Joint2 | Variance of distance between body joints | Floating-point | m2 | 24 |
| std joint1 T Joint2 | Standard deviation of distance between body joints | Floating-point | m | 24 |
| Mean Joint TGr | Mean of distance between Joints and ground | Floating-point | m | 282 |
| variance Joint TGr | Variance of distance between Joints and ground | Floating-point | m2 | 282 |
| std Joint1 TGr | Standard deviation of distance between Joints and ground | Floating-point | m | 282 |
| Rom Joint coordination | Range of movement of 25 joint on x and y coordination | Floating-point | m | 50 |
| MaxStLe | Maximum stride length | Floating-point | m | 1 |
| MaxStWi | Maximum stride width | Floating-point | m | 1 |
| StrLe | Stride length | Floating-point | m | 1 |
| GaCT | Gait cycle time | Time | ms | 1 |
| StaT | Stance time | Time | ms | 1 |
| SwiT | Swing time | Time | ms | 1 |
| Velocity | Velocity | Floating-point | m/s | 1 |
| HaTiLPos | Hand Tip left position relative to SpanBase | Binary (0, 1) | - | 1 |
| HaTiRPos | Hand Tip Right position relative to SpanBase | Binary (0, 1) | - | 1 |
| MaxDBFE and MinDBFE | Maximum and minimum distance between feet | Floating-point | m | 2 |
| Threshold | Using in extract one gait cycle = average of distance between feet | Floating-point | m | 1 |
| Total Number of Features | - | - | - | 1259 |

## 5.2 Performance/ Evaluation metrics

In this research work, we evaluated the performance of the model (combination of algorithms) is done using confusion matrices. Confusion matrix is applied to describe performance, which includes various variables namely, True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN) [13]. These matrices gives indepth knowledge of the model. Hence, the evaluation of the classification model is done using these matrices.

### 5.2.1 Classification accuracy

Accuracy is the ratio of correct predictions (TP and TN) with the total number of predictions. Accuracy value lies in $[0, 1]$ and used to check the prediction ability of model.

$$\text{Accuracy } = \frac{TP + TN}{FP + FN + TP + TN} \tag{26}$$

### 5.2.2 Recall

Recall is the ratio of the properly identified positive values (TP) with total number of correct predictions. Recall is also known as sensitivity and it's value also lies in $[0, 1]$.

$$\text{Recall } = \frac{TP}{TP + FN} \tag{27}$$

### 5.2.3 Precision

Precision is the ratio of properly identified positive values (TP) with total positive predictions. Precision is the accuracy of correct detection of positive values and it's value also lies in $[0, 1]$.

$$\text{Precision } = \frac{TP}{TP + FP} \tag{28}$$

### 5.2.4 F1 score

F1 score is the harmonic mean of accuracy and recall. F1 score is one of the ideal value for analysis purpose. It's value also lies in $[0, 1]$.

$$\text{F1 Score } = 2 \times \frac{\text{precision } \times \text{ recall}}{\text{precision } + \text{ recall}} \tag{29}$$

## 5.3 Results

After evaluation of various combinations of correlation coefficients, nature-inspired optimization algorithms for feature selection and supervised ML classification algorithms, we came to know that combination of GSA and RF achieves an accuracy of 100%. This combination consists of 380 features for relief ranking coefficient, 407 features for PCC and 410 features for SCC. Out of all combinations, this combination (Relief, GSA, RF) turned out to have accuracy of 100% along with feature reduction of 69.817%. Alternatively, if we are concerned for feature reduction and not the accuracy upto that extent, then the combination (Relief, CS, RF) achieves an accuracy of 96.875% using only 4 features out of total 1259 features which corresponds to 99.68% feature reduction. One can choose from these 2 options according to the requirements.

For the problem of overfitting, we use 10-fold cross validation to prove that the model is not overfitted. Nature-inspired optimization algorithm used for feature extraction proved to be highly valuable as it significantly reduces the total computational time while maintaining the accuracy classification.

Table 4 shows the different performance matrix values with the PCC. This table contains a combination of various nature-inspired algorithms and ML algorithms. Similarly, Table 5 and Table 6 show the different performance matrices values with the SCC and relief ranking coefficient respectively. These tables also contain a combination of various nature-inspired algorithms and ML algorithms.

The accuracy analysis is visually represented using histogram which is given by Fig. 4, where different colors represent various nature-inspired algorithms. The histogram is divided into two sections, showing ML algorithms and ranking coefficients used in the evaluation.

Table 4: Result analysis for various combinations of feature selection and classification algorithms with Pearson Correlation Coefficient (PCC)

| Ranking | Nature Inspired Algorithms | ML Algorithms | Accuracy | Recall | | Precision | | F1 Score | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | Autism | Typical | Autism | Typical | Autism | Typical |
| PCC | GSA | **KNN** | 0.99375 | 0.9875 | **1.0** | **1.0** | 0.98765 | 0.99371 | 0.99379 |
| | | **RF** | **1.0** | **1.0** | **1.0** | **1.0** | **1.0** | **1.0** | **1.0** |
| | | **SVM** | 0.75625 | 0.5625 | 0.95 | 0.91837 | 0.6847 | 0.69678 | 0.7958 |
| | BBA | **KNN** | 0.95 | 0.9375 | 0.9625 | 0.96154 | 0.939 | 0.94937 | 0.9506 |
| | | **RF** | 0.99375 | **1.0** | 0.9875 | 0.98765 | **1.0** | 0.99379 | 0.99371 |
| | | **SVM** | 0.7625 | 0.65 | 0.875 | 0.8387 | 0.7143 | 0.73239 | 0.78652 |
| | CS | **KNN** | 0.95 | 0.925 | 0.975 | 0.98684 | 0.94048 | 0.96154 | 0.96341 |
| | | **RF** | 0.94375 | 0.9125 | 0.975 | 0.9737 | 0.92857 | 0.94872 | 0.95122 |
| | | **SVM** | 0.6875 | 0.375 | **1.0** | **1.0** | 0.61538 | 0.54545 | 0.7619 |
| | GA | **KNN** | 0.98125 | 0.9625 | **1.0** | **1.0** | 0.96386 | 0.98089 | 0.9816 |
| | | **RF** | 0.99375 | **1.0** | 0.9875 | 0.9876 | **1.0** | 0.9938 | 0.99371 |
| | | **SVM** | 0.7625 | 0.5625 | 0.9625 | 0.9375 | 0.6875 | 0.7031 | 0.8021 |
| | GWO | **KNN** | 0.96875 | 0.9625 | 0.975 | 0.9747 | 0.963 | 0.9686 | 0.9689 |
| | | **RF** | 0.99375 | 0.9875 | **1.0** | **1.0** | 0.98765 | 0.99371 | 0.9938 |
| | | **SVM** | 0.75 | 0.625 | 0.875 | 0.8334 | 0.7 | 0.7143 | 0.7778 |
| | PSO | **KNN** | 0.975 | 0.9625 | 0.9875 | 0.98718 | 0.96341 | 0.9747 | 0.9753 |
| | | **RF** | 0.99375 | **1.0** | 0.9875 | 0.98765 | **1.0** | 0.9938 | 0.99371 |
| | | **SVM** | 0.71875 | 0.4625 | 0.975 | 0.94872 | 0.64463 | 0.62185 | 0.77612 |
| | WOA | **KNN** | 0.9875 | 0.975 | **1.0** | **1.0** | 0.9756 | 0.98734 | 0.98765 |
| | | **RF** | 0.9875 | 0.9875 | 0.9875 | 0.9875 | 0.9875 | 0.9875 | 0.9875 |
| | | **SVM** | 0.7625 | 0.5625 | 0.9625 | 0.9375 | 0.6875 | 0.70312 | 0.80208 |

The graphs shown in Fig. 5 represents the relationship between the number of features and their frequency of selection when PCC is used. It is clear from the graph that using CS algorithm

results in minimum number of feature (only 4) with an accuracy of 95%. While GSA results in highest accuracy of 100% with 407 features to be taken into consideration.

The graphs depicted in Fig. 6 illustrates the relationship between the number of features and their frequency of selection when the SCC is used. It is clear from the graph that using CS algorithm results in minimum number of feature (only 18) with an accuracy of 95.625%. While GSA results in highest accuracy of 100% with 410 features to be taken into consideration.

The graphs depicted in Fig. 7 illustrates the relationship between the number of features and their frequency of selection when relief ranking coefficient is used. It is clear from the graph that using CS algorithm results in minimum number of feature (only 4) with an accuracy of 96.875%. While GSA results in highest accuracy of 100% with 380 features to be taken into consideration.

Table 5: Result analysis for various combinations of feature selection and classification algorithms with Spearman Correlation Coefficient (SCC)

| Ranking | Nature Inspired Algorithms | ML Algorithms | Accuracy | Recall | | Precision | | F1 Score | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | Autism | Typical | Autism | Typical | Autism | Typical |
| SCC | GSA | KNN | 0.98125 | 0.9625 | **1.0** | **1.0** | 0.96385 | 0.98089 | 0.9816 |
| | | RF | **1.0** | **1.0** | **1.0** | **1.0** | **1.0** | **1.0** | **1.0** |
| | | SVM | 0.75 | 0.5375 | 0.9625 | 0.9348 | 0.67544 | 0.68254 | 0.7938 |
| | BBA | KNN | 0.98125 | 0.9625 | **1.0** | **1.0** | 0.96385 | 0.98089 | 0.9816 |
| | | RF | 0.99375 | **1.0** | 0.9875 | 0.98765 | **1.0** | 0.9938 | 0.99371 |
| | | SVM | 0.725 | 0.4875 | 0.9625 | 0.92857 | 0.65254 | 0.63934 | 0.77779 |
| | CS | KNN | 0.9375 | 0.9375 | 0.9375 | 0.9375 | 0.9375 | 0.9375 | 0.9375 |
| | | RF | 0.95625 | 0.975 | 0.9375 | 0.93976 | 0.974 | 0.957 | 0.9554 |
| | | SVM | 0.85 | 0.725 | 0.975 | 0.96667 | 0.78 | 0.82857 | 0.86667 |
| | GA | KNN | 0.98125 | 0.9625 | **1.0** | **1.0** | 0.96386 | 0.98089 | 0.9816 |
| | | RF | 0.99375 | **1.0** | 0.9875 | 0.98765 | **1.0** | 0.99379 | 0.99371 |
| | | SVM | 0.7625 | 0.5625 | 0.9625 | 0.9375 | 0.6875 | 0.7031 | 0.8021 |
| | GWO | KNN | 0.96875 | 0.9625 | 0.975 | 0.9747 | 0.963 | 0.9685 | 0.96894 |
| | | RF | 0.99375 | 0.9875 | **1.0** | **1.0** | 0.98765 | 0.99371 | 0.99379 |
| | | SVM | 0.75 | 0.625 | 0.875 | 0.83334 | 0.7 | 0.71429 | 0.7778 |
| | PSO | KNN | 0.975 | 0.9625 | 0.9875 | 0.98718 | 0.96341 | 0.9747 | 0.9753 |
| | | RF | 0.99375 | **1.0** | 0.9875 | 0.98765 | **1.0** | 0.99379 | 0.99371 |
| | | SVM | 0.71875 | 0.4625 | 0.975 | 0.94872 | 0.64463 | 0.62185 | 0.77612 |
| | WOA | KNN | 0.9875 | 0.975 | **1.0** | **1.0** | 0.9756 | 0.98734 | 0.98765 |
| | | RF | 0.99375 | 0.9875 | **1.0** | **1.0** | 0.98765 | 0.99371 | 0.99379 |
| | | SVM | 0.7625 | 0.5625 | 0.9625 | 0.9375 | 0.6875 | 0.7031 | 0.80208 |

Table 6: Result analysis for various combinations of feature selection and classification algorithms with Relief Ranking Coefficient

| Ranking | Nature Inspired Algorithms | ML Algorithms | Accuracy | Recall | | Precision | | F1 Score | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | Autism | Typical | Autism | Typical | Autism | Typical |
| Relief | GSA | KNN | 0.975 | 0.9625 | 0.9875 | 0.98718 | 0.96341 | 0.97468 | 0.9753 |
| | | RF | **1.0** | **1.0** | **1.0** | **1.0** | **1.0** | **1.0** | **1.0** |
| | | SVM | 0.81875 | 0.75 | 0.8875 | 0.86956 | 0.78022 | 0.80537 | 0.8304 |
| | BBA | KNN | 0.96875 | 0.95 | 0.9875 | 0.98701 | 0.9518 | 0.96815 | 0.96932 |
| | | RF | 0.9875 | 0.975 | **1.0** | **1.0** | 0.9756 | 0.98734 | 0.98765 |
| | | SVM | 0.71875 | 0.4375 | **1.0** | **1.0** | 0.64 | 0.6087 | 0.78049 |
| | CS | KNN | 0.9625 | 0.9375 | 0.9875 | 0.98684 | 0.94048 | 0.96154 | 0.96341 |
| | | RF | 0.96875 | 0.95 | 0.9875 | 0.98701 | 0.9518 | 0.96815 | 0.96932 |
| | | SVM | 0.8375 | 0.7375 | 0.9375 | 0.92187 | 0.78125 | 0.81945 | 0.85227 |
| | GA | KNN | 0.98125 | 0.9625 | **1.0** | **1.0** | 0.96386 | 0.98089 | 0.9816 |
| | | RF | **1.0** | **1.0** | **1.0** | **1.0** | **1.0** | **1.0** | **1.0** |
| | | SVM | 0.7625 | 0.5625 | 0.9625 | 0.9375 | 0.6875 | 0.70312 | 0.80208 |
| | GWO | KNN | 0.9125 | 0.95 | 0.875 | 0.88372 | 0.94595 | 0.91566 | 0.90909 |
| | | RF | 0.9875 | 0.975 | **1.0** | **1.0** | 0.9756 | 0.98734 | 0.98765 |
| | | SVM | 0.7375 | 0.4875 | 0.9875 | 0.975 | 0.65834 | 0.65 | 0.78999 |
| | PSO | KNN | 0.975 | 0.95 | **1.0** | **1.0** | 0.95238 | 0.97436 | 0.9756 |
| | | RF | 0.9875 | 0.9875 | 0.9875 | 0.9875 | 0.9875 | 0.9875 | 0.9875 |
| | | SVM | 0.74375 | 0.525 | 0.9625 | 0.93334 | 0.66956 | 0.672 | 0.78974 |
| | WOA | KNN | 0.9875 | 0.975 | **1.0** | **1.0** | 0.9756 | 0.98734 | 0.98765 |
| | | RF | 0.99375 | 0.9875 | **1.0** | **1.0** | 0.98765 | 0.99371 | 0.99379 |
| | | SVM | 0.7625 | 0.5625 | 0.9625 | 0.9375 | 0.6875 | 0.70312 | 0.80208 |

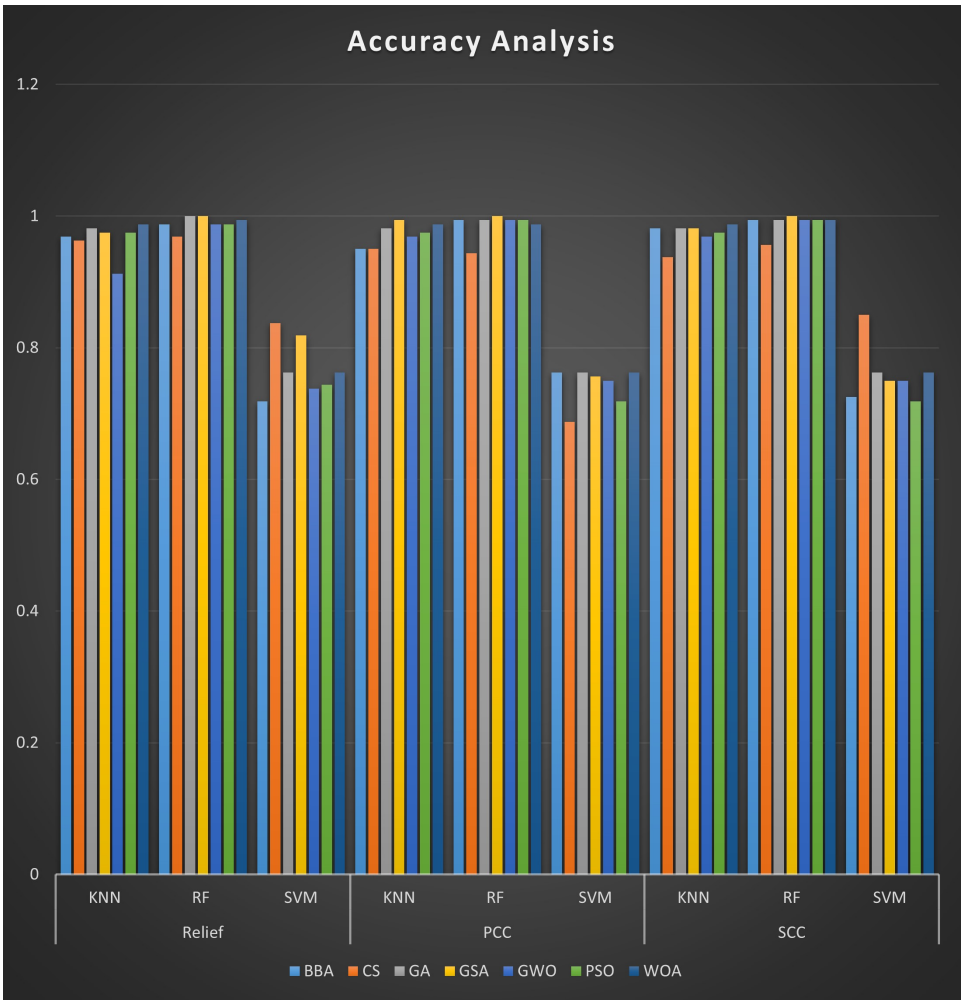
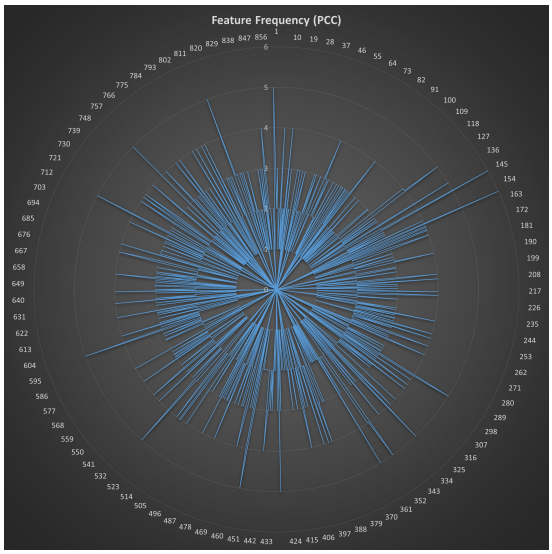
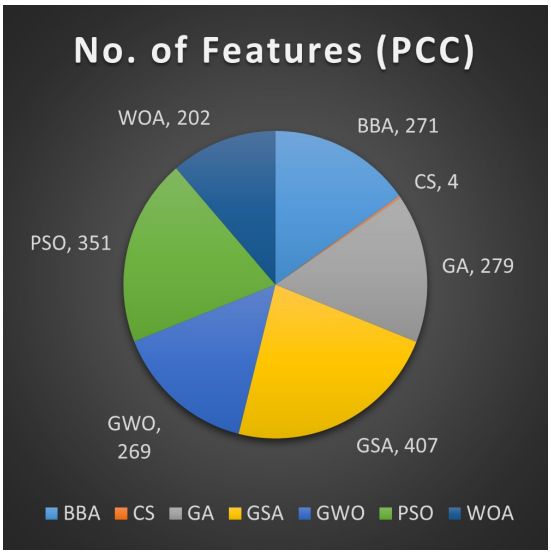



Fig. 4: Accuracy Analysis of various combinations of feature extraction and classification algorithms

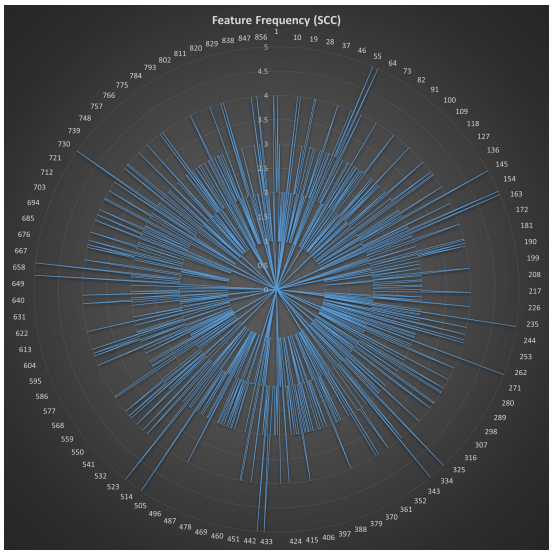



(a) Feature frequency graph

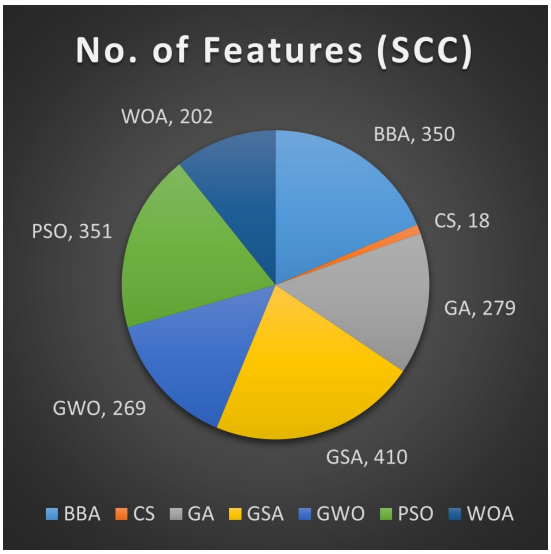



(b) Feature count

Fig. 5: Feature Analysis for Pearson Correlation Coefficient

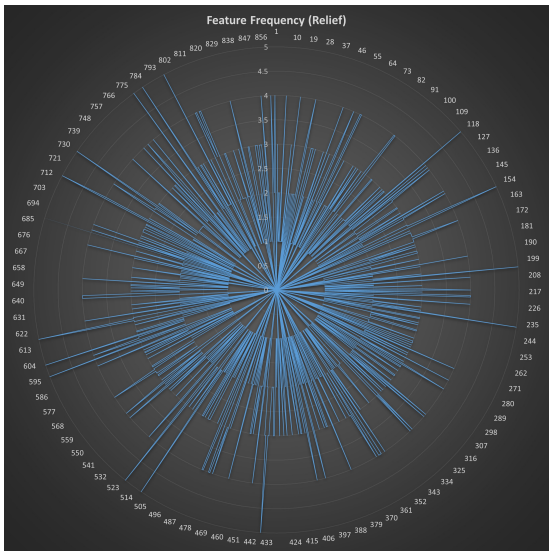



(a) Feature frequency graph

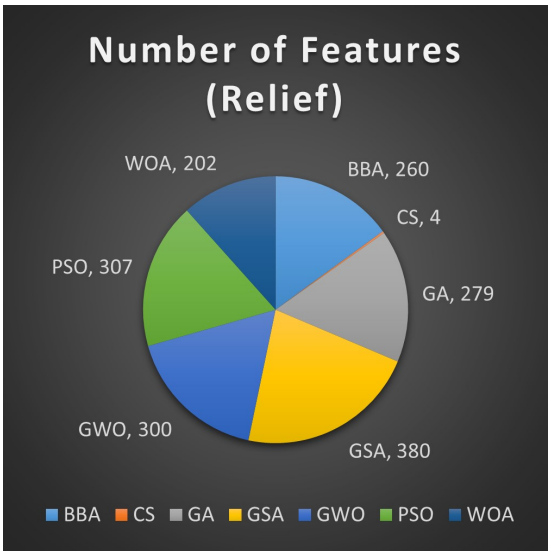



(b) Feature count

Fig. 6: Feature Analysis for Spearman Correlation Coefficient



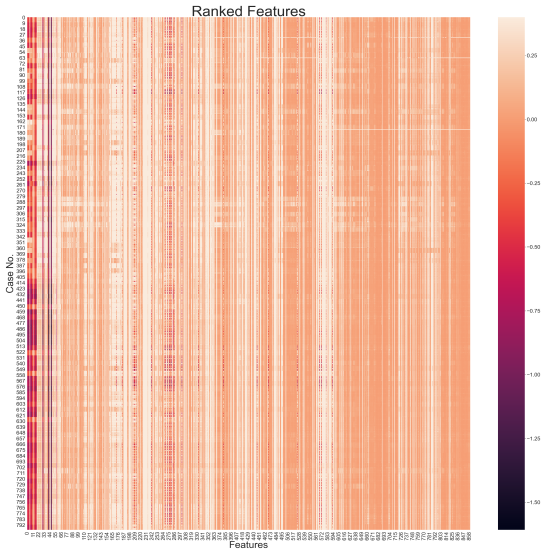

(a) Feature frequency graph



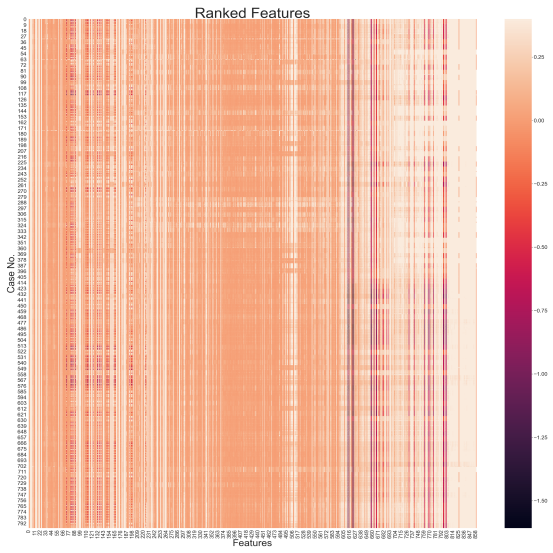

(b) Feature count

Fig. 7: Feature Analysis for Relief Ranking Coefficient

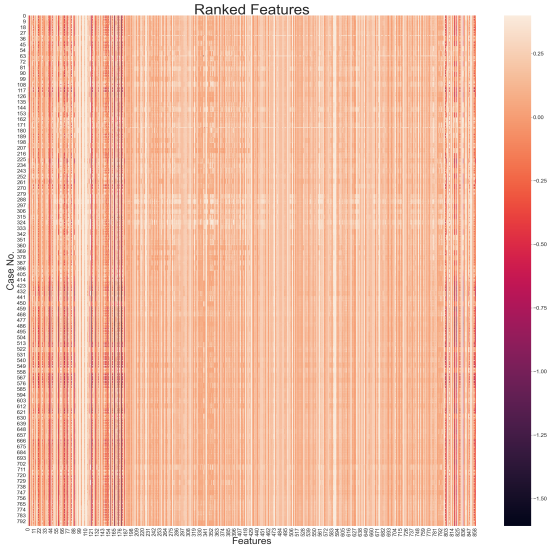
Fig. 8 represents the correlation heatmaps, showing the relationship between the features and cases using different correlation ranking coefficients. Heatmap have x-axis as a number of features (total of 1259) and y-axis as a case number (total of 800).




(a) Pearson Correlation Coefficient




(b) Spearman Correlation Coefficient




(c) Relief Ranking Coefficient

Fig. 8: Correlation Heatmaps

Table 7 shows the 2 best results obtained from the proposed methodology which can be used due to their distinction from the rest of the results. If space and time complexities are concerned then the combination relief ranking coefficient, CS for feature extraction and RF for classification can be used with accuracy of 96.875% along with feature reduction of 99.68%. If accuracy is concerned then the combination relief ranking coefficient, GSA for feature extraction and RF for classification can be used with accuracy of 100% along with feature reduction of 69.81%.

Table 7: Best Results using Relief Ranking coefficient and Random Forest algorithm

| **Feature Selection Algorithm** | **Accuracy** | **No. of Selected Features** | **Feature Reduction** |
|---|---|---|---|
| Gravitational Search Algorithm | **100%** | 380 | 69.81% |
| Cuckoo Search Algorithm | 96.875% | 4 | **99.68%** |

# 6 Conclusion and discussions

The rising cases of ASD around the world make it very crucial for it's early predictions and treatment. Hence, researchers are trying around the world to develop an algorithm that is easy to implement and cost effective which gives higher accuracy results with low total computational time. This paper aims to tackle this problem for 3D walking video dataset and by combining different concepts of correlation coefficients, feature extraction using nature-inspired optimization algorithms and supervise ML classification algorithms. For robustness, no previous work has been done on video dataset for predicting ASD that have such a promising result as ours. The research methodology involves the use of 3D walking video dataset containing data for children with and without ASD. The proposed methodology used in this paper shows significant reduction in total computational time compared to other existing methods. This reduction is due to several factors including selection of initial leading particle for nature-inspired algorithms using correlation analysis specifically, ranking coefficients and feature selection process using nature-inspired optimization algorithms. After evaluating all of the possible combinations, 2 combinations comes out from other combinations due to their distinct results from others. The combination relief ranking coefficient, CS algorithm for feature selection and RF algorithm for classification can be used whenever time and space complexity is concerned. And, the combination relief ranking coefficient, GSA algorithm for feature selection and RF algorithm for classification can be used whenever high accuracy is required. To enhance the proposed methodology, future development and more research is required using bigger and vast variety dataset with higher number of cases. Also, using deep learning and reinforcement models may further boost the generalizability and adaptability of the proposed methodology with practical implementation. As GSA is gives best results in terms of accuracy, so GSA can be further modified for various distance functions and different dimensions which results in better feature extraction optimization for the dataset. Moreover, the proposed methodology have the potential for advancement utilizing the concept of computer vision for the real time ASD predictions. With 100% accuracy, the proposed methodology can be deployed by healthcare professionals in real world for ASD detection. However, continuous research and development in the field of ASD detection is very necessary to fine tune the proposed model and hence, ensuring the accuracy and effectiveness of ASD prediction across diverse populations. Further advancement of the model, through significant contributions in research can improve the early detection and treatment of ASD ultimately, increasing the quality of life for individuals with neurodevelopment disorders.

# Declarations

## Competing Interests

The authors declare that they have no known competing interests.

## Data Availability

The dataset used, "*Three-Dimensional Dataset Combining Gait and Full Body Movement of Children with Autism Spectrum Disorders Collected by Kinect v2 Camera Dataset*," comprising 1259 features and 800 cases (400 typical development, 400 ASD diagnosed), is publicly accessible.

# References

[1] A. AbdulRahman, I. Hadi, and Y. Rajihy, "Generating 3D dataset of Gait and Full body movement of children with Autism spectrum disorders collected by Kinect v2 camera", ijact, vol. 9, no. 8, pp. 3791-3797, Sep. 2020.

[2] P. Kaur and M. Sharma, "Diagnosis of Human Psychological Disorders using Supervised Learning and Nature-Inspired Computing Techniques: A Meta-Analysis", Journal of Medical Systems, vol. 43, no. 7, May 2019, doi: https://doi.org/10.1007/s10916-019-1341-2.

[3] Mahmoud Elbattah, Jean-Luc Guérin, Romuald Carette, F. Cilia, and Gilles Dequen, "NLP-Based Approach to Detect Autism Spectrum Disorder in Saccadic Eye Movement", Dec. 2020, doi: https://doi.org/10.1109/ssci47803.2020.9308238.

[4] J. H. Lee, G. W. Lee, G. Bong, H. J. Yoo, and H. K. Kim, "Deep-Learning-Based Detection of Infants with Autism Spectrum Disorder Using Auto-Encoder Feature Representation", Sensors, vol. 20, no. 23, p. 6762, Nov. 2020, doi: https://doi.org/10.3390/s20236762.

[5] Z. Zheng et al., "Design of an Autonomous Social Orienting Training System (ASOTS) for Young Children With Autism", IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 25, no. 6, pp. 668–678, Jun. 2017, doi: https://doi.org/10.1109/tnsre.2016.2598727.

[6] A. Puli and A. Kushki, "Toward Automatic Anxiety Detection in Autism: A Real-Time Algorithm for Detecting Physiological Arousal in the Presence of Motion", IEEE Transactions on Biomedical Engineering, vol. 67, no. 3, pp. 646–657, Mar. 2020, doi: https://doi.org/10.1109/tbme.2019.2919273.

[7] J. Li, Y. Zhong, J. Han, G. Ouyang, X. Li, and H. Liu, "Classifying ASD children with LSTM based on raw videos", Neurocomputing, vol. 390, pp. 226–238, May 2020, doi: https://doi.org/10.1016/j.neucom.2019.05.106.

[8] A. Saranya and Rangasamy Anandan, "Autism Spectrum Prognosis using Worm Optimized Extreme Learning Machine (WOEM) Technique", Mar. 2021, doi: https://doi.org/10.1109/icacite51222.2021.9404729.

[9] MALATHI, S., and D. KANNAN. "ADAPTIVE WHALE OPTIMIZATION BASED SUPPORT VECTOR MACHINE FOR PREDICTION OF AUTISM SPECTRUM DISORDER.", Journal of Theoretical and Applied Information Technology 100, no. 9 (2022).

[10] C. Jayakumaran and J. Dhalia Sweetlin, "An Optimized Neural Network with Inertia Weight Variation of PSO for the detection of Autism", Oct. 2020, doi: https://doi.org/10.1109/icstcee49637.2020.9277247.

[11] S. K. Sahu and P. Verma, "Classification of Autistic Spectrum Disorder Using Deep Neural Network With Particle Swarm Optimization", International Journal of Computer Vision and Image Processing, vol. 12, no. 1, pp. 1–11, Jan. 2022, doi: https://doi.org/10.4018/ijcvip.290398.

[12] S. Malathi and D. Kannan, "Robust Artificial Bee Colony Optimization Based Classifier for Prediction of Autism Spectrum Disorder", Mar. 2022, doi: https://doi.org/10.1109/icacta54488.2022.9753510.

[13] B. Elshoky, O. Ibrahim, and A. Ali, "MACHINE LEARNING TECHNIQUES BASED ON FEATURE SELECTION FOR IMPROVING AUTISM DISEASE CLASSIFICATION", International Journal of Intelligent Computing and Information Sciences, vol. 21, no. 2, pp. 65– 81, Jul. 2021, doi: https://doi.org/10.21608/ijicis.2021.61582.1058.

[14] Balakrishnan, K., R. Dhanalakshmi, and Utkarsh Mahadeo Khaire. "Detecting Autism spectrum disorder with sailfish optimisation." (2021).

[15] N. Goel, B. Grover, Anuj, D. Gupta, A. Khanna, and M. Sharma, "Modified Grasshopper Optimization Algorithm for detection of Autism Spectrum Disorder", Physical Communication, vol. 41, p. 101115, Aug. 2020, doi: https://doi.org/10.1016/j.phycom.2020.101115.

[16] M. Farrell, Kurniawan Nur Ramadhani, and Suyanto Suyanto, "Combined Firefly Algorithm Random Forest to Classify Autistic Spectrum Disorders", Dec. 2020, doi: https://doi.org/10.1109/isriti51436.2020.9315396.

[17] A. Z. Guo, "Automated Autism Detection based on Characterizing Observable Patterns from Photos", IEEE Transactions on Affective Computing, pp. 1–1, 2020, doi: https://doi.org/10.1109/taffc.2020.3035088.

[18] R. Abitha and S. Mary Vennila, "A Swarm Based Symmetrical Uncertainty Feature Selection Method for Autism Spectrum Disorders", Jan. 2019, doi: https://doi.org/10.1109/icisc44355.2019.9036454.

[19] M. D. Samad, N. Diawara, J. Bobzien, J. A. Harrington, M. A. Witherow, and K. M. Iftekharuddin, "A Feasibility Study of Autism behavioural Markers in Spontaneous Facial, Visual, and Hand Movement Response Data", IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 26, no. 2, pp. 353–361, Feb. 2018, doi: https://doi.org/10.1109/tnsre.2017.2768482.

[20] P. Schober, C. Boer, and L. A. Schwarte, "Correlation coefficients: Appropriate use and interpretation", Anesthesia & Analgesia, vol. 126, no. 5, pp. 1763–1768, 2018, doi: https://doi.org/10.1213/ane.0000000000002864.

[21] Mukaka, Mavuto M. "A guide to appropriate use of correlation coefficient in medical research." Malawi medical journal 24, no. 3 (2012): 69-71.

[22] Porkodi, R. "Comparison of filter based feature selection algorithms: an overview." International journal of Innovative Research in Technology & Science 2, no. 2 (2014): 108-113.

[23] A. Darwish, "Bio-inspired computing: Algorithms review, deep analysis, and the scope of applications", Future Computing and Informatics Journal, vol. 3, no. 2, pp. 231–246, Dec. 2018, doi: https://doi.org/10.1016/j.fcij.2018.06.001.

[24] Binitha, S., and S. Siva Sathya. "A survey of bio inspired optimization algorithms." International journal of soft computing and engineering 2, no. 2 (2012): 137-151.

[25] X.-S. Yang and X. He, "Nature-Inspired Optimization Algorithms in Engineering: Overview and Applications", Studies in computational intelligence, pp. 1–20, Mar. 2016, doi: https://doi.org/10.1007/978-3-319-30235-5_1.

[26] H. Mittal, A. Tripathi, A. C. Pandey, and R. Pal, "Gravitational search algorithm: a comprehensive analysis of recent variants", Multimedia Tools and Applications, vol. 80, no. 5, pp. 7581–7608, Oct. 2020, doi: https://doi.org/10.1007/s11042-020-09831-4.

[27] X. Ma and J. Wang, "Optimized Parameter Settings of Binary Bat Algorithm for Solving Function Optimization Problems", vol. 2018, pp. 1–9, May 2018, doi: https://doi.org/10.1155/2018/3847951.

[28] X.-S. Yang and S. Deb, "Cuckoo Search via Lévy flights", IEEE Xplore, Dec. 2009, doi: https://doi.org/10.1109/NABIC.2009.5393690

[29] Y. Hou, H. Gao, Z. Wang, and C. Du, "Improved Grey Wolf Optimization Algorithm and Application", Sensors, vol. 22, no. 10, p. 3810, May 2022, doi: https://doi.org/10.3390/s22103810.

[30] G.-Y. Ning and D.-Q. Cao, "Improved Whale Optimization Algorithm for Solving Constrained Optimization Problems", Discrete Dynamics in Nature and Society, vol. 2021, pp. 1– 13, Feb. 2021, doi: https://doi.org/10.1155/2021/8832251.

[31] S. Raj and S. Masood, "Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques", Procedia Computer Science, vol. 167, pp. 994–1004, 2020, doi: https://doi.org/10.1016/j.procs.2020.03.399.

[32] O. F.Y, A. J.E.T, A. O, H. J. O, O. O, and A. J, "Supervised Machine Learning Algorithms: Classification and Comparison", International Journal of Computer Trends and Technology, vol. 48, no. 3, pp. 128–138, Jun. 2017, doi: https://doi.org/10.14445/22312803/ijctt-v48p126.

[33] N. A. Ali, "Autism spectrum disorder classification on electroencephalogram signal using deep learning algorithm", IAES International Journal of Artificial Intelligence (IJ-AI), vol. 9, no. 1, p. 91, Mar. 2020, doi: https://doi.org/10.11591/ijai.v9.i1.pp91-99.

[34] S. Xu, J. Yan, and M. Hu, "A new bio-inspired metric based on eye movement data for classifying ASD and typically developing children", Signal Processing: Image Communication, vol. 94, p. 116171, May 2021, doi: https://doi.org/10.1016/j.image.2021.116171.

[35] W. Yin, S. Mostafa, and F. Wu, "Diagnosis of Autism Spectrum Disorder Based on Functional Brain Networks with Deep Learning", Journal of Computational Biology, Oct. 2020, doi: https://doi.org/10.1089/cmb.2020.0252.

[36] [36] A. Gaspar, D. Oliva, S. Hinojosa, I. Aranguren, and D. Zaldivar, "An optimized Kernel Extreme Learning Machine for the classification of the autism spectrum disorder by using gaze tracking images," Applied Soft Computing, vol. 120, p. 108654, May 2022, doi: https://doi.org/10.1016/j.asoc.2022.108654.

[37] K. Khan and Rahul Katarya, "Machine Learning Techniques for Autism Spectrum Disorder: current trends and future directions", Feb. 2023, doi: https://doi.org/10.1109/icitiit57246.2023.10068658.

[38] C. Wu et al., "Machine Learning Based Autism Spectrum Disorder Detection from Videos", Mar. 2021, doi: https://doi.org/10.1109/healthcom49281.2021.9398924.

[39] A. Yadav and D. K. Vishwakarma, "A comparative study on bio-inspired algorithms for sentiment analysis", Cluster Computing, Feb. 2020, doi: https://doi.org/10.1007/s10586-02003062-w.

[40] H. Li et al., "Newly Emerging Nature-Inspired Optimization - Algorithm Review, Unified Framework, Evaluation, and Behavioural Parameter Optimization", vol. 8, pp. 72620–72649, Apr. 2020, doi: https://doi.org/10.1109/access.2020.2987689.

[41] T. Eslami, F. Almuqhim, J. S. Raiker, and F. Saeed, "Machine Learning Methods for Diagnosing Autism Spectrum Disorder and Attention- Deficit/Hyperactivity Disorder Using Functional and Structural MRI: A Survey", Frontiers in Neuroinformatics, vol. 14, Jan. 2021, doi: https://doi.org/10.3389/fninf.2020.575999.

[42] S. Nagpal, S. Arora, S. Dey, and Shreya, "Feature Selection using Gravitational Search Algorithm for Biomedical Data", Procedia Computer Science, vol. 115, pp. 258–265, 2017, doi: https://doi.org/10.1016/j.procs.2017.09.133.

[43] X. Guo, K. C. Dominick, A. A. Minai, H. Li, C. A. Erickson, and L. J. Lu, "Diagnosing Autism Spectrum Disorder from Brain Resting-State Functional Connectivity Patterns Using a Deep Neural Network with a Novel Feature Selection Method", Frontiers in Neuroscience, vol. 11, Aug. 2017, doi: https://doi.org/10.3389/fnins.2017.00460.

[44] R. Guha, B. Chatterjee, Hassan, S. Ahmed, T. Bhattacharyya, and R. Sarkar, "Py_FS: A Python Package for Feature Selection Using Meta-Heuristic Optimization Algorithms", Advances in intelligent systems and computing, pp. 495–504, Jan. 2022, doi: https://doi.org/10.1007/978981-16-2543-5_42.

[45] S.S. Joudar, A.S. Albahri, R.A. Hamid et al. "Artificial intelligence-based approaches for improving the diagnosis, triage, and prioritization of autism spectrum disorder: a systematic review of current trends and open issues", Artificial Intelligence Review, Jun. 2023, doi: https://doi.org/10.1007/s10462-023-10536-x

Journal of Fluorescence
https://doi.org/10.1007/s10895-024-04068-y

RESEARCH

# Enhanced Red Luminescence in $Dy^{3+}/Eu^{3+}$ Co-doped Calcium Alumino Borosilicate Glasses for w-LED Applications

**Videsh Kumar[1] · Sheetal Kumari[1] · Shristy Malik[1] · Ravita[2] · Tripti Sharma[3] · Aman Prasad[3] · A. S. Rao[1]**

Received: 23 August 2024 / Accepted: 4 December 2024
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2025

## Abstract

$Dy^{3+}$ and $Eu^{3+}$ co-doped borosilicate glasses have been synthesised via melt quench technique. Amorphous behaviour of the sample has been verified by XRD study. FT-IR analysis confirmed the presence of various bonds in the host lattice. Energy band gap values of the co-doped glasses were estimated from the UV-VIS data via Tauc Plot. Excitation spectra were recorded under 575 nm ($Dy^{3+}$) and 612 nm ($Eu^{3+}$) emission wavelengths of co-doped glasses. Emission spectra of the co-doped glasses under 350 nm and 393 nm showed a significant shift towards increasing red emission with increasing $Eu^{3+}$ concentrations. Dexter's theory was applied to study the transfer of energy mechanism between $Dy^{3+}$ and $Eu^{3+}$ ions. Time decay curves were plotted for the co-doped glasses for 350 nm excitation and 576 nm emission wavelength. CIE coordinate diagram shows a shift towards red region with increasing excitation wavelength from 350 nm to 465 nm. The enhancement of red component with increasing $Eu^{3+}$ concentrations in the as -prepared borosilicate glasses can be appropriately utilised in various w-LEDs and other optoelectronic devices applications.

**Keywords** Glass · Photoluminescence · CIE coordinates · Time decay · w-LED

## Introduction

Lanthanide doped materials attract immense attention because of their widened applications in the field of solid-state laser, underwater communications, fluorescent lights, photovoltaic devices, high energy lighting devices, display devices, optical switching devices, photonics, optical data storage devices and optical fibres [1–6]. Solid-state lighting (SSL) applications and white light-emitting diodes (W-LEDs) have seen the most sustained development on rare earth (RE) ions doped materials. W-LEDs have increased in acceptance in the SSL area because of their long lifespan and low power consumption [6–8]. The lighting sector consumes 19% of the world's electricity, and energy consumption is rising globally [6]. Economically viable and eco-friendly lighting systems, such as w-LEDs, are a more environmental friendly option. Advancements in SSL technology have led to research on w-LEDs, which are far better than traditional light sources like incandescent bulbs, fluoroscent lamps [9–11].

In recent times, w-LEDs have been produced conventionally by using UV/blue GaN/InGaN LED chips coated with an epoxy, the inner side of which has one or more desired phosphors (YAG: $Ce^{3+}$ for white w-LED). The thickness of the epoxy resin and the amount of phosphor present significantly influence the w-LED emission [12]. The epoxy resin has poor temperature stability and is damaged by high temperatures, which causes it to turn yellow and lose its strength [13–16]. This approach also has certain drawbacks, including a inferior color rendering index (CRI) and low thermal quenching temperature [17]. To resolve these, RE ions doped glass matrix can be synthesized as a suitable substitute for phosphor [18, 19]. Moreover, RE ion-doped glass has low fabrication costs, a simple manufacturing process, and exceptional chemical and thermal stability compared to phosphor [20].

The selection of an appropriate glass host matrix is pivotal for optimizing luminescence in RE ions [11, 21]. Several inorganic glass matrices, including borate, silicate, and phosphate in their oxide forms ($B_2O_3$, $SiO_2$, and $P_2O_5$,

✉ Aman Prasad
amanprasad2891@gmail.com

[1] Department of Applied Physics, Delhi Technological University, New Delhi 110 042, India

[2] Department of Physics, Vaish College, Maharishi Dayanand University, Rohtak 124001, India

[3] Department of Physics and Computer Science, Dayalbagh Educational Institute, Dayalbagh, Agra 282005, India

Published online: 07 January 2025

respectively), are considered suitable host materials for rare earth ion doping [9, 13, 21, 22]. Among these, the borate glass is more suitable than others because of the superior solubility of RE ions, transparency, high thermal stability, strong corrosion resistance, and low melting point [21–23]. These attributes are critical for applications in lighting devices, solar cells, and lasers [23, 24]. However, the borate glass has high phonon energy which can be a limiting factor [4, 17]. While borate exhibits superior glass-forming properties, borosilicate glass emerges as a more suitable host material due to its lower phonon energy. Borosilicate glass offers a range of advantageous characteristics, including exceptional optical clarity, reduced melting point, and a lower coefficient of thermal expansion [21, 24]. Additionally, it demonstrates excellent thermal stability and robust resistance to corrosion. These attributes collectively contribute to its superiority as a glass host material compared to pure borate glass [21]. The combination of boron and silicon is incorporated in the glass that enhances its mechanical strength and improves the refractive index [21, 24]. Moreover, the incorporation of metal oxides such as calcium oxide (CaO) and aluminum oxide ($Al_2O_3$) further improves the glass properties by increasing the critical distance, thermal stability, mechanical strength, transparency and improves the quality of luminescence by eliminating the possibility of RE ions clustering [21, 24, 25]. CaO and $Al_2O_3$ function as property modifiers in this context [24, 25].

In this study, we synthesized $B_2O_3$- $Al_2O_3$- CaO -$SiO_2$ (BACS) glasses to develop a material with superior photonic properties, as well as enhanced mechanical and thermal durability. This glass composition holds potential for direct application in optical devices.

Among the activators, $Dy^{3+}$ has the ability to both absorb and effectively sensitize NUV-blue light to activator ions like $Tb^{3+}$ [26–28] $Eu^{3+}$ [8, 29, 30], $Sm^{3+}$ [31], and $Pr^{3+}$ [32]. The luminescence spectra of $Dy^{3+}$ exhibit multiple distinct lines, the most prominent of which are located in the blue (470–500 nm) and yellow (570–600 nm) regions and associated to the $^4F_{9/2}\rightarrow{}^6H_{15/2}$ and $^4F_{9/2}\rightarrow{}^6H_{13/2}$ transitions respectively [6, 33–35]. Multi-coloured LEDs, w-LEDs, and higher efficient photovoltaic devices can be achieved by adjusting the blue and yellow emission bands' excitation wavelengths and doping concentrations to produce white light. But because of its low CRI and high CCT due to the absence of a red component, the white light produced by $Dy^{3+}$ is neutral and impure, which limits its usefulness in possible applications [36]. Due to its strong fluorescence reliance on the environment, europium ($Eu^{3+}$), another intriguing ion with a $4f^6$ structure, is the most important luminescent probe for spectroscopic studies [37]. Another benefit is that the $^7F_0$ ground state, the lowest $^5D_0$ emitting levels, and the wave functions of the $^7F_J$ levels exhibit pure Russell-Saunders character [37]. Although some transitions are more sensitive and intense in a complex field than in a free-ion environment, the $Eu^{3+}$ emission spectrum can be observed by changing the glass composition [38]. In these borosilicate glasses, the luminescence spectrum characteristics of RE ions, such as $Eu^{3+}$ ($4f^6$) and $Dy^{3+}$ ($4f^9$), have produced fascinating and positive results [11, 30, 37, 39]. Therefore, $Eu^{3+}$ was intended to co-doped with $Dy^{3+}$ to provide warm white light for indoor and SSL applications, given the importance of these two rare-earth ions as reddish-orange and neutral-white light sources. Dexter's theory was applied to better understand the energy transfer among rare-earth ions. This theory primarily addresses the mechanisms of energy transfer between rare-earth ions in solid-state materials. This is particularly relevant to the luminescence properties of the materials studied in this work. Incorporating this theoretical framework provides deeper insights into the energy transfer processes in co-doped glasses, thereby enhancing the clarity and depth of the manuscript. In this research, we synthesized $Dy^{3+}$/$Eu^{3+}$ co-doped BACS glass to achieve warm white light emission, which is suitable for solid-state lighting applications.

## Experimental Section

The borosilicate-based host glass samples have the formula $50B_2O_3{-}15Al_2O_3{-}15CaO\text{-}20SiO_2$ and are referred to as BACS glass. The $Dy^{3+}$ and $Eu^{3+}$ co-doped glass is represented by the formula $50B_2O_3{-}15Al_2O_3$-(15-x-y)CaO-$20SiO_2$ (where x = 1.0 mol% $Dy^{3+}$ and y = 0.1, 0.5, 1.0, 1.5, and 2.0 mol% $Eu^{3+}$). These were prepared using a melt quench technique as described by Ravita et al. [24]. The precursors $B_2O_3$, $Al_2O_3$, CaO, $SiO_2$, $Dy_2O_3$, and $Eu_2O_3$ were taken in stoichiometric ratios, mixed, and ground for 45 min before being moved into the crucible. After transferring the powder mixture in the crucible, the crucible was kept in the furnace at 1100 ˚C for 2 h. The mixture was melted and quenched at ambient temperature using preheated brass plates. The quenched glass was then annealed at 300 ˚C for 3 h to diminish internal stress and strain. After cooling down in the furnace, the glass samples were collected for further characterization. The image of the synthesized glasses is shown in Fig. 1.

The XRD profile of the sample was measured by an X-ray Bruker D8 Advance Diffractometer, with X-rays generated by Cu-Kα radiation (λ = 1.54Å). The excitation and emission of photoluminescence (PL) spectra were observed with a Jasco FP-8300 spectrofluorometer. The UV-Vis-NIR absorption spectra were measured by the Jasco V-770 photo spectrometer. The photoluminescence decay spectra were

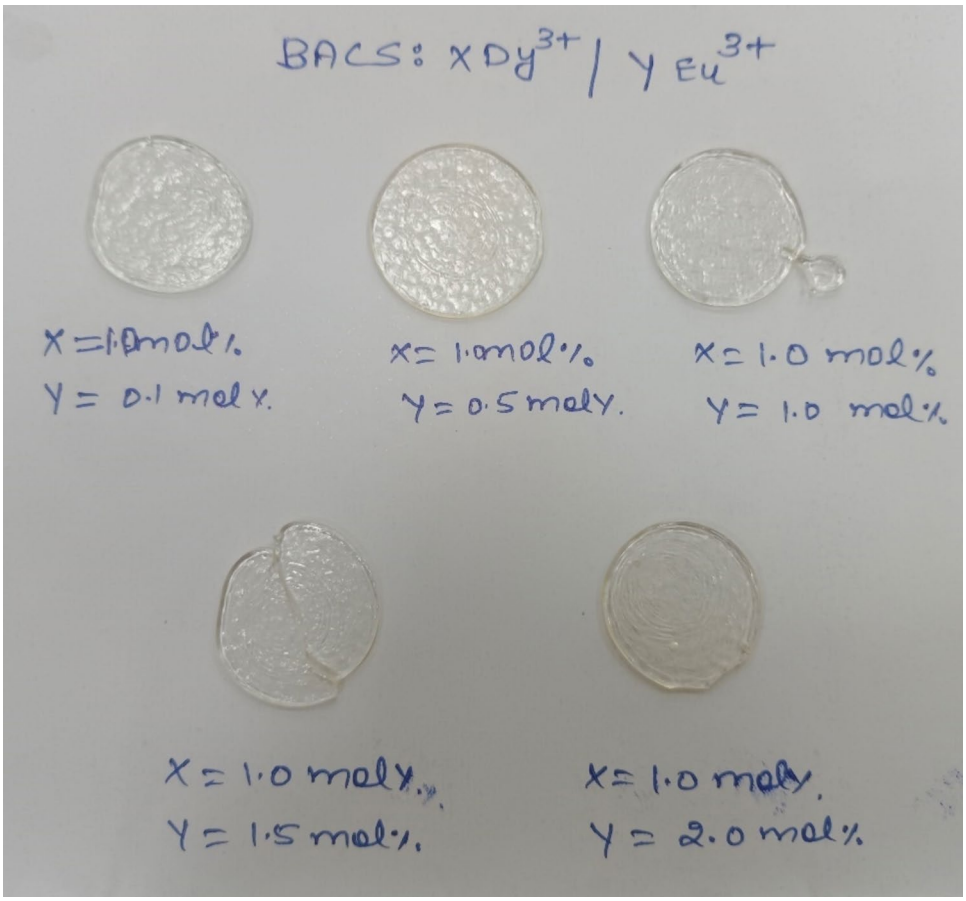



**Fig. 1** Image of the as-synthesized glasses

measured using an Edinburgh FL920 Fluorescence Spectrometer. A Perkin-Elmer Spectrometer was utilized to monitor the FT-IR spectrum.

## Results and Discussion

### XRD & FT-IR

Figure 2 depicts the XRD pattern of the as-synthesised undoped and doped BACS glasses. There is a broad hump instead of any distinct peaks in the XRD pattern, indicating amorphous nature of the prepared glass.

FT-IR data was recorded to reveal information about the vibrational groups found in the BACS glass and is depicted in Fig. 3. The data is of undoped and doped glasses which contains peaks at 456, 688, 1050, 1348, 2730, and 3494 $cm^{-1}$. The band observed at 456 $cm^{-1}$ arises from the bending modes of bridging oxygen in $Si-O-Si$ and $O-Si-O$ glass network [40]. the bending vibrations of B-O-B bond and stretching vibrations of the Si-O bond [40–43]. The band at 688 $cm^{-1}$ and 1050 $cm^{-1}$ is associated with the B-O-B bending and stretching vibrations in the tetrahedral $BO_4^-$ unit respectively [41–49]. The vibrations of asymmetric stretching B-O bond from the pyro and ortho borate groups [$BO_3$] and $BO_2O^-$ units are centered at 1348 $cm^{-1}$ [42, 50]. The small bands around 2730 and 3494 $cm^{-1}$ arises because of hydrogen bonding and water or hydroxyl group [24, 51, 52].

### Absorption Spectral & Optical Bandgap Analysis

Figure 4 illustrates the UV–Vis- NIR absorption spectra for the co-doped BACS glasses in the range of 200–1400 nm under ambient conditions. The absorption bands present in the spectra at wavelengths of 751, 800, 907, 1093, and 1273 nm were caused by $Dy^{3+}$ ions ascribed to $^6H_{15/2} \rightarrow {}^6F_{3/2}$, $^6F_{5/2}$, $^6F_{7/2}$, $^6F_{9/2}+{}^6H_{7/2}$, and $^6F_{11/2}+{}^6H_{9/2}$ transitions respectively [53, 54]. Three bands remain visible for $Eu^{3+}$ ions at wavelengths of 359, 461, and 534 nm, which correspond to the transitions of $^7F_0 \rightarrow {}^5L_6$, $^5D_2$, and $^7F_1 \rightarrow {}^5D_1$, in that order [55]. As seen in Fig. 4, the absorption of $Eu^{3+}$ at 359 nm rises with an increment in $Eu^{3+}$ concentration.

By using the conventional formulas from the prior literature, the optical band gap ($E_g$) was calculated from the absorption data as below [56]:

$$\alpha h\nu = C\left(h\nu - E_g\right)^n \quad (1)$$

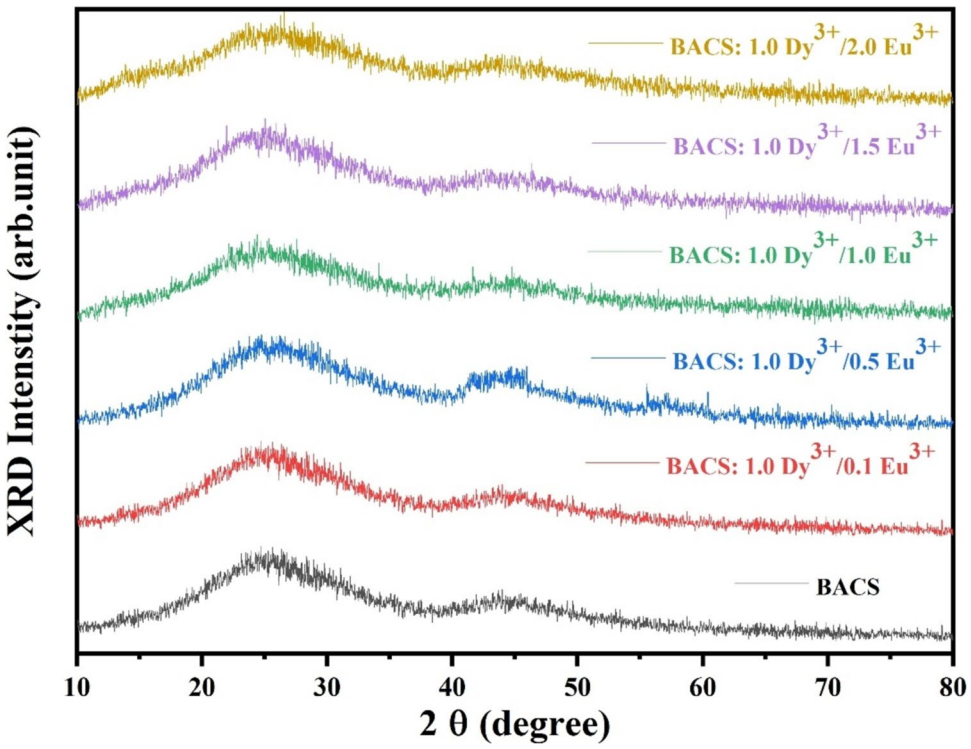



**Fig. 2** XRD profile of the synthesized glasses

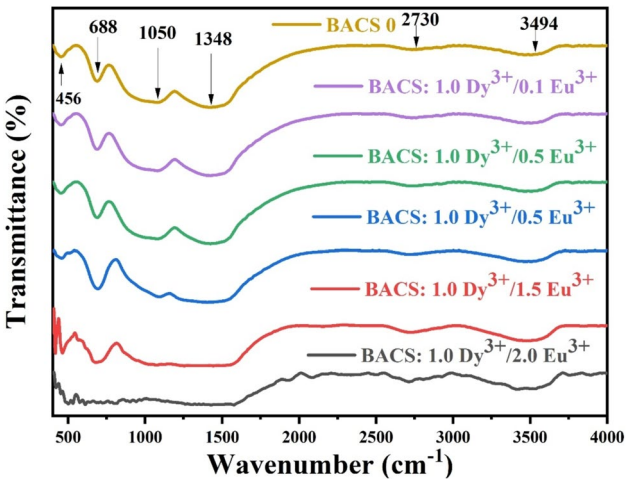



**Fig. 3** FT-IR spectra of the undoped and doped BACS glasses

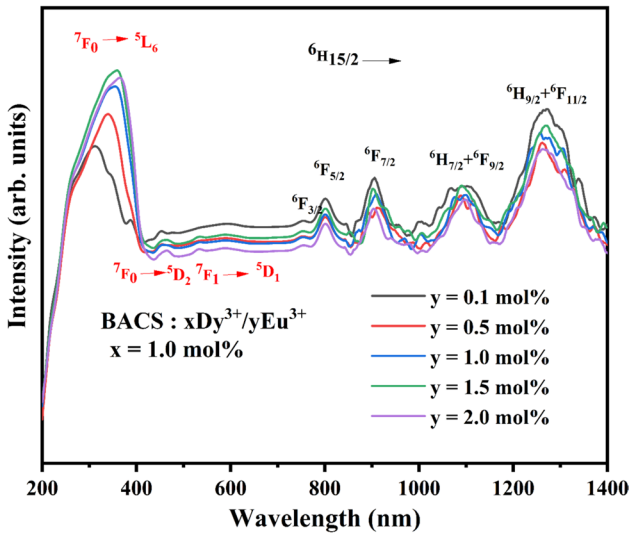



**Fig. 4** UV–Vis- NIR absorption spectra of BACS: x $Dy^{3+}$/y$Eu^{3+}$ (where x = 1.0 mol% & y = 0.1, 0.5, 1.0, 1.5 and 2.0 mol%)

The absorption coefficient represented by α, Plank's constant by h, frequency by v, bonding trailing parameter by C, and index value are represented by n. The possible values for n are 1/2, and 2, which stand for transitions that are direct allowed, and indirect allowed respectively.

With increase in $Eu^{3+}$ ions concentration, the direct bandgap of the glasses decrease, as illustrated in Fig. 5. The co-doped borosilicate glasses exhibit variations in their direct bandgap, ranging from 2.94 eV to 2.87 eV. The generation of non-bridging oxygens (NBOs) may be the cause of the bandgap drop observed with rise in doping of $Eu^{3+}$ ions. The degree of imperfection or disorder, specifically NBO and broken bonds in the as-synthesised glasses, can be determined by the Urbach's energy (ΔE) values. Therefore, luminescence properties are appropriate for low values of ΔE, which indicate less disorder. The empirical formula for Urbach energy is: [57].

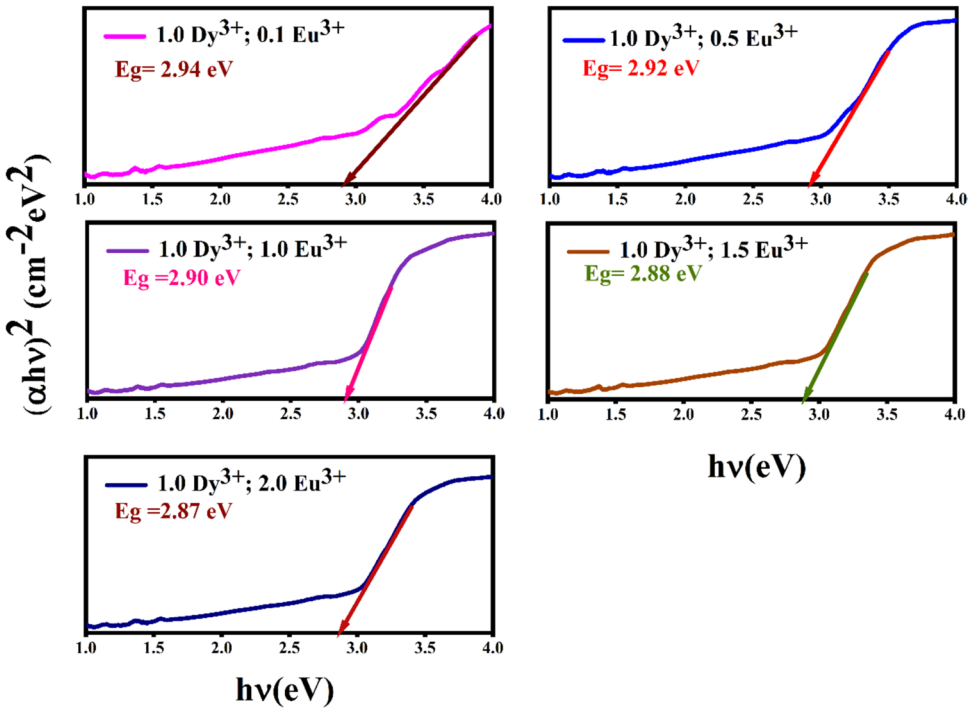



**Fig. 5** Band gap of synthesized BACS glasses

$$\alpha = Cexp\left(\frac{h\vartheta}{\triangle E}\right) \quad (2)$$

Here, $\nu$ stands for radiation frequency, and C is a constant. The value of ΔE is found by calculating the slope of ln(α) vs. photon energy graph. The assessed ΔE values for $Dy^{3+}$ / $Eu^{3+}$ co-doped BACS glasses range from 0.47 eV to 0.93 eV. The comparatively lower ΔE values suggest that the increased luminescence efficiency is due to fewer faults and disorder.

## Photoluminescence Characteristics

### Photoluminescence Studies of Dy3+/Eu3+ Co-doped BACS Glasses Under Different Spectral Wavelengths

A material's eligibility for prospective uses in display devices, SSL, w-LEDs, and other applications is determined in large part by the photoluminescence (PL) analysis. Doping concentrations of $RE^{3+}$ ions in a host matrix are crucial for achieving remarkable luminous intensity and performance. Figure 6(a & b) illustrates the photoluminescence excitation (PLE) spectra of BACS co-doped glasses with $Dy^{3+}$/ $Eu^{3+}$ ions recorded under 575 nm ($Dy^{3+}$) and 612 nm ($Eu^{3+}$) emission wavelengths in the range of 300–500 nm. As seen in Fig. 6(a), glasses exhibit eight excitation bands at 325, 350, 364, 386, 426, 453, 462 and 473 nm wavelengths out of which, the band at 462 nm ($^7F_0 \rightarrow {}^5D_2$) is present because of the $Eu^{3+}$ ions and the remaining seven bands $^6H_{15/2} \rightarrow {}^6P_{3/2}$, $^6P_{7/2}$, $^6P_{5/2}$, $^4I_{13/2}$, $^4G_{11/2}$, $^4I_{15/2}$ and $^4F_{9/2}$ are transitions due to Dy ions respectively. Of all the transitions, the one at 350 nm is comparatively stronger and can be used to excite more doped ions from the ground state. By taking 350 nm ($^6H_{15/2} \rightarrow {}^6P_{7/2}$) as the excitation wavelength, PL emission was monitored for all co-doped glasses and depicted in Fig. 7(a). The PL emission spectra with $Dy^{3+}$ excitation wavelength (350 nm) exhibit six emission bands at 481, 575, 612, 662, 700 and 753 nm out of which three peaks are due to $Dy^{3+}$ ions and three due to $Eu^{3+}$ ions [5, 27]. All transitions about these wavelengths have been shown in Fig. 7(a). $Dy^{3+}$ peak intensity starts diminishing with the increase in $Eu^{3+}$ ion concentration. It shows that the $Dy^{3+}$ ions transfer energy to the $Eu^{3+}$ ions under 350 nm excitation or $Dy^{3+}$ ions act as a suitable sensitizer for $Eu^{3+}$ ions.

Figure 6(b) shows the excitation spectra of all co-doped glasses at 612 nm emission wavelength and it exhibits the seven excitation peaks out of which two are present at 318 nm ($^6H_{15/2} \rightarrow {}^6P_{3/2}$) and 350 nm ($^6H_{15/2} \rightarrow {}^6P_{7/2}$) because of $Dy^{3+}$ ions and the remaining five peaks at 361, 381, 393, 413 and 464 nm are ascribed to the $^7F_0 \rightarrow {}^5D_4$, $^5L_7$, $^5L_6$, $^5D_3$ and $^5D_2$ due to the $Eu^{3+}$ ions respectively. All peaks are at the same position in all glasses, the only variation can be seen in excitation intensity. The band at 393 nm ($^7F_0 \rightarrow {}^5L_6$) is most intense and utilized to monitor the emission spectra as shown in Fig. 7(b). This figure also shows the emission profiles for $Dy^{3+}$/$Eu^{3+}$ co-doped ions. Now, it is recorded under 393 nm which is $Eu^{3+}$ excitation wavelength. Figure 7(b) shows the six emission bands at 481, 575, 591, 612, 650 and 700 nm wavelengths [58]. The first two are

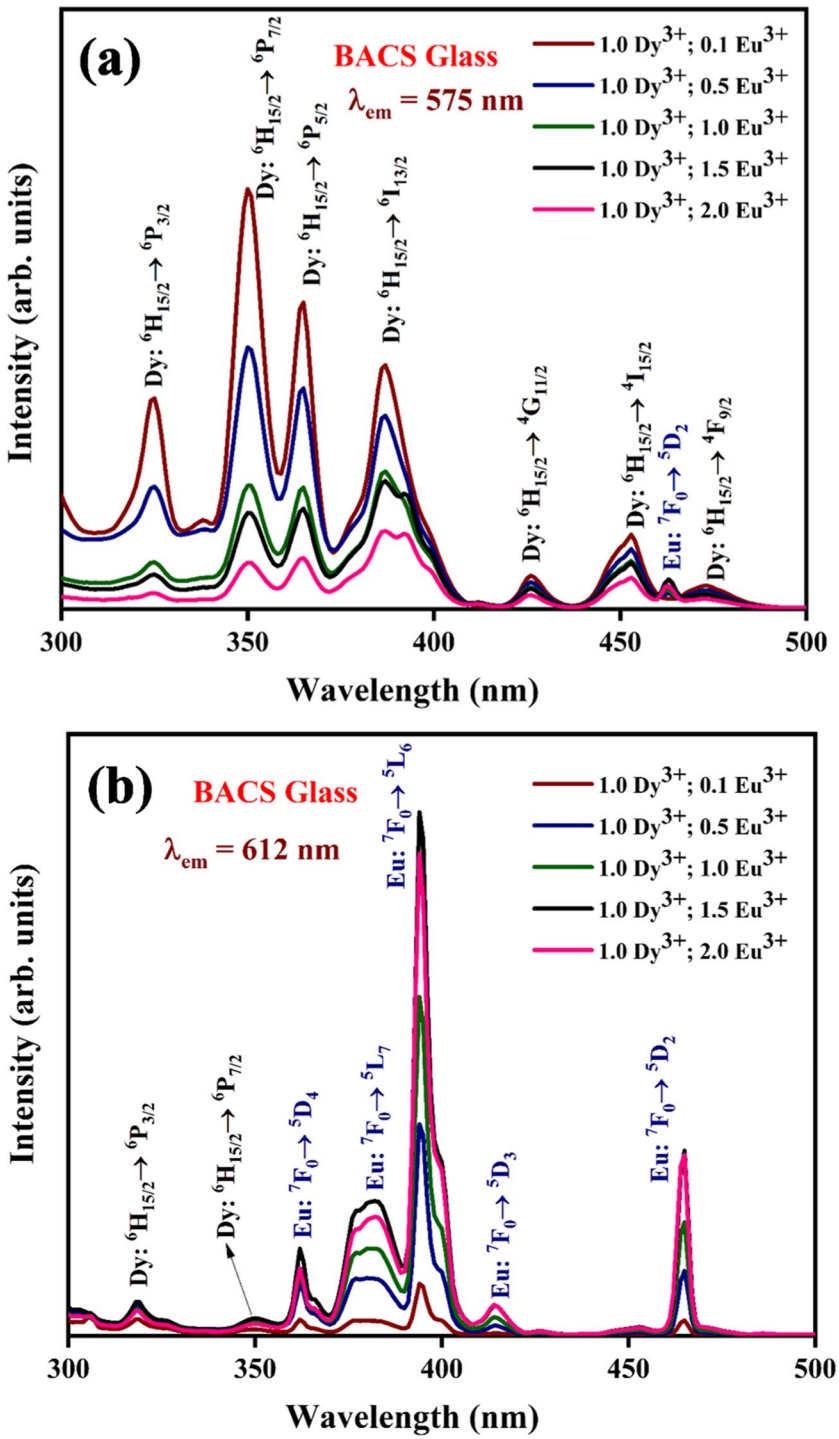



**Fig. 6** **a** PLE spectra of BACS glasses with optimum Dy3+ ions concentration and varying concentration of Eu3+ ions under 575 nm. **b** PLE spectra of BACS glasses with optimum $Dy^{3+}$ ions concentration and varying concentration of $Eu^{3+}$ ions under 612 nm

present because of $Dy^{3+}$ ions and the rest because of $Eu^{3+}$ ions. All the transitions corresponding to these wavelengths have been shown in Fig. 7(b). Again, the intensity of $Dy^{3+}$ peaks is decreasing with an increase in $Eu^{3+}$ ion concentration [59]. Figure 7(c) contains the emission profiles of BACS 1.0$Dy^{3+}$/2.0 $Eu^{3+}$ glass under different excitation wavelengths like viz. 350, 364, and 393 nm. It shows that the variation of excitation wavelengths only leads to the variation in emission intensities rather than the position of the bands. From Fig. 7(a, b & c), it is clear that whether the $Dy^{3+}$/$Eu^{3+}$ co-doped glasses are examined under $Dy^{3+}$ (350 nm) or $Eu^{3+}$ (393 nm) excitation wavelength, $Dy^{3+}$ ions are suitable for sensitizing the $Eu^{3+}$ ions.

The Dexter theory plots are displayed in Fig. 8. They provide further information about the energy transfer ( ET) from $Dy^{3+}$ ions to $Eu^{3+}$ ions. Moreover, Dexter theory

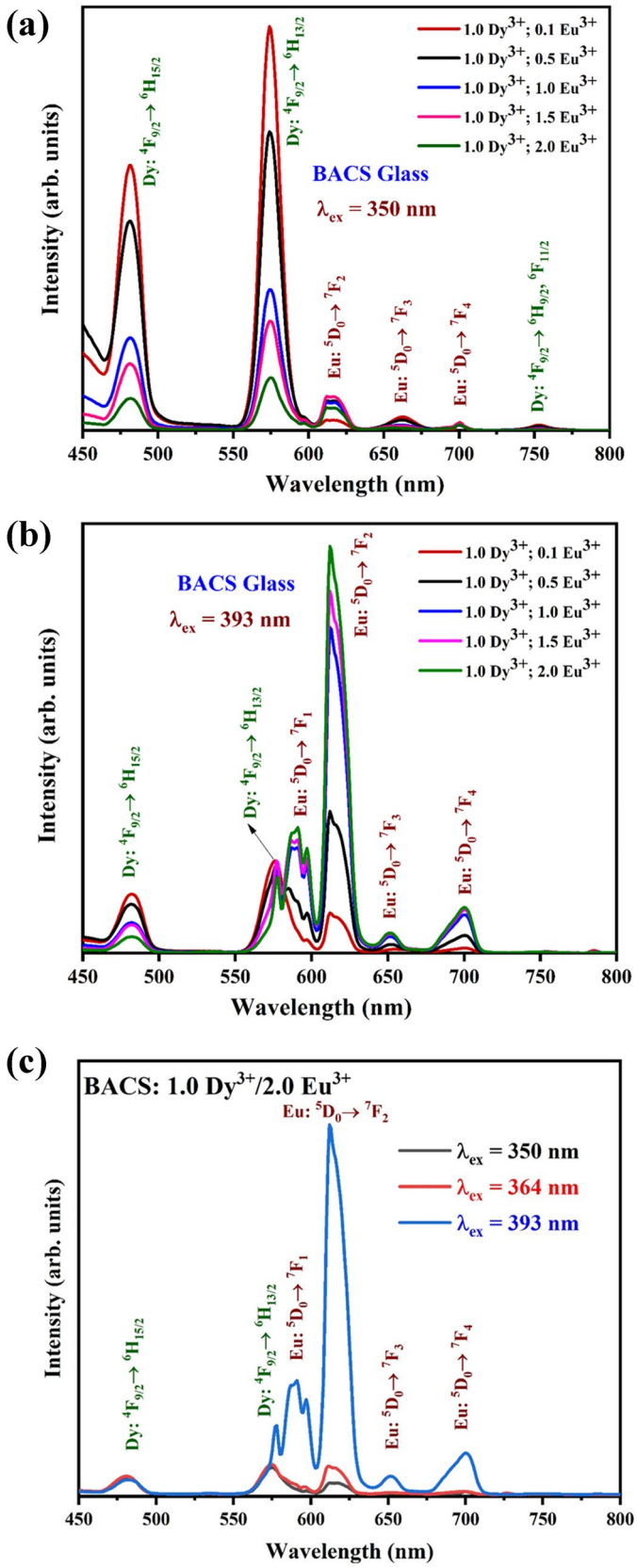



**Fig. 7** **a** Emission spectra of BACS glasses with optimum Dy3+ ions concentration and varying concentration of Eu3+ ions under 350 nm. **b** Emission spectra of BACS glasses wit optimum $Dy^{3+}$ ions concentration and varying concentration of $Eu^{3+}$ ions under 393 nm. **c** Emission spectra of BACS glasses at different wavelength i.e 350 nm, 364 nm and 393 nm

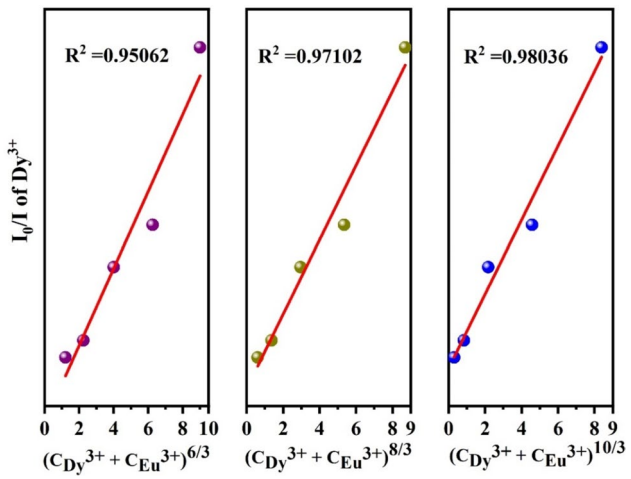



**Fig. 8** Dexter plots

was utilised in accordance with the subsequent equation to explore the kind of multipolar interaction implicated in the ET procedure [60]:

$$\frac{I_0}{I} \propto C^{\frac{\delta}{3}} \tag{3}$$

Where $I_0$ and $I$are the luminescent intensities of $Dy^{3+}$ ion without and with the $Eu^{3+}$ ion in the glass matrix respectively. $C_{Dy}{}^{3+}$ and $C_{Eu}{}^{3+}$ denotes the concentration of $Dy^{3+}$ and $Eu^{3+}$ ion and δ defines the nature of multipolar interaction which is 6, 8 and 10 for dipole-dipole (d-d), dipole-quadrupole (d-q) and quadrupole-quadrupole (q-q) interaction respectively. To understand the mechanism, a graph was plotted between $\frac{I_0}{I}$ and $\left(C_{Dy^{3+}} + C_{Eu^{3+}}\right)^{\frac{\delta}{3}}$ and is shown in Fig. 8. The best linear fit was observed for δ = 10 indicating that the non-radiative ET was through q-q interaction.

All the mechanisms involved in this ET process are shown as a schematic energy level diagram in Fig. 9. It shows all excitation, emission, non-radiative relaxations and the path involved in ET between $Dy^{3+}$ and $Eu^{3+}$. It is visually evident that the $^4F_{9/2}$ level of $Dy^{3+}$ ions is marginally bigger than the $^5D_0$ level of $Eu^{3+}$ ions; hence, there is a significantly higher chance of photons being transferred from $Dy^{3+}$ to $Eu^{3+}$ ions than there is in the reverse process [48]. ET from $Dy^{3+}$ to $Eu^{3+}$ ions happens when $Dy^{3+}/Eu^{3+}$ co-doped samples are excited at wavelengths of 350 nm and 393 nm according to the equation given below:,

$$^7F_2\ (Eu^{3+}) + {}^4F_{9/2}\ (Dy^{3+}) \rightarrow {}^5D_0\ (Eu^{3+}) + {}^6H_{13/2}\ (Dy^{3+})$$

Using the PL emission profiles depicted in Fig. 7(a & b), the ET efficiency from $Dy^{3+}$ to $Eu^{3+}$ ions has been computed using the following formula [61]:

$$\eta = 1 - \frac{I_D}{I_D + I_E} \tag{4}$$

where PL intensities of emission at 575 nm and 612 nm, respectively, are denoted by $I_D$ and $I_E$. For BACS $1.0Dy^{3+}/yEu^{3+}$ glasses, the projected ET efficiencies are tablulated in Table 1. Energy is transferred from $Dy^{3+}$ level ($^4F_{9/2}$) to $Eu^{3+}$ level ($^5D_0$) under both 350 nm and 393 nm illumination. It can be observed that the calculated energy transfer efficiencies follow the linear approximation. The energy transfer efficiency increases linearly in the same concentration range. The maximum energy transfer efficiency achieved by BACS is 86.55%, which is notably higher than the 84.57% reported in a recent study by Vikas et al. for TWKZBi: $Dy^{3+}/Eu^{3+}$ [62]. And a comparative analysis also presented in Table 2.

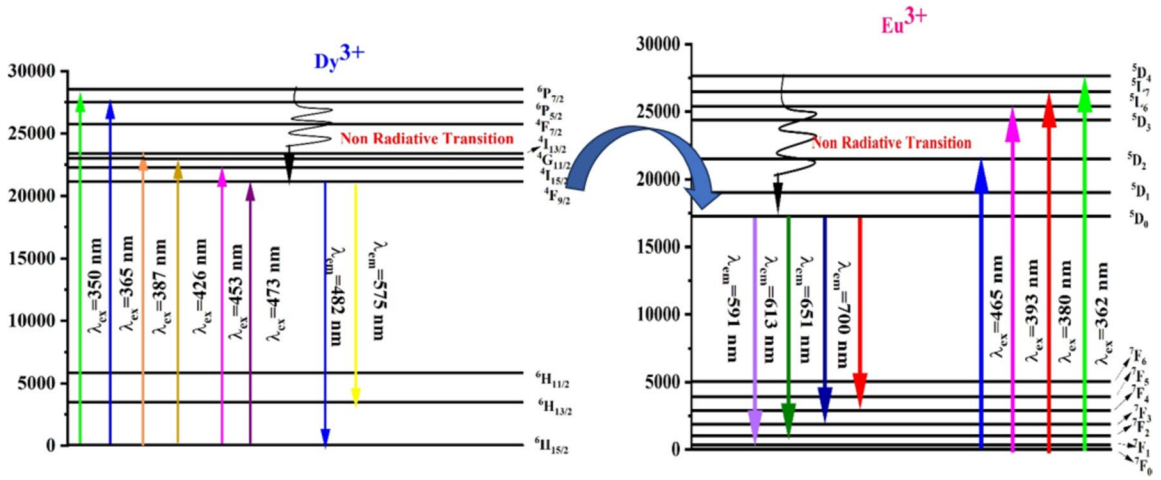



**Fig. 9** Energy transfer mechanism from $Dy^{3+}$ ions to $Eu^{3+}$ ions

**Table 1** CIE coordinates, CCT and energy transfer efficiency values

| BACS: 1.0 $Dy^{3+}$/y$Eu^{3+}$ | | | Energy Transfer Efficiency | |
|---|---|---|---|---|
| Conc. Of $Eu^{3+}$ y mol % | CIE coordinates (x,y) | CCT Values (K) | ET Efficiency at λex = 350 nm (η) (%) | ET Efficiency at λex = 393 nm (η) (%) |
| 0.1 | P (0.3391, 0.3660) | 4265 | 9.06 | 42.56 |
| 0.5 | Q (0.3306, 0.3370) | 4392 | 16.47 | 62.14 |
| 1.0 | R (0.3643, 0.3564) | 4443 | 23.87 | 81.73 |
| 1.5 | S (0.3997, 0.3740) | 4481 | 30.48 | 82.16 |
| 2.0 | T (0.4142, 0.3737) | 4524 | 42.56 | 86.55 |

**Table 2** Energy transfer efficiency comparison between BACS and reported glasses

| Glasses | Energy transfer efficiency (ηmax) | References |
|---|---|---|
| CaBAl: Dy3+/Eu3+ | 14 % | [63] |
| GSBG:Dy3+/Eu3+ | 20.93% | [64] |
| SBBCN: Dy3+/Eu3+ | 2.95% | [65] |
| ZABSDE: Dy3+/Eu3+ | 15.34% | [23] |
| BBZL: Dy3+/Eu3+ | 31.89% | [66] |
| TWKZBi: Dy3+/Eu3+ | 84.57% | [62] |
| BACS: Dy3+/Eu3+ | 86.55% | This work |

### Chromaticity and Color Coordinates

The CIE coordinates for the BACS glass are depicted in Fig. 10(a & b). By using the PL data from Fig. 7(c), we have calculated the CIE coordinates of the glass sample BACS: 1.0$Dy^{3+}$/1.0$Eu^{3+}$ under different excitations wavelengths. The CIE plot of this glass sample has been placed in Fig. 10(a) and we can observe that with the variation of excitation wavelength, CIE coordinates shift towards the red region and the values of the coordinates are A (0.2937, 0.2719), B (0.2888, 0.2478), C (0.2935, 2701), D (0.3758, 0.3257), E (0.55034, 0.4062) corresponding to 350, 362, 365, 393, 465 nm excitation wavelength respectively and tabulated in the Table 3. Figure 10(b) depicted the CIE coordinates of BACS 1.0$Dy^{3+}$/y$Eu^{3+}$ glasses under 350 nm excitation wavelength and emission wavelength 575 nm for various concentrations of $Eu^{3+}$ ions. It again shows the shift towards the red region. The corresponding values of the CIE coordinates are tabulated in Table 1. From both figures, it was concluded that by increasing $Eu^{3+}$ concentration or optimising the excitation wavelength, the red emission from the glasses could be enhanced. So, we can optimise our values by varying the concentration of dopants as well as excitation wavelengths to make the materials suitable in particular applications.

### Decay Profiles

To explore energy transfer, the PL decay lifetime has been measured. The decay curves for $Dy^{3+}$ / $Eu^{3+}$ ions co-doped BACS glasses were observed at 350 nm excitation, as depicted in Fig. 11. The curves fitted well to a first-order exponential decay Eqs. [53, 67].

$$I = I_0 + A_1 * \exp\left(-\frac{t}{\tau}\right) \quad (5)$$

where τ stands for the decay time and $A_1$ represent a constant. The experimental lifetime ($\tau_{exp}$) values for BACS: x $Dy^{3+}$ /y$Eu^{3+}$ (x = 1.0 mol% & y = 0.1, 0.5, 1.0, 1.5 & 2.0 mol%) glasses were 0.81, 0.79, 0.75, 0.72, and 0.68 ms respectively. The measured $\tau_{exp}$ values show that the calculated lifespan values decrease with the $Eu^{3+}$ doping concentration. The decrease in lifetime after doping with $Eu^{3+}$ ions consistently demonstrates the energy transfer from $Dy^{3+}$ to $Eu^{3+}$ ions. This is because, when energy transfer occurs, the excited $Dy^{3+}$ ions lose energy non-radiatively to nearby $Eu^{3+}$ ions, leading to a faster decay of $Dy^{3+}$ emission and, consequently, a shorter lifetime.

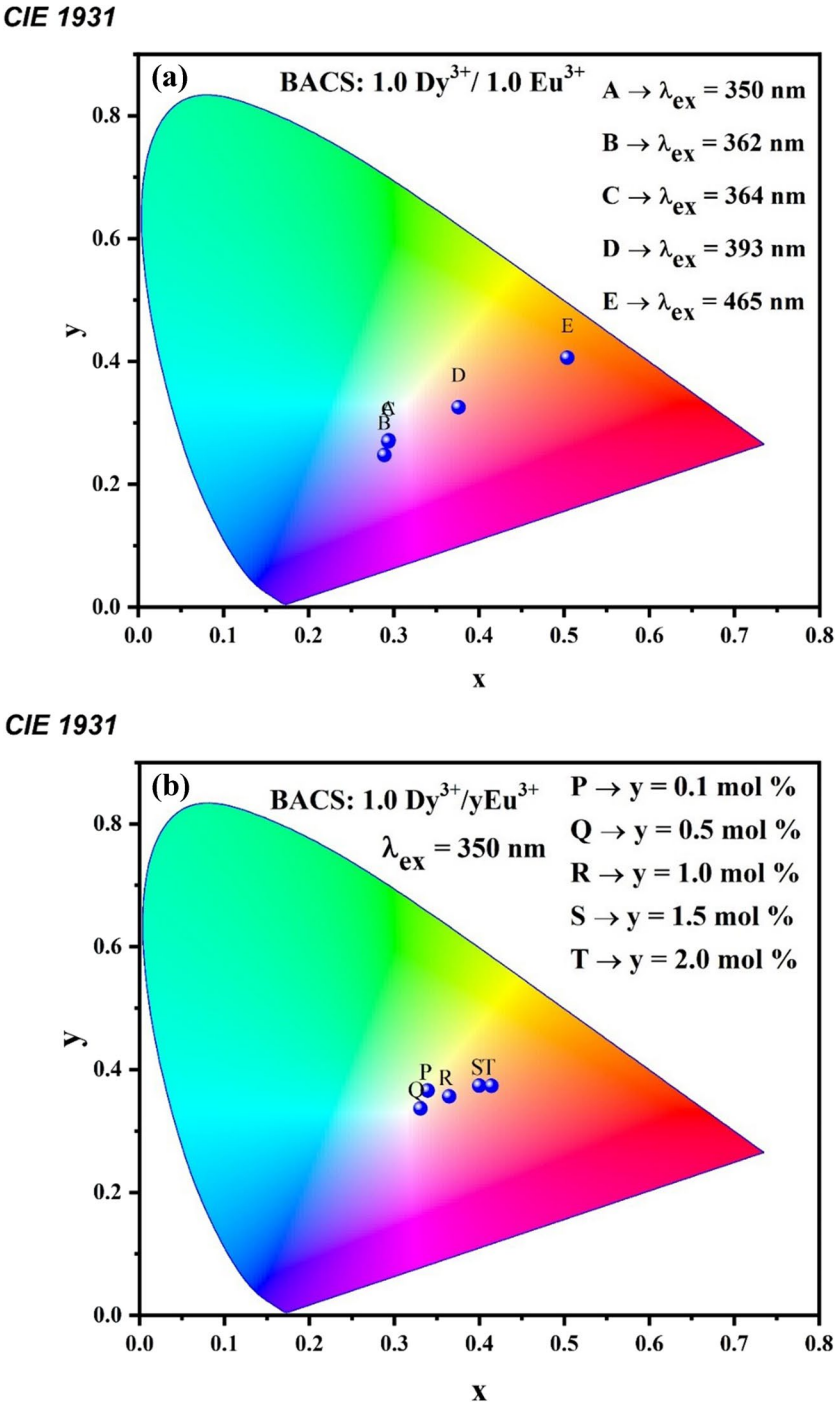



**Fig. 10** **a** CIE plot of synthesized BACS: 1.0 $Dy^{3+}$/1.0 $Eu^{3+}$ at different excitation wavelengths. **b** CIE plot of synthesized glass BACS :1.0 $Dy^{3+}$/y $Eu^{3+}$ (where y = 0.1, 0.5, 1.0, 1.5 and 2.0 mol%) at excitation wavelength 350 nm and emission wavelength 575 nm

**Table 3** CIE plot of synthesized BACS: 1.0 $Dy^{3+}$/1.0 $Eu^{3+}$ at different excitation wavelengths

| BACS: 1.0 $Dy^{3+}$/1.0$Eu^{3+}$ | |
|---|---|
| Excitation Wavelength | CIE coordinates (x,y) |
| 350 nm | A (0.2937, 0.2719) |
| 362 nm | B (0.2888, 0.2478) |
| 365 nm | C (0.2935, 0.2701) |
| 393 nm | D (0.3758, 0.3257) |
| 465 nm | E (0.5034, 0.4062) |

## Conclusion

Borosilicate glasses doped with $Dy^{3+}$/$Eu^{3+}$ were synthesized using a melt quench technique. The glasses show an amorphous nature with no short or long range of order as was confirmed by the XRD analysis. Various vibrational groups were confirmed to be present in the host lattice via FT-IR study. UV-Vis-NIR analysis of the co-doped glasses was conducted in 200–1400 nm range. Various bands were seen in the spectra where transitions from $^6H_{15/2}$ to $^6F_{3/2}$,

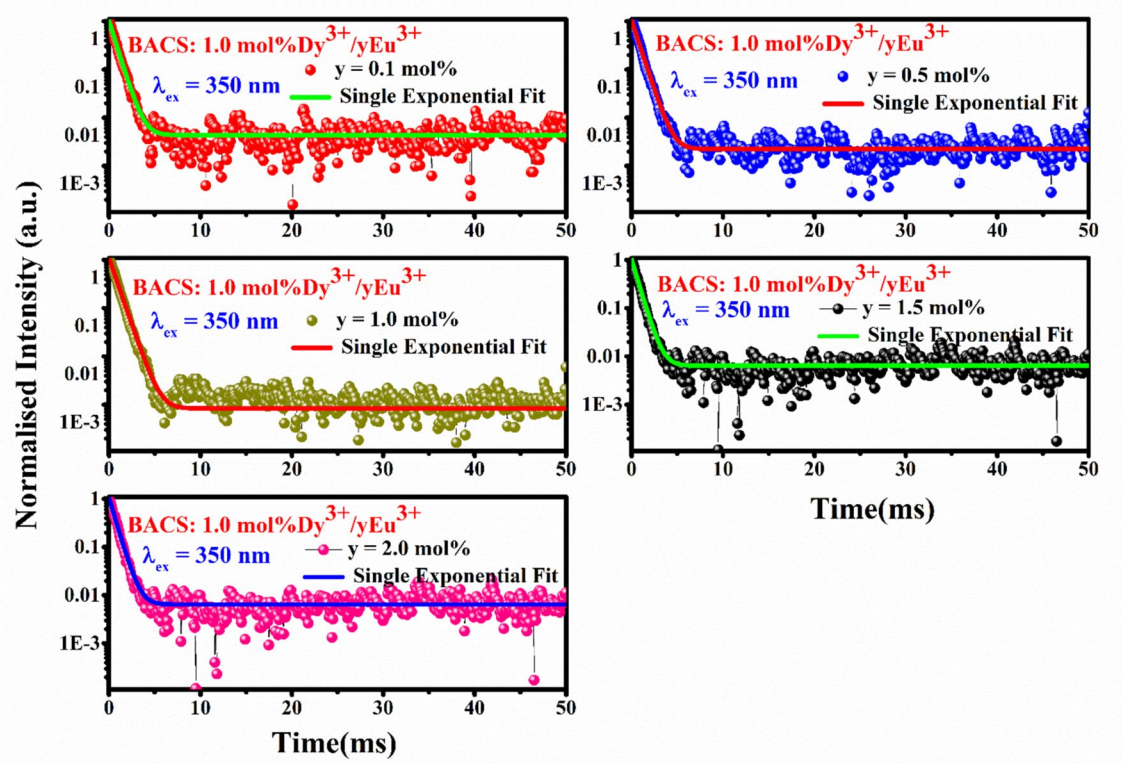



**Fig. 11** Decay profile of BACS :1.0 $Dy^{3+}$/y $Eu^{3+}$ (where y = 0.1, 0.5, 1.0, 1.5 and 2.0 mol%) at excitation wavelength 350 nm

$^6F_{5/2}$, $^6F_{7/2}$, $^6F_{9/2}$+$^6H_{7/2}$, and $^6F_{11/2}$+$^6H_{9/2}$, corresponded to $Dy^{3+}$ and $^7F_0 \rightarrow {}^5L_6$, $^5D_2$ and $^7F_1 \rightarrow {}^5D_1$ corresponded to $Eu^{3+}$. Energy bandgap values of the co-doped glasses were calculated and lie in 2.87–2.94 eV range. Excitation spectra were recorded at 575 nm and 612 nm emission wavelengths. In both of these spectra, many bands corresponding to both $Dy^{3+}$ and $Eu^{3+}$ ions were seen wherein peaks at 350 nm ($^6H_{15/2} \rightarrow {}^6P_{7/2}$ ) and 393 nm ($^7F_0 \rightarrow {}^5L_6$) had the highest intensity. Emission spectra of the samples were then recorded under 350 nm and 393 nm excitation wavelengths. It was observed that with increasing $Eu^{3+}$ ion concentrations, the luminescence peaks of $Dy^{3+}$ decreased, thereby showing energy transfer from $Dy^{3+}$ to $Eu^{3+}$ ions. Dexter theory was applied to study the mechanism of energy transfer between the dopant ions. It was observed that the energy transfer between the ions is quadrupole-quadrupole in nature. Decay curves for 575 emission wavelength under 350 nm excitation were recorded. The curves fit well to a single order exponential function and the lifetime values of the samples lie in the range of 0.81–0.68 ms. The CIE coordinates for the co-doped glasses under 350 nm excitation show a gradual shift towards the red region with increasing $Eu^{3+}$ concentrations. Also, shift towards the red region is again observed with increasing excitation wavelengths from 350 nm to 465 nm thereby signalling energy transfer from $Dy^{3+}$ to $Eu^{3+}$ ions. Therefore, the enhancement of red luminescence under varying excitation wavelengths as well as concentrations of $Eu^{3+}$ ions these $Dy^{3+}$/$Eu^{3+}$ co-doped glasses can be utilised in w-LEDs, display and other optoelectronic applications.

**Acknowledgements** One of the authors, Ms. Sheetal Kumari is very much thankful to CSIR-UGC for the award of a Junior Research Fellowship (JRF) (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).

**Author Contributions** V.K. : Conceptualisation, characterization, writing S.K. : Synthesis, characterization, writing S.M. : Data Analysis R.P. : Characterization, data analysis T.S. : Data Analysis A.P. : Conceptualisation, review, writing A.S. R. : Review, supervision.

**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.

## Declarations

**Competing Interests** The authors declare no competing interests.

## References

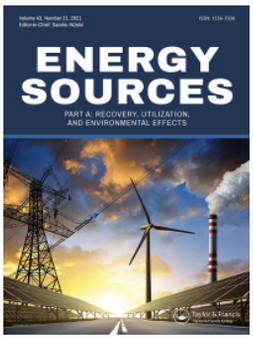

1. Suneel Kumar A, Narendrudu T, Suresh S, Sambasiva Rao MV, Chinna Ram G, Krishna Rao D (2017) Physical and spectroscopic features of cobalt ions in multi-component $CaF_2$–ZnO–$Bi_2O_3$–$P_2O_5$ glass ceramics. J Alloys Compd 699:392–400. https://doi.org/10.1016/j.jallcom.2016.12.334
2. Wang JS, Vogel EM, Snitzer E (1994) Tellurite glass: a new candidate for fiber devices. Opt Mater 3:187–203. https://doi.org/10.1016/0925-3467(94)90004-3
3. Anand Pandarinath M, Upender G, Narasimha Rao K, Suresh Babu D (2016) Thermal, optical and spectroscopic studies of boro-tellurite glass system containing ZnO. J Non-cryst Solids 433:60–67. https://doi.org/10.1016/j.jnoncrysol.2015.11.028
4. Kaur A, Khanna A, González-Barriuso M, González F (2021) Thermal and light emission properties of rare earth ($Eu^{3+}$, $Dy^{3+}$ and $Er^{3+}$), alkali ($Li^+$, $Na^+$ and $K^+$) and $Al^{3+}$-doped barium tellurite and boro-tellurite glasses. J Mater Sci: Mater Electron 32:17266–17281. https://doi.org/10.1007/s10854-021-06228-3
5. Reddy DVK, Rao TS, Taherunnisa S, Suchocki A, Zhydachevskyy Y, Piasecki M, Reddy MR (2019) 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 513:167–182. https://doi.org/10.1016/j.jnoncrysol.2019.03.011
6. Milewska K, Maciejewski M, Žitňan M, Velázquez JJ, Galusek D, Sadowski W, Kościelska B (2024) Tunable emission and energy transfer of $B_2O_3$–$Bi_2O_3$–$AlF_3$ glass system doped with $Eu^{3+}$/$Dy^{3+}$. J Lumin 269:120440. https://doi.org/10.1016/j.jlumin.2024.120440
7. Caldiño U, Lira A, Meza-Rocha AN, Camarillo I, Lozada-Morales R (2018) Development of sodium-zinc phosphate glasses doped with $Dy^{3+}$, $Eu^{3+}$ and $Dy^{3+}$/$Eu^{3+}$ for yellow laser medium, reddish-orange and white phosphor applications. J Lumin 194:231–239. https://doi.org/10.1016/j.jlumin.2017.10.028
8. Meza-Rocha AN, Camarillo I, Lozada-Morales R, Caldiño U (2017) Reddish-orange and neutral/warm white light emitting phosphors: $Eu^{3+}$, $Dy^{3+}$ and $Dy^{3+}$/$Eu^{3+}$ in potassium-zinc phosphate glasses. J Lumin 183:341–347. https://doi.org/10.1016/j.jlumin.2016.11.068
9. Erol E, Vahedigharehchopogh N, Kibrisli O, Ersundu MC, Ersundu AE (2021) Recent progress in lanthanide-doped luminescent glasses for solid-state lighting applications - a review. J Phys Condens Matter 33. https://doi.org/10.1088/1361-648X/ac22d9
10. Lewandowski T, Seweryński C, Walas M, Łapiński M, Synak A, Sadowski W, Kościelska B (2018) Structural and luminescent study of $TeO_2$-BaO-$Bi_2O_3$-Ag glass system doped with $Eu^{3+}$ and $Dy^{3+}$ for possible color-tunable phosphor application. Opt Mater 79:390–396. https://doi.org/10.1016/j.optmat.2018.03.031
11. Rajesh D, Brahmachary K, Ratnakaram YC, Kiran N, Baker AP, Wang GG (2015) Energy transfer based emission analysis of $Dy^{3+}$/$Eu^{3+}$ co-doped ZANP glasses for white LED applications. J Alloys Compd 646:1096–1103. https://doi.org/10.1016/j.jallcom.2015.05.138
12. Deopa N, Rao AS (2022) Structural and luminescence characteristics of thermally stable $Dy^{3+}$ doped oxyfluoride strontium zinc borosilicate glasses for photonic device applications. Opt Laser Technol 154:108328. https://doi.org/10.1016/j.optlastec.2022.108328
13. Zhou W, Wang G, Zheng X, Yu L, Zhang J, Qiu Z, Lian S (2019) Tunable colors and applications of $Dy^{3+}$/$Eu^{3+}$ co-doped CaO-$B_2O_3$-$SiO_2$ glasses. J Am Ceram Soc 102:5890–5898. https://doi.org/10.1111/jace.16442
14. Zhang R, Lin H, Yu Y, Chen D, Xu J, Wang Y (2014) A new-generation color converter for high-power white LED: transparent $Ce^{3+}$: YAG phosphor-in-glass. Laser Photonics Revi 8:158–164. https://doi.org/10.1002/lpor.201300140
15. Peng Y, Wang S, Li R, Li H, Cheng H, Chen M, Liu S (2016) Luminous efficacy enhancement of ultraviolet-excited white light-emitting diodes through multilayered phosphor-in-glass. Appl Opt 55:4933. https://doi.org/10.1364/ao.55.004933
16. Fukui T, Sakuta H, Mishiro K, Miyachi T, Kamon K, Hayashi H, Nakamura N, Uchida Y, Kurai S, Taguchi T (2008) Development of white light emitting diodes by multi-layered red, green, and blue phosphors by excited near-ultraviolet light emitting diodes. J Light Visual Environ 32:43–45. https://doi.org/10.2150/jlve.32.43
17. Deopa N, Sahu MK, Rani PR, Punia R, Rao AS (2020) Realization of warm white light and energy transfer studies of $Dy^{3+}$/$Eu^{3+}$ co-doped $Li_2O$-PbO-$Al_2O_3$-$B_2O_3$ glasses for lighting applications. J Lumin 222:117166. https://doi.org/10.1016/j.jlumin.2020.117166
18. Bouchouicha H, Panczer G, De Ligny D, Guyot Y, Baesso ML, Andrade LHC, Lima SM, Ternane R (2016) Synthesis and luminescent properties of $Eu^{3+}$/$Eu^{2+}$ co-doped calcium aluminosilicate glass-ceramics. J Lumin 169:528–533. https://doi.org/10.1016/j.jlumin.2014.11.054
19. Hari Babu B, Ravi Kanth VV Kumar (2014) White light generation in Ce3 + $Tb^{3+}$ $Sm^{3+}$ codoped oxyfluoroborate glasses. J Lumin 154:334–338. https://doi.org/10.1016/j.jlumin.2014.05.010
20. Kumar M, Ratnakaram YC, Vijayalakshmi RP, Raman TR (2022) Probing into structural and spectroscopic properties of $Dy^{3+}$ doped and $Eu^{3+}$/$Dy^{3+}$ co-doped bismuth phosphate ($BiPO_4$) glass ceramics with different modifier fluorides. J Lumin 245:118786. https://doi.org/10.1016/j.jlumin.2022.118786
21. Sharma S, Kaur S, Rao AS, Kishor K (2023) Thermally stable illuminating characteristics of $Tb^{3+}$/$Sm^{3+}$ ions activated BaO–SrO–$Al_2O_3$–$B_2O_3$–$SiO_2$ glasses for photonic applications. Opt Mater 145:1–9. https://doi.org/10.1016/j.optmat.2023.114446
22. Maheshwari K, Prasad A, Pilania R, Mahamuda S, Tayal Y, Rao AS (2023) Thermally stable multi-color emitting $Dy^{3+}$/$Eu^{3+}$ co-doped BaO–ZnO–$Li_2O$–$P_2O_5$ glasses for w-LEDs. J Mater Sci: Mater Electron 34:1–15. https://doi.org/10.1007/s10854-023-11394-7
23. Monisha M, Sayyed MI, Mazumder N, Arayro J, Kamath SD (2023) Color tuneability behaviour and energy transfer analysis on $Dy^{3+}$-$Eu^{3+}$ co-doped glasses for NUV-WLEDs application. J Mater Sci: Mater Electron 34:1–13. https://doi.org/10.1007/s10854-022-09602-x
24. Ravita A, Prasad P, Rohilla R, Bajaj, Anu R, Punia AS, Rao (2023) Luminescence studies on $Dy^{3+}$ Doped Calcium Aluminum Borosilicate (CABS) glasses for white light emission and applications in w-LEDs. J Fluoresc 34:1783–1792. https://doi.org/10.1007/s10895-023-03389-8
25. Deopa N, Rao AS, Choudhary A, Saini S, Navhal A, Jayasimhadri M, Haranath D, Vijaya G, Prakash (2018) Photoluminescence investigations on $Sm^{3+}$ ions doped borate glasses for tricolor w-LEDs and lasers. Mater Res Bull 100:206–212. https://doi.org/10.1016/j.materresbull.2017.12.019
26. Pisarska J, Kos A, Pietrasik E, Pisarski WA (2014) Energy transfer from Dy $^{3+}$ to $tb^{3+}$ in lead borate glass. Mater Lett 129:146–148. https://doi.org/10.1016/j.matlet.2014.05.035
27. Ravita, Rao AS (2021) Effective energy transfer from $Dy^{3+}$ to $Tb^{3+}$ ions in thermally stable KZABS glasses for intense green emitting device applic. J Lumin 239:118325. https://doi.org/10.1016/j.jlumin.2021.118325
28. Meza-Rocha AN, Lozada-Morales R, Speghini A, Bettinelli M, Caldiño U (2016) White light generation in $Tb^{3+}$/$Eu^{3+}$/$Dy^{3+}$ triply-doped $Zn(PO_3)_2$ glass. Optical materials 51:128–132. https://doi.org/10.1016/j.optmat.2015.11.032

29. Deng Y, Yi S, Huang J, Xian J, Zhao W (2014) White light emission and energy transfer in $Dy^{3+}/Eu^{3+}$ co-doped $BaLa_2WO_7$ phosphors. Mater Res Bull 57:85–90. https://doi.org/10.1016/j.materresbull.2014.05.035
30. Carrillo-Torres RC, Saavedra-Rodríguez G, Alvarado-Rivera J, Caldiño U, Sánchez-Zeferino R, Alvarez-Ramos ME (2019) Tunable emission and energy transfer in $TeO_2$-$GeO_2$-ZnO and $TeO_2$-$GeO_2$-$MgCl_2$ glasses activated with $Eu^{3+}/Dy^{3+}$ for solid state lighting applications. J Lumin 212:116–125. https://doi.org/10.1016/j.jlumin.2019.04.008
31. Jose A, Krishnapriya T, Jose TA, Joseph C, Unnikrishnan NV, Biju PR (2021) Color tunable luminescence characteristics and energy transfer analysis of $Dy^{3+}/Sm^{3+}$ co-doped multicomponent borosilicate glasses. Scripta Mater 203:114088. https://doi.org/10.1016/j.scriptamat.2021.114088
32. Naresh V, Lee N (2021) $Dy^{3+}/Pr^{3+}$ co-doped fluoro-borosilicate glasses: energy transfer induced color-tunable luminescence. Mater Res Bull 142:111381. https://doi.org/10.1016/j.materresbull.2021.111381
33. Blasse G, Grabmaier BC (1994) A General introduction to luminescent materials. Luminescent Mater 1–9. https://doi.org/10.1007/978-3-642-79017-1_1.
34. Blasse G, Grabmaier BC (1994) Luminescent Materials - Blasse and Grabmaier – 1994.pdf. Springer Ser Fluoresc 7:121
35. Fhoula M, Koubaa T, Dammak M (2020) White photoluminescence and energy transfer properties of dysprosium and europium singly and codoped $Na_2ZnP_2O_7$ phosphors. Optics Laser Technolog 130:106352. https://doi.org/10.1016/j.optlastec.2020.106352
36. Vijayalakshmi L, Kumar KN, Baek JD (2023) Multicolor and warm white luminescence from $Dy^{3+}/Eu^{3+}$ co-activated glasses for indoor and solid-state lighting applications. Ceram Int 49:5013–5021. https://doi.org/10.1016/j.ceramint.2022.11.315
37. Ofelt GS (1963) Structure of the f6 configuration with application to rare-earth ions. J Chem Phys 38:2171–2180. https://doi.org/10.1063/1.1733947
38. Oomen EWJL, Van Dongen AMA (1989) Europium (III) in oxide glasses: dependence of the emission spectrum upon glass composition. Journal of non-crystalline solids 111(2–3):205–213. https://doi.org/10.1016/0022-3093(89)90282-2
39. Jang S-H, Motome Y (2024) Exploring rare-earth Kitaev magnets by massive-scale computational analysis. Commun Mater 1–8:1. https://doi.org/10.1038/s43246-024-00634-w
40. Wahab SAA, Matori KA, Aziz SHA, Zaid MHM, Kechik MMA, Azman AZK, Khaidir REM, Khiri MZA, Effendy N (2020) Effect of ZnO on the phase transformation and optical properties of silicate glass frits using rice husk ash as a $SiO_2$ source. J Mater Res Technol 9:11013–11021. https://doi.org/10.1016/j.jmrt.2020.08.005
41. Rada S, Pascuta P, Culea M, Maties V, Rada M, Barlea M, Culea E (2009) The local structure of europium-lead-borate glass ceramics. J Mol Struct 924–926.https://doi.org/10.1016/j.molstruc.2008.12.032
42. Subhadra M, Kistaiah P (2012) Infrared and Raman spectroscopic studies of alkali bismuth borate glasses: evidence of mixed alkali effect. Vib Spectrosc 62:23–27. https://doi.org/10.1016/j.vibspec.2012.07.001
43. Ahmed MR, Ashok B, Ahmmad SK, Hameed A, Chary MN, Shareefuddin M (2019) Infrared and raman spectroscopic studies of $Mn^{2+}$ ions doped in strontium alumino borate glasses: describes the role of $Al_2O_3$. Spectrochimica Acta - Part A: Mol Biomolecular Spectroscopy. 210:308–314. https://doi.org/10.1016/j.saa.2018.11.053
44. Singh GP, Kaur P, Kaur S, Singh DP (2011) Role of $V_2O_5$ in structural properties. Mater Phys Mech 12:58–63
45. Doweidar H, Saddeek YB (2009) FTIR and ultrasonic investigations on modified bismuth borate glasses. J Non-cryst Solids 355:348–354. https://doi.org/10.1016/j.jnoncrysol.2008.12.008
46. Rani S, Sanghi S, Agarwal A, Kishore N (2010) Study of structure and $Li^+$ ions dynamics in presence of $Fe_2O_3$ in $Bi_2O_3{\cdot}B_2O_3$ glasses. Solid State Phenom 161:51–61
47. Gautam C, Yadav AK, Singh AK (2012) A Review on infrared spectroscopy of borate glasses with effects of different additives. ISRN Ceramics. 2012:1–17. https://doi.org/10.5402/2012/428497
48. Marzouk MA, Elbatal HA, Abdelghany AM (2003) Effect of $Gd_2O_3$ and $sm_2O_3$ addition on the properties of $CeO_2$. J Korean Inst Electr Electron Mater Eng 16:979–986. https://doi.org/10.4313/jkem.2003.16.11.979
49. Ouis MA, Abd-Allah WM, Sallam OI (2022) Gamma ray interaction with soda lime silicate glasses doped with $V_2O_5$, CuO or SrO. Appl Phys A: Mater Sci Process 128:1–12. https://doi.org/10.1007/s00339-022-05522-z
50. Balachander L, Ramadevudu G, Shareefuddin M, Sayanna R, Venudharc YC (2013) IR analysis of borate glasses containing three alkali oxides. ScienceAsia. 39:278–283. https://doi.org/10.2306/scienceasia1513-1874.2013.39.278
51. Esawii HA, Salama E, Sayed El-ahll L, Moustafa M, Saleh HM (2022) High impact tungsten-doped borosilicate glass composite for gamma and neutron transparent radiation shielding. Prog Nucl Energy 150:104321. https://doi.org/10.1016/j.pnucene.2022.104321
52. Gao C, Gao Q, Bao X, Li Y, Teramoto A, Abe K (2011) Preparation and in vitro bioactivity of novel mesoporous borosilicate bioactive glass nanofibers. J Am Ceram Soc 94:2841–2845. https://doi.org/10.1111/j.1551-2916.2011.04434.x
53. Anilkumar K, Damodaraiah S, Babu S, Prasad VR, Ratnakaram YC (2019) Emission spectra and energy transfer studies in $Dy^{3+}$ and $Dy^{3+}/Eu^{3+}$ co-doped potassium fluorophosphate glasses for white light applications. J Lumin 205:190–196. https://doi.org/10.1016/j.jlumin.2018.09.007
54. Niu L, Zhu C, Zhang M, He Z, Liu J, Xu Q, Wang Z (2019) Eu and Dy doped borophosphosilicate glass-ceramics for near ultraviolet based light-emitting diode applications. Vacuum 169:108877. https://doi.org/10.1016/j.vacuum.2019.108877
55. Yu P, Guo W, Zhang R, Su L, Xu J (2021) White and tunable light emission in $Eu^{3+}$, $Dy^{3+}$ codoped phosphate glass. Opt Mater 114:110939. https://doi.org/10.1016/j.optmat.2021.110939
56. Kumari S, Anu, Prasad A, Rohilla P, Rao AS (2023) Prospective applications of thermally stable $Dy^{3+}$ doped potassium zinc strontium borate (KZSB) glasses in w-LEDs. J Mater Sci: Mater Electron 34:1–17. https://doi.org/10.1007/s10854-023-10272-6
57. Monisha M, Mazumder N, Lakshminarayana G, Mandal S, Kamath SD (2021) Energy transfer and luminescence study of $Dy^{3+}$ doped zinc-aluminoborosilicate glasses for white light emission. Ceram Int 47:598–610. https://doi.org/10.1016/j.ceramint.2020.08.167
58. Caldiño U, Muñoz G, Camarillo HI, Speghini A, Bettinelli M (2015) Down-shifting by energy transfer in $Tb^{3+}/Dy^{3+}$ co-doped zinc phosphate glasses. J Lumin 161:142–146. https://doi.org/10.1016/j.jlumin.2015.01.001
59. Ravita ASR (2021) Effective sensitization of $Eu^{3+}$ visible red emission by $Sm^{3+}$ in thermally stable potassium zinc alumino borosilicate glasses for photonic device applications. J Lumin 118689. https://doi.org/10.1016/j.jlumin.2021.118689
60. Jha K, Jayasimhadri M (2018) Effective sensitization of Eu3 + and energy transfer in $Sm^{3+}/Eu^{3+}$ co-doped ZPBT glasses for CuPc based solar cell and w-LED applications. J Lumin 194:102–107. https://doi.org/10.1016/j.jlumin.2017.09.049
61. Ramachari D, Moorthy LR, Jayasankar CK (2014) Energy transfer and photoluminescence properties of $Dy^{3+}$/ $tb^{3+}$ co-doped oxy fl

uorosilicate glass – ceramics for solid-state white lighting. Ceram Int 40:11115–11121. https://doi.org/10.1016/j.ceramint.2014.03.136
62. Sangwan V, Jayasimhadri M, Haranath D (2025) Thermally stable white light emission and energy transfer analysis of tungstate-tellurite glasses co-activated with $Dy^{3+}/Eu^{3+}$ for optoelectronic applications. J Mol Struct 1322:140577. https://doi.org/10.1016/j.molstruc.2024.140577
63. Lodi TA, Sandrini M, Medina AN, Barboza MJ, Pedrochi F, Steimacher A (2018) Dy:Eu doped CaBAl glasses for white light applications. Opt Mater 76:231–236. https://doi.org/10.1016/j.optmat.2017.12.043
64. Liu R, Wang D, Chen M, Liu L, Zhou Y, Zeng F, Su Z (2021) Luminescence, energy transfer properties of $Dy^{3+}/Eu^{3+}$ coactivated neutral and warm white emissions GSBG glasses. J Lumin 237:118180. https://doi.org/10.1016/j.jlumin.2021.118180
65. An J, Zhang S, Liu R, Hu G, Zhang Z, Qiu Y, Zhou Y, Zeng F, Su Z (2021) Luminescent properties of $Dy^{3+}/Eu^{3+}$ doped fluorescent glass for white LED based on oxyfluoride matrix. J Rare Earths 39:26–32. https://doi.org/10.1016/j.jre.2020.01.013
66. Xu S, Fu J, Liu X, Yuan M, Zhang C, Wang H, Liu G, Cui S (2023) Analysis of thermal chromaticity stability, energy transfer and colour tunablity of $Dy^{3+}$-$Eu^{3+}$ doped BBZL glasses for high-power W-LED applications. Ceramics International 49:5228–5238. https://doi.org/10.1016/j.ceramint.2022.10.04
67. Manasa P, Ran F, Basavapoornima C, Depuru SR, Jayasankar CK (2020) Optical characteristics of ($Eu^{3+}$,$Nd^{3+}$) co-doped lead-fluorosilicate glasses for enhanced photonic device applications. J Lumin 223:117210. https://doi.org/10.1016/j.jlumin.2020.117210

**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.



Energy Sources, Part A: Recovery, Utilization, and Environmental Effects

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/ueso20

# Experimental investigation of Homogeneous Charge Compression Ignition (HCCI) engine with vaporized fuel injection

Naveen Kumar, Pawan Kishore Jha & R. C. Singh

**To cite this article:** Naveen Kumar, Pawan Kishore Jha & R. C. Singh (2025) Experimental investigation of Homogeneous Charge Compression Ignition (HCCI) engine with vaporized fuel injection, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 47:1, 3074-3092, DOI: 10.1080/15567036.2024.2442057

**To link to this article:** https://doi.org/10.1080/15567036.2024.2442057

Published online: 20 Jan 2025.

Submit your article to this journal

Article views: 12

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=ueso20

ENERGY SOURCES, PART A: RECOVERY, UTILIZATION, AND ENVIRONMENTAL EFFECTS
2025, VOL. 47, NO. 1, 3074–3092
https://doi.org/10.1080/15567036.2024.2442057

Taylor & Francis
Taylor & Francis Group

Check for updates

# Experimental investigation of Homogeneous Charge Compression Ignition (HCCI) engine with vaporized fuel injection

Naveen Kumar, Pawan Kishore Jha, and R. C. Singh

Centre for Advanced Studies and Research in Automotive Engineering, Department of Mechanical Engineering, Delhi Technological University, Delhi, India

**ABSTRACT**

Homogeneous Charge Compression Ignition (HCCI) has recently been a promising option for meeting stringent emissions norms. HCCI engine results in higher thermal efficiency and simultaneous reduction in NOx emissions. The present study aimed to investigate the various performance and emission characteristics of HCCI engines and compare them with baseline data. Port fuel injection and external fuel vaporizer have been used to achieve HCCI combustion for Diesel, Ethanol, and Compressed Natural Gas (CNG) at various loads in the present work. The $NO_X$ and $CO_2$ emissions were decreased by 16% and 17%, respectively, at maximum load using HCCI mode operation. The results obtained from this exhaustive trial provide new dimensions to compare the HCCI combustion with different kinds of fuels at varying loads and open new pathways of optimizing the combustion stages and comprehending the intricate relationships throughout different operating situations to exploit HCCI technology's advantages fully.

**ARTICLE HISTORY**
Received 22 March 2024
Revised 18 November 2024
Accepted 9 December 2024

**KEYWORDS**
CNG; emission; engine performance; ethanol; HCCI diesel engine

## Introduction

Homogeneous Charge Compression Ignition (HCCI) technology is an innovative internal combustion engine design approach that marries the best features of gasoline and diesel engines to enhance fuel efficiency and reduce harmful emissions (Mogra and Gupta 2020). In HCCI engines, a well-mixed air-fuel mixture is compressed until it auto-ignites without needing a spark plug or direct injection at the point of combustion, as in a diesel engine's operation, but with a homogeneous mixture similar to that of a gasoline engine. This method allows for more efficient combustion and significantly lower emissions of nitrogen oxides ($NO_X$) and particulate matter (PM), as it operates at lower temperatures and pressures, ensuring a more uniform temperature distribution across the combustion chamber. Despite its advantages, challenges such as controlling ignition timing across various conditions and extending the operational range remain (Zhao et al. 2019). Nonetheless, HCCI represents a promising advancement toward achieving more sustainable and environmentally friendly internal combustion engines (Pandey et al. 2018).

The promise of HCCI technology in improving fuel economy and engine performance is well recognized. However, the cost-effective pathway to achieving these benefits remains a challenge. The intrinsic advantages of HCCI enhanced fuel efficiency and reduced emissions are offset by technical hurdles that need to be addressed to make this technology viable for widespread adoption in vehicle production (Nishi, Kanehara, and Iida 2016). These challenges include mastering the precise control of ignition timing across varying operational conditions and expanding the technology's limited operational range to suit a more comprehensive array of driving scenarios.

**CONTACT** Pawan Kishore Jha Jha.pawan81@gmail.com Centre for Advanced Studies and Research in Automotive Engineering, Department of Mechanical Engineering, Delhi Technological University, Main Bawana Road, Delhi 110042, India
© 2025 Taylor & Francis Group, LLC

The compression ratio plays a critical role in determining the thermal efficiency and emissions characteristics of HCCI engines. A higher compression ratio can enhance thermal efficiency by achieving better combustion of the air-fuel mixture. However, it may also increase the formation of nitrogen oxides ($NO_X$) due to higher combustion temperatures. Thus, optimizing the compression ratio is essential to balance efficiency and emissions in HCCI engines with vaporized fuel injection. The compression ratio affects the cylinder pressure, temperature, and heat release rate, influencing the combustion process and emissions. Therefore, a detailed investigation into the impact of varying compression ratios on the performance and emissions of HCCI engines with vaporized fuel injection is essential. This understanding is supported by Yamagiwa (Yamagiwa et al. 2023), discussed the implications of compression ratio adjustments in HCCI and CAI engines.

Extensive research and development are crucial to transitioning HCCI from a promising concept to a practical reality in daily vehicle production. Studies focusing on HCCI engines have explored numerous aspects of their operation, aiming to optimize performance, and overcome existing limitations. This includes investigating the impact of different variables on the engine's efficiency and emissions, such as air-fuel ratio, compression ratio, engine load and speed, and the temperature and pressure conditions within the combustion chamber. Moreover, the potential of HCCI technology is more comprehensive than that of traditional petrol and diesel fuels. Research has also delved into alternative fuels that could be utilized in HCCI engines to maximize efficiency and minimize environmental impact. These alternatives include biofuels, ethanol, and natural gas, among others. Each fuel type presents advantages and challenges within the HCCI context, uniquely influencing the engine's performance, efficiency, and emissions profile (Riyadi et al. 2023).

Optimizing ignition timing is vital for achieving efficient and stable combustion in HCCI engines. Proper ignition timing ensures the air-fuel mixture is ignited at the optimal point during the compression stroke, maximizing combustion efficiency. Therefore, further optimization of ignition timing can enhance the combustion efficiency of the HCCI engine. Ignition timing significantly influences the combustion phasing, heat release rate, and stability. Thus, a comprehensive study on optimizing ignition timing in HCCI engines can provide valuable insights into improving combustion efficiency and reducing emissions. This is supported by Fathi et al. (Fathi, Jahanian, and Shahbakhti 2021), the authors explored the impact of ignition timing on HCCI combustion, emphasizing its critical role in achieving efficient and stable engine operation. Yang et al. (2020) were investigated the effects of ignition timing on HCCI combustion and highlighted its crucial role in optimizing engine performance and emissions control.

This broad spectrum of research is instrumental in pushing the boundaries of HCCI technology toward practical application. By thoroughly understanding the interactions between different fuels, operational parameters, and engine design, researchers can devise strategies to mitigate the current limitations of HCCI engines. This ongoing research aims to enable the mass production and implementation of HCCI technology in the automotive industry, offering a more efficient and environmentally friendly alternative to conventional internal combustion engines. Future research in this field will undoubtedly play a pivotal role in realizing the full potential of HCCI engines, marking a significant step forward in pursuing sustainable automotive technologies (Sharma, Rao, and Murthy 2016; Verma et al. 2022).

It was observed by Verma et al. (2022) that the $NO_X$ emission was found to be approximately 2500 ppm when the CI engine was fueled with diesel. However, CNG, as fuel in the HCCI engine, has $NO_X$ lower than 15 ppm. The research reveals that higher emission control could be possible in HCCI engines. Test fuels for the experiments included mixtures of ethanol and diethyl ether in varying volumetric ratios, such as 30% ethanol–70% diethyl ether (E30/D70), 40% ethanol–60% diethyl ether (E40/D60), 50% ethanol–50% diethyl ether (E50/D50), and 100% diethyl ether (DEE) and to analyze the role of blending in HCCI engine the Polat et al. (2016) found during the investigation that blends of diethyl ether (DEE) with ethanol in an HCCI engine shown the improvements in indicated thermal efficiency about 22%.

The investigation by Biswas et al. (2023) is a noteworthy exploration of the use of alternative fuels within the context of HCCI technology to enhance engine performance while minimizing environmental impact. In their study, biodiesel derived from Mahua (a type of tree found in South Asia) was initially used as an ignition source due to its suitable properties for initiating combustion under compression. Concurrently, CNG and ethanol were employed as the primary energy sources, effectively replacing a significant portion of the biodiesel during combustion. This innovative approach leverages the distinct advantages of each fuel type – biodiesel for its bio-based, renewable qualities and favorable ignition characteristics and CNG/ethanol for its cleaner combustion profiles (Mishra and Naveen Kumar 2017).

The research findings indicated that the combined use of CNG and biodiesel leads to lower emissions of non-hydrocarbon compounds (NHC), carbon monoxide (CO), and carbon dioxide ($CO_2$) compared to conventional diesel fuel. This outcome can be attributed to several factors inherent to the combustion properties of these fuels. Biodiesel, an oxygenated fuel, can enhance combustion, leading to more complete fuel burning and thus reducing CO and particulate matter emissions. Conversely, CNG is a cleaner-burning fuel that inherently produces less $CO_2$ and $NO_X$ emissions due to its higher hydrogen-to-carbon ratio, contributing to a reduction in greenhouse gas emissions and improving air quality (Vuilleumier et al. 2016).

Port fuel injection (PFI) in HCCI engines ensures a well-mixed air-fuel mixture by injecting fuel into the intake air before it enters the combustion chamber, promoting even combustion and reducing emissions. Preheating of intake air aids in reaching the fuel-air mixture to the necessary temperature for the auto-ignition, especially under conditions where compression heat alone might be insufficient. Data acquisition systems collect real-time operational data to analyze engine performance and combustion efficiency, facilitating heat release analysis and exhaust emission measurements. This comprehensive approach allows for optimizing HCCI engines to improve efficiency and minimize environmental impact by adjusting engine parameters based on the gathered data and analysis (Kumar Maurya and Kumar Agarwal 2015).

To retrofit a diesel engine for HCCI operation, additional components like a fuel vaporizer unit, fuel pump, fuel tank, fuel injector, and a data acquisition system are integrated. The fuel vaporizer plays a crucial role by vaporizing the fuel before it enters the cylinder, ensuring it mixes thoroughly with the incoming air in the intake manifold to form a homogeneous air-fuel mixture. As the engine operates, this mixture is drawn into the combustion chamber during the suction stroke, where it auto-ignites due to the high temperature and pressure conditions characteristic of HCCI combustion. This process leads to simultaneous combustion at multiple points within the chamber, promoting a more uniform and lower peak temperature than traditional diesel combustion. As a result, nitrogen oxide ($NO_X$) and carbon dioxide ($CO_2$) emissions are significantly reduced due to the cooler overall combustion temperature. However, there might be a slight increase in hydrocarbon (HC) and carbon monoxide (CO) emissions, but these remain within acceptable limits for compression ignition engines. This setup illustrates how adapting diesel engines to HCCI technology can enhance efficiency and reduce specific emissions.

The research gap addressed in our study is that more work needs to be conducted on Homogeneous Charge Compression Ignition (HCCI) engines using liquid fuel, mainly where an external fuel vaporizer is utilized. This gap is significant because HCCI engines have the potential to offer higher efficiency and lower emissions compared to conventional engines. However, using liquid fuels in HCCI engines with external vaporizers still needs to be explored. Our research aims to fill this gap by investigating the performance and emissions characteristics of HCCI engines under these specific conditions, thereby contributing valuable insights and advancements in engine technology.

Maintaining stable combustion at higher loads is a significant challenge in HCCI engines due to the risk of misfire or incomplete combustion. Potential strategies to address this issue include optimizing the air-fuel mixture, adjusting the compression ratio, and controlling the ignition timing. Recent studies support these strategies. For instance, Liu et al. (2018) discuss the optimization of the air-fuel

mixture to improve combustion stability at higher loads, while Niklawy et al. (2022) highlight the importance of adjusting the compression ratio and ignition timing in achieving stable combustion.

Minimizing heat losses is crucial for improving the overall efficiency of HCCI engines. Heat losses can occur through various mechanisms, including cooling systems, exhaust, and frictional losses. Potential strategies to reduce heat losses include improving the engine design, optimizing the combustion process, and using advanced thermal management techniques. Recent studies support these strategies. For example, Wang et al. (2019) discuss the impact of advanced engine design and thermal management techniques on reducing heat losses, while Christensen et al. (2016) highlight the importance of optimizing the combustion process to minimize energy dissipation.

Controlling combustion phasing is vital for optimizing HCCI engine performance and reducing emissions. Specific methods to control combustion phasing include adjusting the ignition timing, optimizing the air-fuel mixture, and using advanced combustion control strategies. Recent studies support these methods. For instance, Duan et al. (2020) highlight the role of optimizing the air-fuel mixture. Additionally, Onishi et al. (2018) explore advanced combustion control strategies and their impact on HCCI engine performance.

Improving air-fuel mixture homogeneity is essential for enhancing HCCI engine performance. Advancements in air-fuel mixture homogeneity can be achieved through optimizing the fuel injection strategy, improving the intake air distribution, and using advanced intake air preheating techniques. These advancements are supported by recent studies. For instance, Wang et al. (2021) discuss the optimization of fuel injection strategies to achieve better mixture homogeneity, while Chen et al. (2019) highlight the importance of improving intake air distribution. Li et al. (2018) also explore the impact of advanced intake air preheating techniques on air-fuel mixture homogeneity.

Achieving complete combustion is a challenge in HCCI engines due to the homogeneous nature of the air-fuel mixture and the risk of misfire or incomplete combustion. Innovative approaches to overcome this limitation include optimizing the air-fuel mixture, controlling the ignition timing, and using advanced combustion control strategies. These approaches are supported by recent studies. For instance, Zheng et al. (2019) discuss optimizing the air-fuel mixture to reduce the risk of incomplete combustion.

Reducing hydrocarbon (HC) emissions without compromising combustion efficiency is crucial for improving the environmental performance of HCCI engines. Innovative approaches to achieve this include optimizing the air-fuel mixture, controlling the combustion phasing, and using advanced exhaust after-treatment technologies. These approaches are supported by recent studies. For instance, Gao et al. (2020) discuss optimizing the air-fuel mixture to reduce HC emissions, while Kokjohn et al. (2013) highlight the importance of controlling combustion phasing. Additionally, Cheng et al. (2017) further explore the use of advanced exhaust after-treatment technologies to reduce HC emissions.

Exhaust gas recirculation (EGR) can significantly optimize the performance of HCCI engines with vaporized fuel injection by reducing $NO_X$ emissions and improving efficiency. EGR reduces the oxygen concentration in the intake air, lowering the combustion temperature and reducing the formation of $NO_X$. These benefits are supported by recent studies. For instance, Sjoberg et al. (2005) discuss the impact of EGR on reducing $NO_X$ emissions and improving combustion stability, while Zhang et al. (2020) highlight the effectiveness of EGR in optimizing the performance and efficiency of HCCI engines with vaporized fuel injection.

Modifying the design of the HCCI engine can improve volumetric efficiency across varying operating conditions by optimizing the intake and exhaust systems, adjusting the valve timing, and improving the combustion chamber design. Recent studies support these design modifications. For instance, Xie et al. (2018) discuss the impact of optimizing intake and exhaust systems on volumetric efficiency, while Zhu et al. (2017) highlight the importance of valve timing adjustments. Additionally, Lee et al. (2019) explore the improvements in combustion chamber design that can enhance volumetric efficiency.

Reducing $NO_X$ emissions while maintaining high efficiency is a critical challenge in HCCI engines. Specific techniques include optimizing the air-fuel mixture, controlling the combustion phasing, using

exhaust after-treatment technologies, and implementing advanced combustion control strategies. These techniques are supported by recent studies. For instance, Zhang et al. (2019) discuss the optimization of the air-fuel mixture to minimize $NO_X$ emissions.

To deepen the understanding of HCCI combustion and its implications for engine performance and emissions control, additional studies or experiments could be conducted, including a detailed analysis of the impact of different fuels and fuel blends on HCCI combustion and emissions, as discussed by Saxena and Bedoya (2013). Investigation of the effects of varying engine operating conditions, such as load, speed, and temperature, on HCCI combustion and emissions is also crucial, as highlighted by Reitz et al. (2020). Additionally, experimental validation of advanced combustion control strategies, such as closed-loop combustion control and predictive combustion control, can provide significant insights, as explored by Tan et al. (2018).

Therefore, there is a need to investigate the potential of HCCI with port fuel injection method and external fuel vaporizer for improving the performance and emissions of internal combustion engines. The performance and emissions of HCCI combustion, such as specific energy consumption, brake thermal efficiency, volumetric efficiency, exhaust gas temperature (EGT), and emissions profiles ($NO_X$, $CO_2$, HC), should be experimentally investigated using different liquid and gaseous fuels.

The primary goal of this research is to experimentally compare the performance and emission characteristics of Homogeneous Charge Compression Ignition (HCCI) combustion using different liquid and gaseous fuels in a modified single-cylinder diesel engine.

### *Novelty*

Our research aims to contribute to the field by investigating the performance and emission characteristics of HCCI combustion using various liquid and gaseous fuels. Existing research has explored the advantages of HCCI combustion in terms of efficiency and emissions reduction only.

Our study offers several novel aspects. Firstly, we conducted experiments using a modified single-cylinder diesel engine to achieve HCCI mode via port fuel injection strategy using an external fuel vaporizer, providing practical insights into real-world implementation. This approach differs from previous studies that rely on simulation or laboratory setups.

Secondly, our comparative assessment includes conventional diesel fuel also, offering a comprehensive analysis of diesel fuel with HCCI combustion. This comparative aspect adds novelty to our research as it provides valuable insights into the performance variations and emission characteristics.

Furthermore, our investigation delves into the effects of load variation on HCCI combustion, shedding light on the engine's behavior under different operating conditions. This aspect extends beyond existing literature, often focusing on specific load conditions or steady-state operations.

## Materials and methodology

### *Engine setup*

The experiment was carried out on a single-cylinder, four-stroke water-cooled diesel engine Figure 1. The technical specifications of the engine are given in Table 1.

Eddy current dynamometer was coupled with the engine to measure the output power. Temperature sensors LM35 were used to measure the temperature. An exhaust gas analyzer was used for measuring HC (ppm), $NO_X$ (ppm), CO (% by vol.), $CO_2$ (%by vol.), and $O_2$ (% by vol.).

The experimental setup comprised crucial components, including a fuel vaporizer, fuel injection system, data acquisition system, and an exhaust gas analyzer. The modification to the engine intake manifold facilitated mixing intake air with vaporized diesel. The fuel vaporizer,

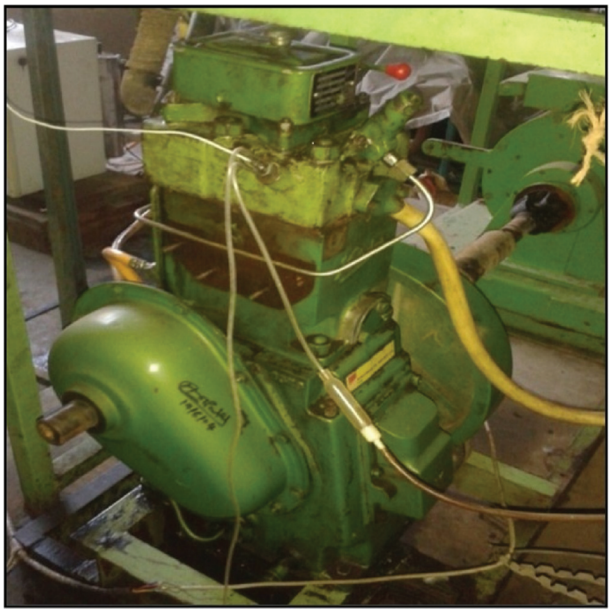


**Figure 1.** Test engine.

**Table 1.** Engine specification.

| Engine | Kirloskar (AVI) |
|---|---|
| Dynamometer | Eddy current, air-cooled |
| Bore (mm) | 87.5 |
| Stroke (mm) | 110 |
| Displacement ($cm^3$) | 661 |
| Compression ratio | 17.5 |
| R.P.M. | 1500 |
| H.P. | 5.2 |

a key element, featured a primary vaporizing chamber made of copper. An electric band heater enveloped the external surface of the vaporizing chamber, generating sufficient heat for the vaporization of various liquid fuels such as diesel, ethanol, or butanol. This configuration allowed for precise control over the vaporization process and ensured that the vaporized fuel could seamlessly mix with the intake air in the modified intake manifold. The integration of these components provided a comprehensive setup for experimentation, allowing researchers to assess the impact of vaporized fuel on combustion characteristics and emissions in the engine. The external fuel injection system of the HCCI engine consists of a fuel pump, fuel tank, fuel injector, and an injector control circuit (Data Acquisition System). The fuel pump supplies the liquid fuel from the fuel tank to the fuel injector. The fuel injector was placed inside the fuel vaporizer. When the receiver receives the optical rays passing through the window of the pulley, it signals the fuel injector to inject the fuel for the defined time into the vaporizer surface. When injected diesel comes into contact with the hot surfaces of the vaporizer chamber, the diesel gets vaporized and mixed with the intake air to make a homogeneous mixture. During the suction stroke, this homogeneous mixture will suck inside the combustion chamber.

In the HCCI engine, the temperature and pressure in the combustion chamber rise as the air-fuel mixture is compressed during the compression stroke. This compression process increases the temperature of the homogeneous charge, which is a well-mixed combination of fuel and air. As the compression stroke nears its end, the homogeneous charge reaches its

auto-ignition temperature. At this critical point, spontaneous ignition occurs without needing an external ignition source, such as a spark plug. Auto-ignition initiates the combustion of the homogeneous charge, resulting in the release of energy and the engine's propulsion. This auto-ignition characteristic distinguishes HCCI engines from traditional spark-ignition and compression-ignition engines, providing a unique approach to achieving efficient and controlled combustion.

## Modifications of diesel engine for the HCCI engine

### Installation of injector and fuel vaporizer

The vaporizer is crucial in converting liquid fuel into a gaseous form, optimizing its mixture with the incoming air to create a homogeneous charge. An external injector atomizes the fuel before it enters the vaporizing chamber, where atomized fuel transforms into a gaseous state. This gaseous form allows for easier and more thorough mixing with the inlet air, facilitating the formation of a homogeneous charge. The vaporization process is vital for uniformly mixing the liquid fuel with the inlet air in the engine cylinder.

The vaporizer is strategically installed on the extension pipe of the engine's inlet manifold to accomplish this. This extended pipe is equipped with a heater controlled by the Data Acquisition (DAQ) system. Precise temperature control is essential to prevent the unintended ignition of fuel within the vaporizer. The controlled heating ensures that the fuel undergoes vaporization without the risk of ignition, maintaining a safe and efficient process. Control circuit sketch of heater is shown in Figure 2.

At the end of the vaporizer, an injector is attached at an angle, allowing for complete exposure of the heated pipe to the injected fuel. This arrangement optimizes the interaction between the gaseous fuel and the heated surface, enhancing the efficiency of the vaporization process. The overall design of the vaporizer, along with the controlled heating and injector placement, contributes to producing a well-vaporized and homogeneously mixed fuel-air charge for combustion in the engine cylinder.

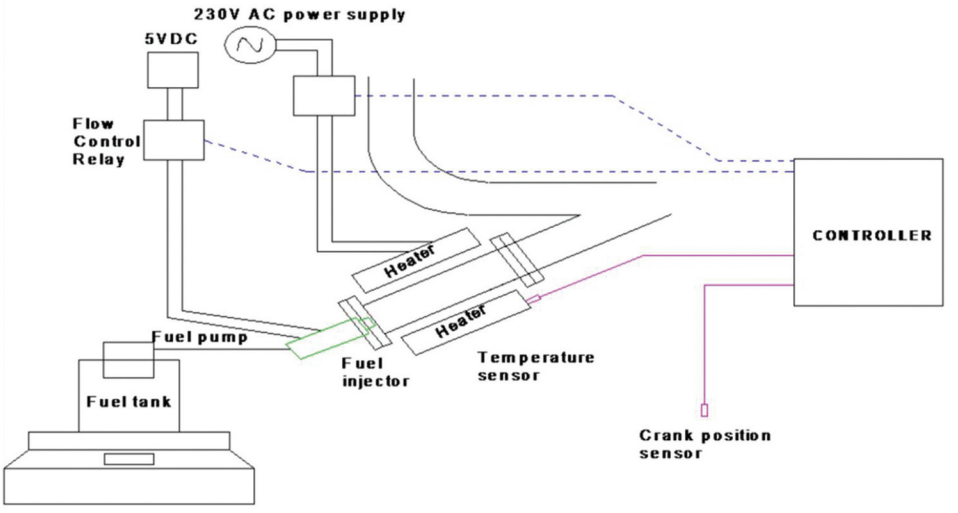



**Figure 2.** Control circuit for heater.

### *Injector cut-off setup*

Experimental investigation introduces an optical sensor-controlled fuel injection system for HCCI engines. A pulley with a window, rotating at engine speed, triggers the injector when an optical ray passes through, precisely controlling fuel injection timing. The system's flexibility is enhanced by adjusting the injection angle through pulley window rotation, providing optimized combustion near the Top Dead Centre. Experimental results demonstrate the system's efficacy in achieving precise injection timing, influencing combustion efficiency and heat release rates. This innovation offers a promising avenue for enhancing HCCI engine performance and contributing to cleaner, more efficient combustion technologies. Control circuit sketch of the fuel injector is shown in Figure 3.

### *Speed measurement mechanism*

Speed measurement is accomplished in this setup using the pulley employed in the injector cutoff mechanism. Magnets on the pulley's periphery generate a magnetic field detected by a proximate sensor affixed to a C-shaped acrylic clamp. The sensor senses changes in the magnetic field as the pulley rotates, enabling accurate speed measurement. The sensed data from this mechanism is then utilized to display the engine speed. This dual functionality of the pulley illustrates an efficient integration, simplifying the experimental setup while providing essential data for speed measurement and injector control in HCCI engines.

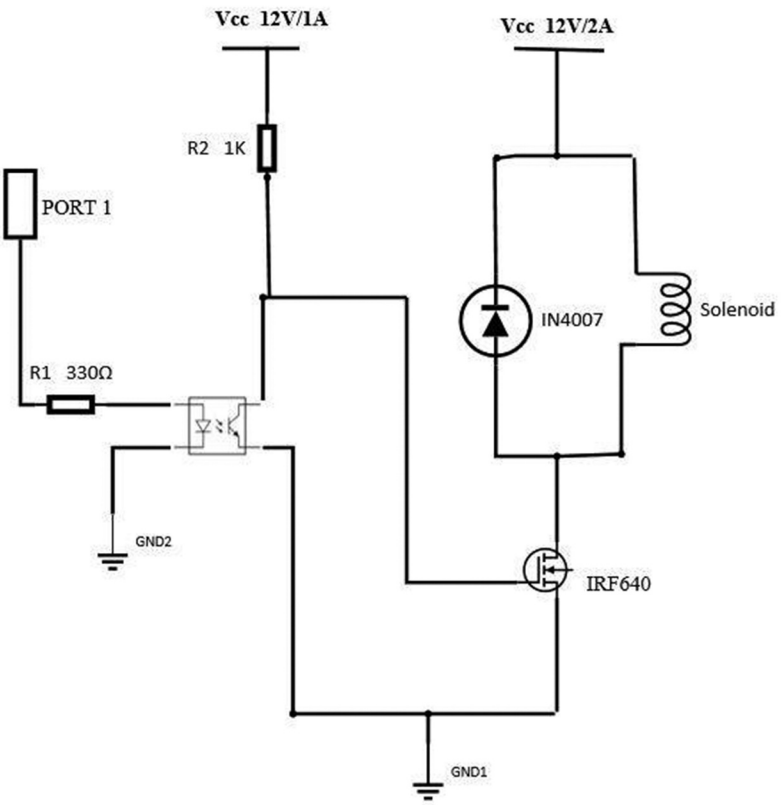
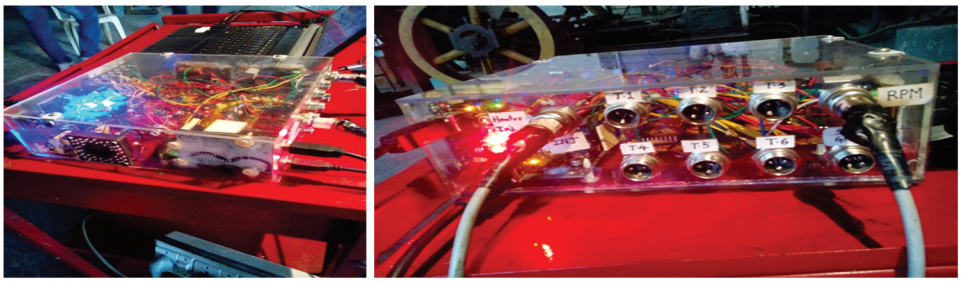



**Figure 3.** Control circuit for fuel injector.




(a) (b)

**Figure 4.** Data acquisition system.

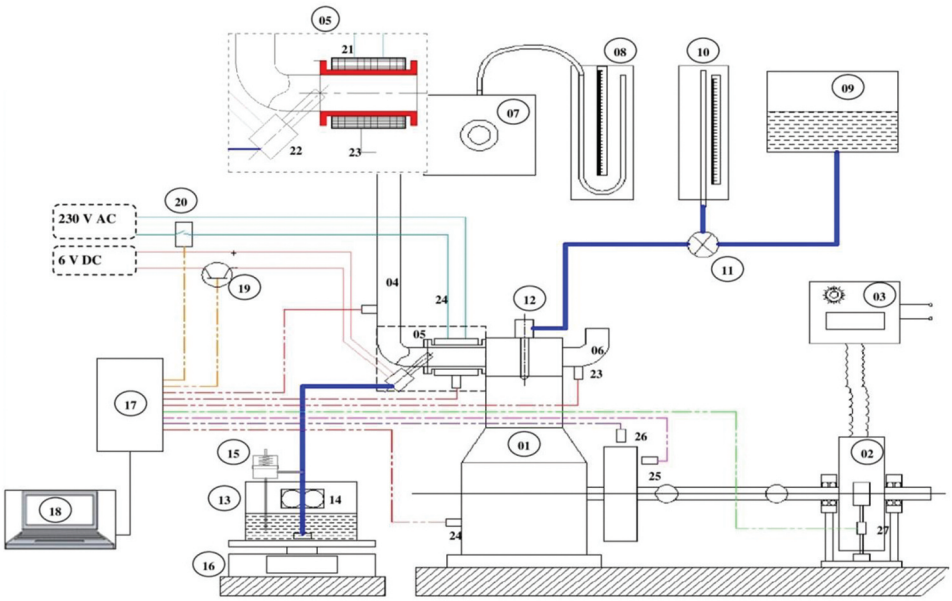



**Figure 5.** Schematic diagram of experimental setup.

(1) Diesel Engine
(2) Eddy Current Dynamometer
(3) Excitation control Rheostat
(4) Air Inlet
(5) Fuel Vaporizer
(6) Exhaust Gas Outlet
(7) Air Surge Tank with orifice meter
(8) Manometer
(9) Diesel Fuel Tank
(10) Burette
(11) Three Way valve
(12) Conventional Diesel Injector
(13) Fuel Tank for fuel vaporizer
(14) Fuel Pump
(15) Pressure Regulating Valve
(16) Weighting Scale
(17) High-speed Data Acquisition System with injector control module
(18) Computer
(19) MOSFET for Injector control
(20) Relay for heater supply control
(21) Heater Tube
(22) Fuel Injector
(23) Thermocouple
(24) Temp. Sensor
LM-35
(25) High Precision Crank Position sensor
(26) Tachometer
(27) Load Cell

### *Data acquisition unit*

The Data Acquisition Unit (DAU) in this experimental setup serves as a hub for connecting various sensors and incorporates four LED lights to provide immediate visual. DAU is shown in Figure 4. The following components are numbered in the setup that activate when the heater is ON, indicating the heating element's operation. The second LED lights up to signal the initiation of the injector trigger, providing a clear indication of the fuel injection process. The third and fourth LEDs are crucial diagnostic indicators. The third LED alerts users to a potential failure in the thermocouples connected to the exhaust gas. In contrast, the fourth LED indicates a failure in the thermocouples associated with the vaporizer. These LED indicators offer a quick and efficient way to monitor the system's operational status and identify any temperature-related issues during the experimental process. Figure 5 is the sketch diagram of the experimental set up.

## Results and discussion

### *Effect on engine performance parameters*

#### *Specific energy consumption*

Experimental results revealed that increasing the HCCI percentage leads to a corresponding rise in specific energy consumption, especially at higher loads. This section delves into the data and implications of this trend, particularly in comparison to diesel mode, as shown in Figure 6. In the 10% HCCI, 30% HCCI, 50% HCCI, 70% HCCI, and 90% HCCI modes, the increase in specific energy consumption at higher loads was 2.81%, 5.2%, 5.3%, 5.4%, and 5.9%, respectively. These findings align with those observed by Gharehghani et al. (2019), who also noted that higher HCCI proportions lead to increased energy consumption due to the higher demands for achieving the compression conditions necessary for auto-ignition.

This increase in energy consumption is primarily due to the greater compression work required at elevated loads, a well-documented phenomenon in HCCI engines. Gharehghani et al. (2019) further highlighted the need for advanced control strategies to optimize the combustion process, a challenge that was also observed in our study. Specifically, maintaining combustion stability and control at higher loads remains a significant hurdle, and addressing these challenges is critical to improving energy efficiency in HCCI engines.

In addition to our findings, similar trends have been reported by Purayil et al. (2021), who emphasized the increased work required to maintain optimal combustion conditions as HCCI operation is increased. Collectively, these studies underscore the necessity for further research and technological advancements in optimizing HCCI engine performance. Balancing the emission

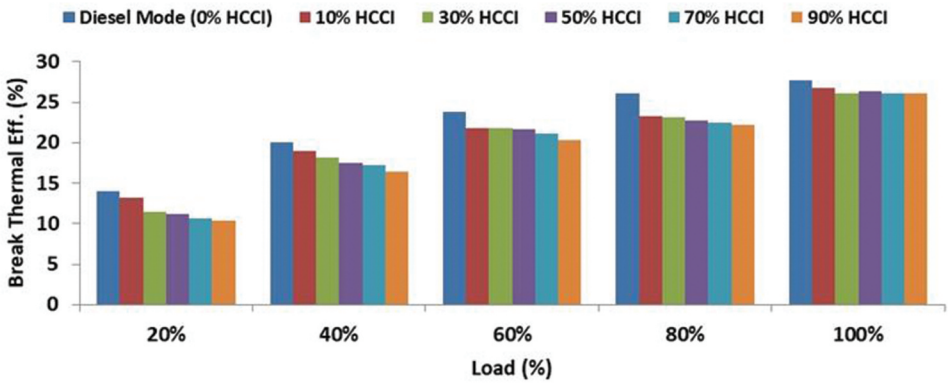



Figure 6. Effect of HCCI on specific energy consumption.

reduction benefits with the challenge of increased specific energy consumption, particularly at higher loads, remains a key focus for future applications of HCCI technology.

### *Brake thermal efficiency*

The experimental results indicate that brake thermal efficiency (BTE) decreases with an increase in HCCI percentage at higher loads, revealing the complex behavior of the engine under varying operating conditions. As shown in Figure 7, BTE in the 10% HCCI, 30% HCCI, 50% HCCI, 70% HCCI, and 90% HCCI modes decreased by 2.84%, 4.8%, 5.21%, 5.39%, and 5.90%, respectively, compared to diesel. This trend deviates from conventional engines, where BTE typically increases with load due to improved combustion efficiency and reduced heat losses.

The reduction in BTE with higher HCCI percentages can be attributed to challenges in maintaining combustion stability, especially at elevated loads. Factors such as combustion phasing control, increased heat losses, and difficulties in achieving complete combustion become more pronounced in HCCI engines (Telli et al. 2021). These findings are consistent with those of Gharehghani et al. (2019), who observed a similar decrease in BTE with increased HCCI percentages. Their study also pointed to the challenges of efficient energy conversion and stable combustion at higher loads.

Furthermore, Verma et al. (2021) emphasized the importance of precise control over ignition timing and air-fuel mixture preparation to mitigate heat losses and enhance combustion completeness. Our results reinforce the necessity of advanced control strategies and optimization techniques to improve brake thermal efficiency and engine performance, particularly under high-load conditions with elevated HCCI percentages.

It can be concluded that, while HCCI technology holds promise for reducing emissions, achieving optimal brake thermal efficiency under varying loads and HCCI percentages requires addressing significant technical challenges. Advanced optimization techniques are essential to fully harness the potential of HCCI combustion, ensuring both efficiency and stability across a wide range of operating conditions.

### *Volumetric efficiency*

Experimental results revealed that volumetric efficiency decreases as the load increases with higher HCCI blending. Specifically, in the 10% HCCI, 30% HCCI, 50% HCCI, 70% HCCI, and 90% HCCI modes, the volumetric efficiency decreased by 7.34%, 8.97%, 9.32%, 10.59%, and 12.89%, respectively, as shown in Figure 8. These results indicate a significant reduction in volumetric efficiency compared to diesel mode, which is consistent with the challenges posed by HCCI combustion.

In conventional internal combustion engines, volumetric efficiency typically decreases at higher loads due to increased pumping losses and reduced intake time. However, the introduction of higher HCCI percentages amplifies this effect, as the intake air is partially replaced by vaporized fuel,

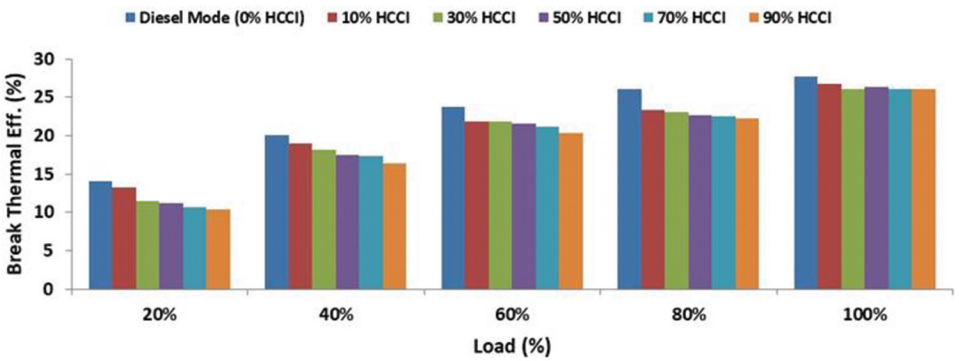



Figure 7. Effect of HCCI on break thermal efficiency.

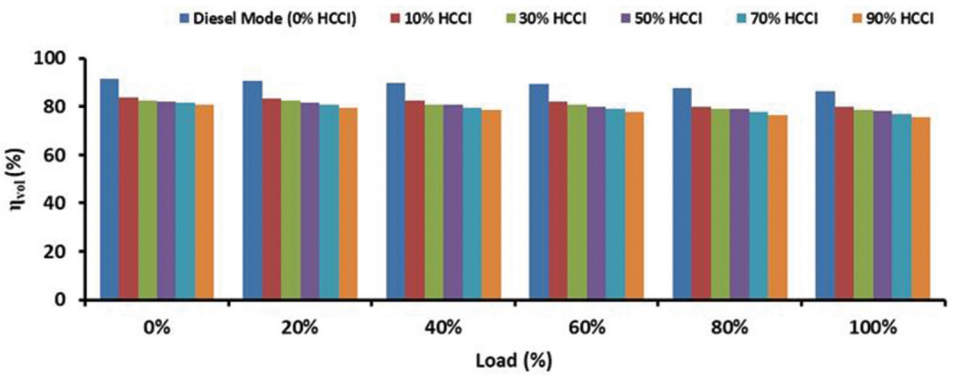



**Figure 8.** Effect of HCCI on volumetric efficiency.

impacting the air-fuel mixture dynamics (Telli et al. 2021). This reduction in fresh air intake is a trade-off that complicates the optimization of HCCI engines, where balancing combustion stability, emissions, and efficiency is essential across varying operating conditions.

Similar trends have been reported by Verma et al. (2021), who observed a decrease in volumetric efficiency with increased HCCI percentages due to the displacement of intake air by vaporized fuel. Gao et al. (2020) also emphasized the critical need to optimize the air-fuel mixture to balance the trade-offs between combustion stability, emissions, and efficiency in HCCI engines. These findings reinforce the importance of advanced engine management strategies to optimize volumetric efficiency while maintaining the benefits of HCCI technology, particularly under high-load conditions with varying HCCI percentages.

## *Effect on engine exhaust parameters*

### *Exhaust gas temperature*

Experimental results demonstrated that exhaust gas temperature (EGT) decreases as the HCCI percentage increases at higher loads. Specifically, in the 10% HCCI, 30% HCCI, 50% HCCI, 70% HCCI, and 90% HCCI modes, the EGT was observed to be 2.89%, 4.92%, 4.9%, 7.24%, and 10.14% lower, respectively, compared to diesel-run compression ignition (CI) engines, as shown in Figure 9. This trend highlights the distinct combustion characteristics of HCCI engines compared to conventional diesel engines.

In conventional diesel engines, as load increases, combustion generates higher flame temperatures, leading to elevated exhaust gas temperatures. However, HCCI engines, which rely on a homogeneous air-fuel mixture that auto-ignites under compression, exhibit more controlled combustion. This leads to lower peak flame temperatures and a broader spatial distribution of the combustion process (Chaichan and Abaas 2015; Hoang 2021; Mofijur et al. 2016). As a result, the increase in EGT with load is more moderate in HCCI engines than in conventional diesel engines.

Chaichan et al. (2015) observed a similar trend, reporting that HCCI combustion results in lower exhaust gas temperatures due to the more controlled and distributed nature of the combustion process. Their findings align with our results, as HCCI combustion reduces peak temperatures and heat release rates. Zhang et al. (2020) also supported this observation, noting that the unique characteristics of HCCI combustion contribute to improved thermal efficiency and lower emissions by reducing exhaust gas temperatures.

These findings underscore the potential of HCCI technology to achieve lower exhaust gas temperatures, which enhances engine performance and provides environmental benefits. The ability of HCCI engines to maintain lower temperatures under varying operating conditions, including

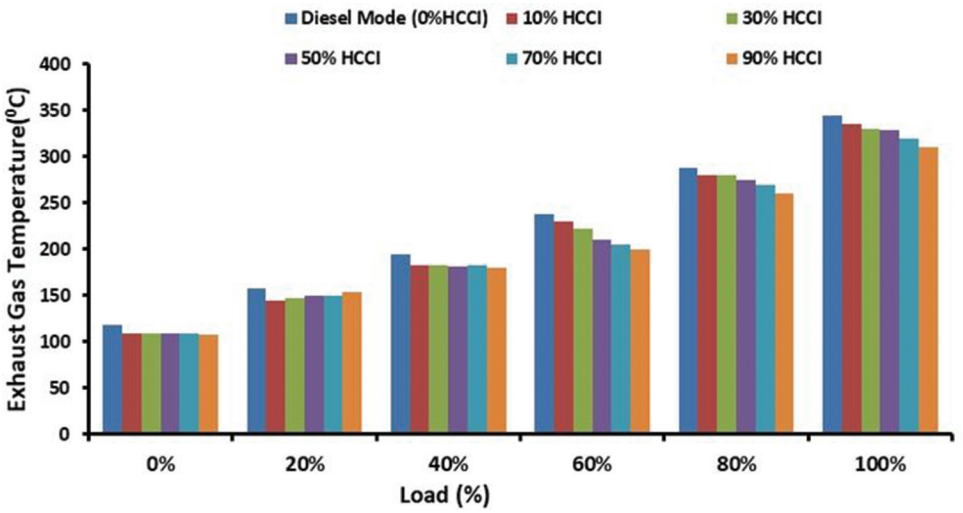



**Figure 9.** Effect of HCCI on exhaust gas temperature.

increased loads, makes it a promising technology for future applications in achieving a balance between efficiency and emissions.

### *$NO_X$ emissions*

The experimental data revealed a significant reduction in nitrogen oxide (NOx) emissions as the HCCI percentage increased at higher engine loads. Specifically, in the 10% HCCI, 30% HCCI, 50% HCCI, 70% HCCI, and 90% HCCI modes, NOx emissions were reduced by 1.5%, 3.48%, 6.72%, 8.13%, and 16.3%, respectively, compared to diesel-run compression ignition (CI) engines, as shown in Figure 10. This reduction can be attributed to the lower combustion temperatures in HCCI engines.

In conventional diesel engines, higher loads result in increased combustion temperatures and pressures, which promote the formation of NOx compounds. However, the combustion process in HCCI engines is characterized by a more uniform air-fuel mixture that auto-ignites under compression, leading to more controlled combustion with lower peak temperatures (Bhurat et al. 2021). The reduced oxygen availability and lower temperatures in HCCI engines significantly limit the formation of NOx emissions compared to diesel engines.

These findings are consistent with those reported by Verma et al. (2021), who observed a similar trend of reduced NOx emissions as the HCCI percentage increased. Their study emphasized that the lower peak combustion temperatures inherent to HCCI technology play a crucial role in minimizing NOx formation. Additionally, Kumar et al. confirmed that the controlled and homogeneous combustion process in HCCI engines is a key factor in reducing NOx emissions.

In summary, the ability of HCCI engines to significantly reduce NOx emissions while maintaining efficient combustion makes them a promising technology for environmentally friendly and fuel-efficient internal combustion engines. These results highlight the potential of HCCI technology in contributing to cleaner engine operation and advancing emissions control strategies under varying operating conditions.

### *$CO_2$ emission*

Our experimental results showed that carbon dioxide (CO2) emissions increased with engine load as HCCI percentages rose. Specifically, in the 10% HCCI, 30% HCCI, 50% HCCI, 70% HCCI, and 90% HCCI modes, CO2 emissions were found to be 4.26%, 5.32%, 6.79%, 11.45%, and 17.57% higher, respectively, compared to diesel, as shown in Figure 11. This trend can be attributed to the increased fuel consumption required to meet higher power demands at elevated loads.

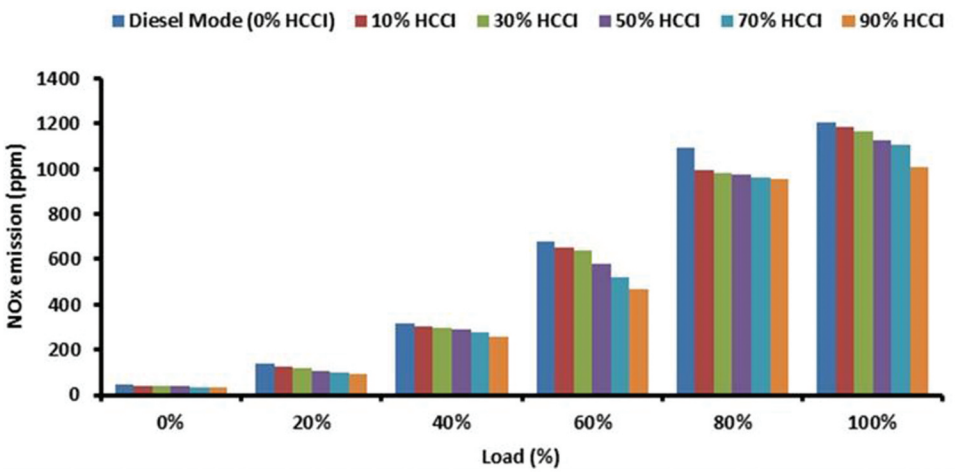



**Figure 10.** Effect of HCCI on $NO_x$ emission.

In conventional internal combustion engines, an increase in load leads to higher combustion temperatures and pressures, which results in more complete combustion of fuel and a corresponding increase in CO2 emissions. However, in HCCI (Homogeneous Charge Compression Ignition) engines with Exhaust Gas Recirculation (EGR), the combustion process operates at lower peak temperatures and pressures, and the homogeneous air-fuel mixture tends to promote more efficient combustion. The use of EGR, which re-circulates exhaust gases back into the combustion chamber, further moderates combustion temperatures by diluting the intake air mixture (Han et al. 2016).

Despite the efficiency benefits of HCCI and EGR, the overall CO2 emissions can still increase at higher loads due to the need for additional fuel to maintain engine performance. This finding aligns with the results of Han et al. (2016), who also observed higher CO2 emissions with increasing HCCI percentages. Their research indicated that while HCCI can improve combustion efficiency, the increase in fuel consumption at higher loads contributes to the rise in CO2 emissions. Similarly, Liu et al. reported that although HCCI engines reduce some emissions, CO2 emissions tend to increase with higher HCCI percentages and loads.

These results underscore the trade-offs involved in optimizing HCCI engines for both performance and emissions control. While HCCI technology and EGR offer advantages in controlling certain

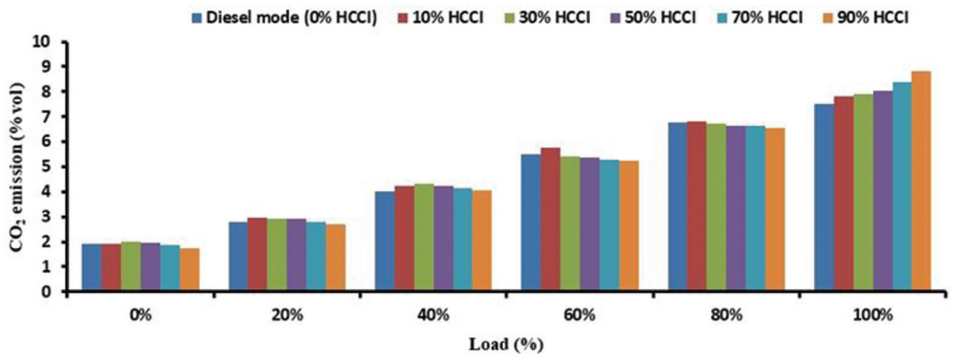



**Figure 11.** Effect of HCCI on $CO_2$ emission.

pollutants, achieving an optimal balance between CO2 emissions and engine performance requires careful consideration of engine design and operating conditions.

### HC emission

The experimental results indicated that hydrocarbon (HC) emissions increased as the HCCI percentage and engine load rose. Specifically, in the 10% HCCI, 30% HCCI, 50% HCCI, 70% HCCI, and 90% HCCI modes, HC emissions were observed to be 28.78%, 40.9%, 46.96%, 65.15%, and 98.48% higher, respectively, compared to diesel, as shown in Figure 12. This increase can be attributed to the lower peak temperatures in HCCI combustion, which, under certain conditions, may result in incomplete combustion and higher unburned hydrocarbon emissions.

HCCI (Homogeneous Charge Compression Ignition) engines are known for their lower combustion temperatures, which, while beneficial for reducing NOx emissions, can lead to higher HC emissions if complete combustion is not achieved. The challenge with HCCI engines lies in finding the right balance between combustion efficiency and emissions control across a wide range of operating conditions (Ramana, Maheswar, and UmameheswerGowd 2015).

As correctly noted, the use of catalytic converters is a well-established method for reducing HC emissions. Catalytic converters facilitate additional oxidation reactions in the post-combustion process, converting unburned hydrocarbons into less harmful compounds. This technology is essential for meeting stringent emission standards and improving the environmental performance of both conventional and advanced combustion engines.

Kumar Maurya and Kumar Agarwal (2015) reported similar findings, showing that higher HCCI percentages tend to increase HC emissions due to the lower peak combustion temperatures. Their study emphasized the importance of controlling combustion conditions to minimize unburned hydrocarbons. Similarly, Gharehghani et al. (2019) demonstrated that although HCCI technology offers several emission benefits, the risk of higher HC emissions exists under certain operating conditions. These studies highlight the crucial role of catalytic converters in mitigating hydrocarbon emissions and optimizing the environmental impact of HCCI engines.

### Oxygen emission

In internal combustion engines, oxygen (O2) emission is not typically measured as a pollutant, as oxygen is not a by-product of combustion. Instead, O2 concentration in exhaust gases is used to assess combustion efficiency. In conventional engines, variations in the air-fuel ratio and combustion completeness influence the oxygen content in the exhaust. A more complete combustion process leaves less unburned oxygen in the exhaust gases.

In the context of HCCI (Homogeneous Charge Compression Ignition) engines, where combustion occurs differently from conventional engines, changes in O2 emissions tend to be minimal or

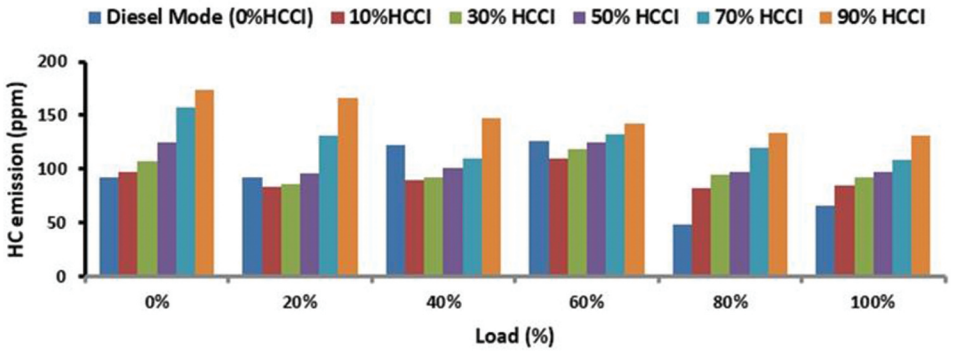



Figure 12. Effect of HCCI on HC emission.

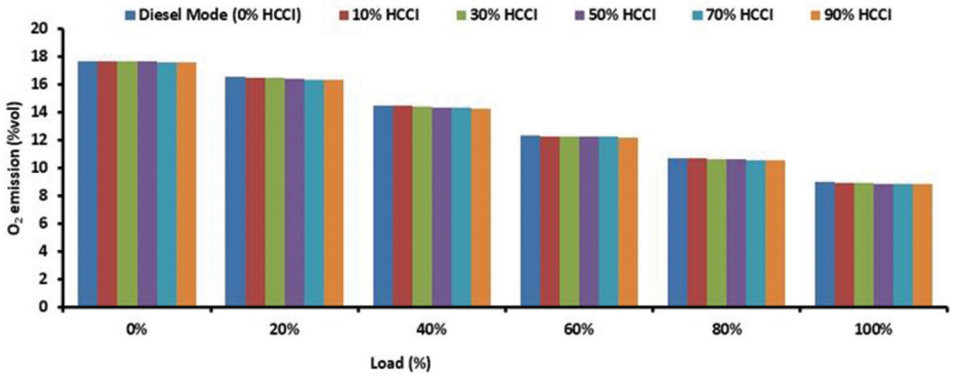



**Figure 13.** Effect of HCCI on $O_2$ Emission.

negligible. HCCI combustion, characterized by a more homogeneous air-fuel mixture and lower peak temperatures, affects the combustion chemistry and the amount of oxygen available for reactions. While O2 emissions are not a direct regulatory focus, monitoring changes in O2 concentration can provide critical insights into combustion efficiency and overall engine performance. Figure 13 illustrates the oxygen emissions observed during the experiment conducted in HCCI mode, highlighting the consistency of O2 concentration across different operating conditions.

Wang et al. (2019) observed that O2 emissions in HCCI engines showed minimal variations compared to conventional engines. Their research indicated that the stable O2 concentrations in the exhaust were due to the homogeneous air-fuel mixture and lower combustion temperatures typical of HCCI technology. Similarly, Lee et al. (2019) supported these findings, highlighting that while O2 emissions are not directly regulated, monitoring O2 concentrations is essential for evaluating combustion efficiency and optimizing the combustion process in HCCI engines.

These studies emphasize the importance of understanding O2 emission patterns to improve combustion efficiency and engine performance, even though O2 emissions themselves are not a primary concern in emission regulations.

## Conclusions

The above experimental investigation reveals that Homogeneous Charge Compression Ignition (HCCI) holds promise for the future of Compression Ignition (CI) engines. During the investigation, the following outcomes were identified:

The specific energy consumption increased with engine load compared to diesel. A decreasing trend in brake thermal efficiency was observed at higher loads. Volumetric efficiency showed a declining trend with increased engine load, and Exhaust Gas Temperature (EGT) also decreased as the engine load increased. Notably, $NO_X$ emissions decreased by approximately 16% compared to neat diesel under increased engine load. However, $CO_2$ emissions increased by up to 17% at higher loads than diesel. Hydrocarbon (HC) emissions also rose significantly, increasing by approximately 95% as engine load was enhanced.

While these findings suggest that HCCI can improve specific performance metrics and emissions characteristics of CI engines, further research is necessary to optimize its combustion phases and thoroughly analyze its effects under various operating conditions. This research provides a crucial step toward understanding the potential of HCCI technology, indicating its advantages and the areas that require further investigation to harness its benefits fully.

In conclusion, HCCI technology shows a favorable future for CI engines, particularly in reducing $NO_X$ emissions and optimizing specific energy consumption. Continued exploration and refinement

are essential to address the observed challenges and to enhance the overall efficiency and emissions profile of HCCI engines.

The following outcomes have been identified during the investigation:

- The specific energy consumption increased in engine load compared to diesel.
- A decreasing trend was observed for brake thermal efficiency at higher loads.
- The decreasing trend was observed during the evaluation of volumetric efficiency. The enhancement in the engine load was reduced.
- EGT was observed to have a reducing trend upon increasing the engine load.
- It was a favorable observation that $NO_X$ decreased when the engine load increased. Approximately 16% $NO_X$ lower than neat diesel.
- The $CO_2$ emission was found to increase by up to 17% at higher load, which was achieved compared to diesel.
- The HC emission was found to increase when the engine load was. It increased by approximately 95% as compared to diesel.

For moderate operating conditions, it was adopted, but more studies were needed for its combustion phases and analysis.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Notes on contributor

*Naveen Kumar*, Centre for Advanced Studies and Research in Automotive Engineering, Department of Mechanical Engineering, Delhi Technology University, Delhi Pawan Kishore Jha, Research Scholar, Centre for Advanced Studies and Research in Automotive Engineering, Department of Mechanical Engineering, Delhi Technology University, Delhi R. C. Singh, Professor, Centre for Advanced Studies and Research in Automotive Engineering, Department of Mechanical Engineering, Delhi Technology University, Delhi

## ORCID

Naveen Kumar http://orcid.org/0000-0002-4408-9997

## Abbreviations

| | |
|---|---|
| BDC | Bottom Dead Center |
| BP | Brake Power |
| BSFC | Brake-Specific Fuel Consumption |
| BTE | Brake Thermal Efficiency |
| CA | Crank Angle |
| CI | Compression-Ignition |
| CO | Carbon Monoxide |
| $CO_2$ | Carbon Dioxide |
| CR | Compression Ratio |
| DI | Direct Injection |
| EGR | Exhaust Gas Recirculation |
| ER | Equivalence Ratio |
| HC | Hydro Carbon |
| HCCI | Homogeneous Charge Compression Ignition |
| IA | Ignition Advance |
| ICE | Internal Combustion Engine |
| $NO_X$ | Nitrogen Oxides |

| | |
|---|---|
| PM | Particulate Matter |
| SI | Spark Ignition |
| TDC | Top Dead Center |

## References


Bhurat, S. S., S. Pandey, V. Chintala, and P. S. Ranjit. 2021. Technical barriers and their solutions for deployment of HCCI engine technologies–a review. *International Journal of Ambient Energy* 42 (16):1922–35. doi: 10.1080/01430750.2019.1611644.

Biswas, S., A. Sengupta, D. Kakati, and R. Banerjee. 2023. The transition from conventional biodiesel combustion to RCCI with CNG/Ethanol induction in CI engine: A comparative combustion analysis and relative effects on performance-emissions. *International Journal of Engine Research* 24 (6):2505–22. doi: 10.1177/14680874221123615.

Chaichan, M. T., and K. I. Abaas. 2015. EGR and Injection timing variation effects of an engine run in HCCI mode performance and emitted emissions. *International Journal of Engineering Trends and Technology* 19 (3):121–30. doi: 10.14445/22315381/IJETT-V19P221.

Chen, L., J. Liu, Z. Wang, and M. Yao. 2019. Influence of intake air distribution on HCCI combustion and emissions. *Applied Energy* 239:218–26. doi: 10.1016/j.apenergy.2019.01.211.

Cheng, X., H. Zhao, and Z. Wang. 2017. Advanced exhaust after-treatment technologies for hydrocarbon emission reduction in HCCI engines. *Applied Energy* 204:945–56.

Christensen, M., A. Hultqvist, and B. Johansson. 2016. Demonstrating the multi fuel capability of a homogeneous charge compression ignition engine with variable compression ratio. *SAE Transactions* 105:407–22.

Duan, X., M. C. Lai, M. Jansons, G. Guo, and J. Liu. 2020. A review of controlling strategies of the ignition timing and combustion phase in Homogeneous Charge Compression Ignition (HCCI) engine. *Energy Reports* 6:433–51.

Fathi, M., O. Jahanian, and M. Shahbakhti. 2021. Modeling and Controller design architecture for cycle-by-cycle combustion control of homogeneous charge compression ignition (HCCI) engines – a comprehensive review. *Energy Reports* 7:892–908.

Gao, Z., Y. Zhang, and J. Wang. 2020. Optimization of air-fuel mixture to reduce hydrocarbon emissions in HCCI engines. *Fuel* 278:118337.

Gharehghani, A. 2019. Load limits of an HCCI engine fueled with natural gas, ethanol, and methanol. *Fuel* 239 (April 2018):1001–14. doi: 10.1016/j.fuel.2018.11.066.

Han, H. S., K. R. Han, Y. Wang, C. J. Kim, C. H. Sohn, and C. Nam. 2016. Effects of natural-gas blending on ignition delay and pollutant emission of diesel fuel for the condition of homogeneous charge compression ignition engine. *Energy (Oxford)* 103:1–10.

Hoang, A. T. 2021. Combustion behavior, performance and emission characteristics of diesel engine fuelled with biodiesel containing cerium oxide nanoparticles: A review. *Fuel Processing Technology* 218 (March):106840. doi: 10.1016/j.fuproc.2021.106840.

Kokjohn, S. L., R. M. Hanson, D. A. Splitter, and R. D. Reitz. 2013. Fuel reactivity controlled compression ignition (RCCI): A pathway to controlled high-efficiency clean combustion. *International Journal of Engine Research* 14 (3):210–26.

Kumar Maurya, R., and A. Kumar Agarwal. 2015. Combustion and emission characterization of N-Butanol fueled HCCI engine. *Journal of Energy Resources Technology* 137 (1). doi: 10.1115/1.4027898.

Lee, S., D. Kim, and S. Park. 2019. Enhancing volumetric efficiency through combustion chamber design improvements in HCCI engines. *Fuel* 236:145–53.

Li, Y., H. Liu, and M. Yao. 2018. Advanced intake air preheating techniques for improved air-fuel mixture homogeneity in HCCI engines. *Energy Conversion and Management* 166:545–55. doi: 10.1016/j.enconman.2018.04.040.

Liu, H., S. Ma, Y. Li, and M. Yao. 2018. Combustion characteristics and NOx emissions of a HCCI engine fueled with n-Butanol–Diesel blends under various air-fuel ratios. *Fuel* 235:778–89.

Mishra, D., and R. C. Naveen Kumar. 2017. A comprehensive review on performance, combustion and emission characteristics of biodiesel fuelled diesel engines. *Renewable and Sustainable Energy Reviews* 79:1134–59. doi: 10.1016/j.rser.2017.05.176.

Mofijur, M., M. G. Rasul, J. Hyde, A. K. Azad, R. Mamat, and M. M. K. Bhuiya. 2016. Role of biofuel and their binary (diesel–biodiesel) and ternary (ethanol–biodiesel–diesel) blends on internal combustion engines emission reduction. *Renewable and Sustainable Energy Reviews* 53:265–78. doi: 10.1016/j.rser.2015.08.046.

Mogra, A., and K. K. Gupta. 2020. A review on new technology in internal combustion engine-HCCI engine. IOP Conference Series: Materials Science and Engineering Mumbai India, 12054, IOP Publishing.

Niklawy, W., M. Shahin, M. I. Amin, and A. Elmaihy. 2022. Comprehensive analysis of combustion phasing of multi-injection HCCI diesel engine at different speeds and loads. *Fuel* 314:116731. doi: 10.1016/j.fuel.2021.123083.

Nishi, M., M. Kanehara, and N. Iida. 2016. Assessment for innovative combustion on HCCI engine by controlling EGR ratio and engine speed. *Applied Thermal Engineering* 99:42–60. doi: 10.1016/j.applthermaleng.2015.11.126.

Onishi, K., K. Nakamura, K. Moriya, and M. Mori. 2018. Advanced combustion control strategies for high efficiency and low emission in HCCI engines. *International Journal of Engine Research* 19 (6):701–12.
Pandey, S., P. Diwan, P. K. Sahoo, and S. S. Thipse. 2018. A review of combustion control strategies in diesel HCCI engines. *Biofuels* 9 (1):61–74. doi: 10.1080/17597269.2016.1257315.
Polat, S. 2016. An experimental study on combustion, engine performance and exhaust emissions in a HCCI engine fuelled with diethyl ether–ethanol fuel blends. *Fuel Processing Technology* 143:140–50. doi: 10.1016/j.fuproc.2015.11.021.
Purayil, S. T. P., M. O. Hamdan, S. A. B. Al-Omari, M. Y. E. Selim, and E. Elnajjar. 2021. Review of hydrogen–gasoline SI dual fuel engines: Engine performance and emission. *Energy Reports* 7:1239–54.
Ramana, P. V., D. Maheswar, and B. UmameheswerGowd. 2015. Review on research and development of HCCI technology. *International Journal of IT, Engineering and Applied Sciences Research* 4 (5):5–16.
Reitz, R. D., H. Ogawa, R. Payri, T. D. Fansler, S. L. Kokjohn, Y. Moriyoshi, and M. Moulai. 2020. IJER editorial: The future of the internal combustion engine. *International Journal of Engine Research* 21 (1):3–10. doi: 10.1177/1468087419877990.
Riyadi, T. W. B., M. Spraggon, S. G. Herawan, M. Idris, P. A. Paristiawan, N. R. Putra, R. Veza, and I. Silambarasan. 2023. Biodiesel for HCCI engine: Prospects and challenges of sustainability biodiesel for energy transition. *Results in Engineering* 17:100916. doi: 10.1016/j.rineng.2023.100916.
Saxena, S., and I. D. Bedoya. 2013. Fundamental processes affecting the HCCI combustion and the factors influencing its control. *Renewable and Sustainable Energy Reviews* 24:118–27.
Sharma, T. K., G. A. P. Rao, and K. M. Murthy. 2016. Homogeneous charge compression ignition (HCCI) engines: A review. *Archives of Computational Methods in Engineering* 23 (4):623–57. doi: 10.1007/s11831-015-9153-0.
Sjoberg, M., and J. E. Dec. 2005. Effects of EGR and Its constituents on HCCI autoignition of ethanol. *Proceedings of the Combustion Institute* 30:2719–26.
Tan, Y., Y. Zhang, Y. Li, and M. Yao. 2018. Advanced combustion control strategies in HCCI engines: Closed-loop combustion control and predictive combustion control. *Energy (Oxford)* 157:845–57.
Telli, G. D., C. R. Altafini, C. A. Costa, J. S. Rosa, M. E. Martins, and L. A. O. Rocha. 2021. A comprehensive review of homogeneous charge compression ignition (HCCI) engines: Advantages, challenges and evolution 2020 SAE Brasil Congress & Exhibition.
Verma, S. K., S. Gaur, T. Akram, S. Gautam, and A. Kumar. 2022. Emissions from Homogeneous Charge Compression Ignition (HCCI) engine using different fuels: A review. *Environmental Science and Pollution Research* 29 (34):50960–69. doi: 10.1007/s11356-021-15602-x.
Verma, S. K., S. Gaur, T. Akram, Samsher, and A. Kumar. 2021. Performance characteristic of HCCI engine for different fuels. *Materials Today: Proceedings* 47 (xxxx):6030–34. doi: 10.1016/j.matpr.2021.04.609.
Vuilleumier, D., I. Taritas, B. Wolk, D. Kozarac, S. Saxena, and R. W. Dibble. 2016. Multi-level computational exploration of advanced combustion engine operating strategies. *Applied Energy* 184:1273–83. doi: 10.1016/j.apenergy.2016.05.043.
Wang, Z., X. Han, X. Zhang, and Z. Huang. 2019. Thermal management strategies for improving energy efficiency of HCCI engines. *Energy Conversion and Management* 196:1068–77.
Wang, Z., X. Zhang, X. Han, and Z. Huang. 2021. Optimization of fuel injection strategies for improved air-fuel mixture homogeneity in HCCI engines. *Fuel* 288:119570. doi: 10.1016/j.fuel.2020.119570.
Xie, H., H. Xu, and K. Deng. 2018. Optimization of intake and exhaust systems for improving volumetric efficiency in HCCI engines. *Energy (Oxford)* 151:55–64.
Yamagiwa, R., Y. Manabe, S. Mito, A. Iijima, S. Yoshihara, T. Yamaguchi, and S. Miyamoto. 2023. A study of autoignition and combustion characteristics in an HCCI engine using a blended fuel of DME and City gas. 2023-32-0017. JSAE/SAE Powertrains, Energy and Lubricants International Meeting Kyoto, Japan.
Yang, B., H. Liu, and M. Yao. 2020. Impact of ignition timing on combustion and emissions of a gasoline HCCI engine. *Energy Conversion and Management* 214:112909.
Zhang, Y., Z. Zheng, H. Liu, and M. Yao. 2020. Effect of EGR on the performance of a gasoline HCCI engine with vaporized fuel injection. *Fuel* 267:117262. doi: 10.1016/j.fuel.2020.117154.
Zhang, Z., Y. Chen, and Z. Huang. 2019. Optimization of air-fuel mixture for NOx reduction in HCCI engines. *Fuel* 244:1–12. doi: 10.1016/j.fuel.2019.01.129.
Zhao, S., S. Niu, H. Yu, Y. Ning, X. Zhang, X. Li, Y. Zhang, C. Lu, and K. Han. 2019. Experimental investigation on biodiesel production through transesterification promoted by the La-dolomite catalyst. *Fuel* 257:116092. doi: 10.1016/j.fuel.2019.116092.
Zheng, Z., H. Liu, and M. Yao. 2019. Optimizing the air-fuel mixture for complete combustion in HCCI engines. *Energy (Oxford)* 174:143–54. doi: 10.1016/j.energy.2019.02.003.
Zhu, G., H. Zhao, and C. F. Lee. 2017. The role of variable valve timing in improving the performance of HCCI engines. *Applied Energy* 204:145–55.

Optical and Quantum Electronics (2025) 57:96
https://doi.org/10.1007/s11082-024-08007-w

# Highly coherent on-chip mid-infrared supercontinuum generation from 1.5 to 25 μm in CdTe-based Rib waveguide

**Deepak Garg**[1] · **Ajeet Kumar**[1]

Received: 24 September 2024 / Accepted: 19 December 2024
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2025

**Abstract**
We present design and computational analysis of a rib waveguide composed of a core of CdTe chalcogenide glass having upper and lower cladding of Zinc Selenide (ZnSe) for high coherence on-chip supercontinuum (SC) generation. CdTe chalcogenide glass, known for its wide transparency range extending from visible to mid-infrared region, is utilized in our design to achieve broad spectral coverage. Additionally, its higher non-linear refractive index enhances the supercontinuum generation efficiency. Using Finite Element Method (FEM), we conducted the dispersion engineering by analyzing the influence of waveguide geometrical parameters. This optimization is aimed at minimizing the dispersion values and to obtain a flat top dispersion profile. The proposed design offers a Zero Dispersion Wavelength (ZDW) approximately near 2.6 μm. Our simulation reports a supercontinuum spectrum broadening covering 1.5–25 μm, when pumped with 2.5 μm secant hyperbolic pulses having pulse width 50 fs and peak power 16 kW in only 1 mm length of waveguide. The high coherence of the final generated spectra positions our proposed design as a promising candidate for various non-linear applications including non-linear microscopy, optical coherence tomography, frequency comb generation and non-linear spectroscopy.

**Keywords** CdTe · Coherence · SPM · Supercontinuum generation · Non-linearity

## 1 Introduction

Over the span of years, optical waveguide based supercontinuum generation sources has gained wide popularity for various applications in the domain of optical coherence tomography (Moon and Kim 2006), non-linear microscopy (Medjouri and Djamel 2020), biomedical sensing (Kaur et al. 2022), frequency comb generation (Cheng et al. 2020a, b), pulse compression (Cheng et al. 2020a, b) etc. due to small size and enhanced bandwidth. On chip supercontinuum sources are more robust, more power and cost efficient than their counterpart Photonic Crystal Fiber (PCF) sources. The electromagnetic region spanning 1–25 μm is crucial as it covers near and mid-infrared regions which can

✉ Ajeet Kumar
ajeetdph@dtu.ac.in

[1] Advanced Photonics Simulation Research Laboratory, Department of Applied Physics, Delhi Technological University, Bawana Road, Shahbad Daulatpur, Delhi 110042, India

Published online: 06 January 2025

open door to many applications like bio-photonic diagnostics (Saini and Sinha 2021), microscopy, non-destructive testing. Mid-infrared region also includes two optical windows spanning, 3–5 µm and 8–13 µm known as atmospheric transparent windows (Ballester et al. 2020) making it suitable for sensing and free space communication. Most of the biomolecules have fundamental frequencies i.e. possesses their absorption spectra in mid-IR region making it suitable for precise identification and differentiation of molecules. Therefore, this specific region is commonly referred as molecular "fingerprint region" (Fiore and Pellerito 2021).

Supercontinuum generation (SCG) is a non-linear phenomenon of spectral broadening of an ultrashort high intensity pulse when it is made to travel through a non-linear optical medium such as optical fibers and bulk mediums. This broadening is due to the interplay of several linear and non-linear effects. When linear effects like dispersion and nonlinear effects like self-phase modulation (SPM), cross-phase modulation (XPM), modulation instability (Yupapin et al. 2022), Self-steepening (Chowdary et al. 2023), stimulated Brillouin scattering (SBS), stimulated Raman scattering (SRS), four wave mixing (FWM), and soliton dynamics at pump wavelength (Saini et al. 2020) combines to generate a broad band of spectra from an initial narrowband pulse. In past few years a substantial amount of study has been carried out on SC sources based on PCFs and waveguides. The most commonly used background material for supercontinuum generation in optical media is silica but due to the presence of multiphonon transmission edge phenomena and high absorption coefficient, the broadened SC spectra is limited to 2.5 µm only (Saini and Sinha 2021). Therefore, researchers started exploring other materials like ZBLAN (Yang et al. 2020), tellurite (Liao et al. 2009), fluoride (Gauthier et al. 2018) and chalcogenide glasses, as these materials possesses high non-linearity, wider transparency range in near and mid-IR region and low losses (Dai et al. 2018).

Among all these materials, chalcogenide (ChG) materials are best choice for SC sources as certain compositions of ChG glasses possesses optical transparency range upto 25 µm which encompasses near as well as mid-infrared region. They also possess high Kerr nonlinearity, lower material absorption and lower response time and the nonlinear refractive index of chalcogenide glasses is ten to hundred times greater than fluoride and tellurite compositions. Cadmium telluride (CdTe) as a core material is chosen because of its favorable properties which can used in an efficient way for SC sources. CdTe is an important group II–IV semiconductor having an ultra-broad transparency range of 0.9–25 µm covering the entire near and mid-IR region (Long et al. 2022). As compared to most mid-IR semiconductors and dielectrics like SiGe, $Si_3N_4$, Ge, Si etc. CdTe has the largest transparency window. It also offers high kerr non-linearity and possesses third order non-linear refractive index ($n_2$) of ~ $5.2 \times 10^{-17}$ $m^2/W$ which is two order greater than that of $Si_3N_4$ and several order greater than that of silica and is thermally stable over wide range of temperatures. Thus, CdTe is an excellent choice of material for supercontinuum generation. Few years back, in 2015, Hu et al. observed the supercontinuum phenomena in CdTe/ZnSe quantum dots (Hu et al. 2015). ZnSe as upper and lower cladding material is also an excellent choice as it also offers transparency upto 22 µm (Ming et al. 2021) and helps in the efficient mode confinement in CdTe core. Moreover, the lattice mismatch between CdTe and ZnSe is relatively small (Ray and Tarafder 2023), making them compatible for epitaxial growth. The fabrication of chalcogenide based rib waveguides can be done with good control and high precision by using the subsisting fabrication techniques like ion etching (Hiscocks et al. 2008), chemical vapour deposition (Matsuura and Harrington 1997), and photolithography (Tung et al. 2005). Further, the epitaxial layers of semiconductor group materials like

CdTe can be grown with the help of subsisting semiconductor deposition techniques like sputtering depositions techniques (Hunsperger 2009).

There have been a lot of research works done on supercontinuum sources using both waveguides and PCF geometries. During the last few years efforts have been made by the researchers for the on-chip SC generation. In 2022, Long et al. theoretically proposed a CdTe/CdS/S heterostructure waveguide demonstrating the SC broadening from 3.5 to 20 µm when a 0.25 cm long waveguide is pumped with 9 µm wavelength having peak power 15 kW (Long et al. 2022). Attempts have also been made for SC generation using other semiconductor group materials. Zhang et al. theoretically demonstrated two octave SC spectra spanning 1.150–2.425 µm using $Si_3N_4$ waveguide (Zhang et al. 2023). Saini et al. reported a dispersion engineered $As_2Se_3$ waveguide which is capable of generating SC spectra ranging from 2 to 15 µm in 4 mm long waveguide when pumped with 2.5 µm pulses (Saini et al. 2017). In 2021, Saini and Supradeepa reported the SC span from 0.88 to 5.22 µm using Tellurium-oxide coated silicon–nitride waveguide when pumped with 1.5 µm wavelength having peak power 10 kW and pule width 50 fs (Saini and Supradeepa 2021). Zhao et al. experimentally reported an SC spectrum broadening using Ge–As–Se–Te as a core material from 1.5 to 14 µm when a 23 cm long step index (SI) fiber is pumped with 4.5 µm wavelength (Zhao et al.2016). In 2018, Sinobad et al. reported supercontinuum generation from 3 to 8.5 µm using SiGe/Si waveguide pumped at 4 µm with high average power of 10 mW (Sinobad et al. 2018). Additionally, the recent advancements in the fabrication technology have made production of integrated waveguides increasingly feasible (Bres et al. 2023). Over the past years, several studies based on integrated platforms using chalcogenide materials have been demonstrated (Yu et al. 2016; Du et al. 2018), indicating that waveguides based on chalcogenide materials are well suited for integrated chip based SC generation. However, it is worth mentioning here that to achieve full and efficient integration of SC sources, it is necessary to develop several key components together. These component includes on chip fiber lasers to eliminate the external optical parametric oscillators (OPOs) or fiber based lasers as well as on chip pulse compression and amplification to achieve the required pump wave characteristics for efficient supercontinuum broadening.

In this work, we theoretically design and present a new and efficient on-chip SC source based on Cadmium Telluride (CdTe)/Zinc Selenide (ZnSe) capable of producing an ultra-broad highly coherent SC spectrum ranging from 1.5 to 25 µm which almost covers the whole transmission range of CdTe and to the best of our knowledge this has been accomplished for the first time using CdTe. Using dispersion engineering i.e. by tailoring the dimensions of waveguide parameters, we have optimized the dispersion profile of the waveguide. Supercontinuum has been obtained from the proposed waveguide structure by simulating it through generalised non-linear Schrödinger equation (GNLSE). SC spectrum ranging from 1.5 to 25 µm has been obtained using 2.5 µm hyperbolic secant pulses having pulse width 50 fs and peak power 16 kW. The highly coherent and ultra-broadened final spectrum makes our proposed design a good candidate for a diverse field of non-linear applications including environmental monitoring, optical coherence tomography and biomedical sensing.

## 2 Proposed waveguide design

The lateral cross sectional profile of the reported computational design is depicted in Fig. 1. The core of the designed waveguide is made up of CdTe and the upper and lower cladding of ZnSe. The good contrast between the refractive indices of CdTe and ZnSe

which is beneficial for proper mode confinement within the core area. The width of core region is denoted as $a$, core rib height as $b$ and core thickness as $t$, shown in Fig. 1a. The proposed waveguide structure is a rib-waveguide structure on which we have performed dispersion engineering by altering the geometrical parameters and finally proposed an optimized structure having a flat dispersion profile. The core region of the proposed design exhibits several modes and to prevent the inter-modal coupling, mode coupling suppression conditions are met by the effective indices of the mode. During the simulation and while calculating the dispersion characteristics only the effective indices of fundamental transverse electric (TE) mode is taken into account as the higher order modes have high confinement losses. The electric field profile of fundamental TE mode at 2.5 μm is shown in Fig. 1b.

## 3 Numerical analysis

### 3.1 Linear characteristics

The linear refractive index is wavelength dependent and can be calculated using an empirical relation widely known as Sellmeier equation, which in general can be written as (Sharma et al. 2020):

$$n^2(\lambda) = A_0 + \sum_{n=1}^{3} \frac{A_n \lambda^2}{\lambda^2 - B_n^2} \tag{1}$$

where, $A_0$, $A_n$, $B_n$ are fitting constants. The values of the fitting constants for our core material (CdTe) and cladding material (ZnSe) is summarized in Table 1.

We have utilized a Finite Element Method (FEM) based computational software "COMSOL Multiphysics" for modal analysis of the proposed structure. It gives the idea about the behavior of electric field and provides with the values of effective mode indices ($n_{eff}$) of the various propagating modes. In our simulations, we have used wave optics module with Electromagnetic wave, frequency domain (EWFD) study in COMSOL. To accurately model the waveguide geometry and minimize back reflections, we implemented perfectly matched layers.

Group velocity dispersion (GVD) plays a significant role in the SC process as it influences the degree to which various spectral component of an ultra-short pulse travel at

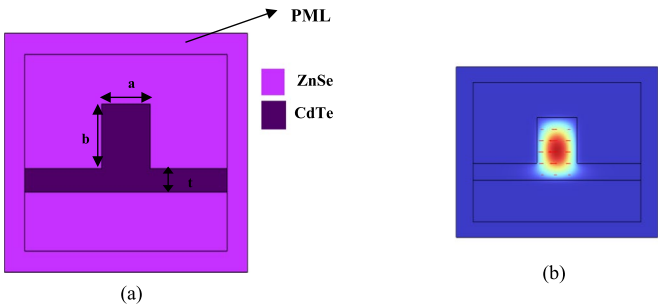



**Fig. 1** **a** Transverse cross section of proposed waveguide, **b** Fundamental mode profile at 2.5 μm

**Table 1** Sellmeier fitting coefficients for core and cladding materials

| Sellmeier coefficient | CdTe (core) | ZnSe (upper and lower cladding) |
|---|---|---|
| $A_0$ | 1 | 1 |
| $A_1$ | 6.19779 | 4.9257 |
| $A_2$ | 3.2238 | $3.2477e^{-5}$ |
| $A_3$ | – | 22.2063 |
| $B_2^2$ | 0.010106 | 0.0513 |
| $B_1^2$ | 27,873,310.3123 | 19.315 |
| $B_3^2$ | – | 16,574 |

different phase velocity within waveguide structure. The group velocity dispersion, $D$ ($\lambda$) can be calculated using the wavelength dependent refractive indices of propagating modes as (Sharma et al 2020):

$$D(\lambda) = \frac{-\lambda}{c}\frac{d^2\mathrm{Re}(n_{\mathrm{eff}})}{d\lambda^2} \tag{2}$$

Another important linear parameter is Effective mode area ($A_{\mathrm{eff}}$) of the fundamental propagating mode. $A_{\mathrm{eff}}$ can be calculated using the formulae given below (Garg et al. 2024):

$$A_{eff} = \frac{\left(\iint_{-\infty}^{\infty}|E|^2 dxdy\right)^2}{\iint_{-\infty}^{\infty}|E|^4 dxdy} \tag{3}$$

where, $E$ is the electric field distribution obtained by solving the eigenvalue problem derived from Maxwell's equations.

### 3.2 Non-linear characteristics

For SC to be broad and efficient, material non-linearity is the most significant parameter. The non-linear coefficient ($\gamma$) is related to the non-linear refractive index of the material and effective mode area of the modes propagating in waveguide as follows (Wang et al. 2024; Chauhan et al. 2019):

$$\gamma = \frac{2\pi n_2}{\lambda A_{eff}(\lambda)} \tag{4}$$

where, $n_2$ is the non-linear refractive index of the material which in our case is ~ $5.2 \times 10^{-17}$ $m^2$/W at 1.55 µm and $1.3 \times 10^{-17}$ $m^2$/W at 9 µm, $\lambda$ is the pump wavelength. For better SC broadening, non-linear coefficient should be maximized and effective mode area should be as low as possible. Non-linearity can be increased by either choosing the material having large non-linear refractive index which in our case is largest among chalcogenides and/or by altering the waveguide parameters in such a way that the effective mode area comes out to be small.

To simulate SC, a generalised nonlinear Schrödinger wave equation (GNLSE) is solved using split-step Fourier transform method which is given as:

$$\frac{\partial A}{\partial Z}+\frac{\alpha}{2}A-\left(\sum_{n\geq 2}\frac{i^{n+1}}{n!}\beta_n\frac{\partial^n A}{\partial T^n}\right)=i\gamma\left(1+\frac{i}{\omega_0}\frac{\partial}{\partial T}\right)\left(A(z,t)\int_{-\infty}^{\infty}R(T)|A(z,T-T\prime)|^2dT+i\hbar_R(z,T)\right) \tag{5}$$

The L.H.S in the above equation addresses all linear optical effects of the waveguide structure while the R.H.S pertains to non-linear effects. The dispersion coefficients are represented by $\beta_n$. The slowly varying optical wave envelope is denoted by '$A$'. $\alpha$ is the power loss in the waveguide as the optical mode propagates through it. Spatial co-ordinate across the waveguide length is denoted as $z$ with the time variable represented as $t$. The Raman term $\hbar_R$ is taken as zero since the waveguide length we are working with is very small. $R(T)$, non-linear response function, incorporating both instantaneous electronic and delayed Raman response and can be given by the following equation (Long et al. 2022):

$$R(t)=\left(1-f_R\right)\delta(t)+f_R h_R(t) \tag{6}$$

where the value of $f_R$ for CdTe is taken as $f_R=0.115$ (Long et al. 2022) The Raman response function $h_R(t)$ is expressed as follows:

$$h_R(t)=\frac{{\tau_1}^2+{\tau_2}^2}{\tau_1{\tau_2}^2}\exp\left[\frac{-t}{\tau_2}\right]\sin\left(\frac{t}{\tau_1}\right) \tag{7}$$

$\tau_1$ and $\tau_2$ in the above equation are related to non-oscillation frequency of photon and damping time of vibrating molecules. It is worth noting that the Raman parameters for CdTe are not known yet, hence we have used the values of chalcogenide glass and simulated SC for different values of Raman parameters (not shown here). There is no observable change in the SC spectra or coherence. For simulating SC we have used a tunable hyperbolic secant pulse expressed as (Khamaru and Kumar 2024):

$$A(z=0,t)=\sqrt{P_0}\,\mathrm{sech}\left(\frac{T}{T_0}\right) \tag{8}$$

where, $P_0$ represents the peak power of input tunable pulse, $T_0=T_{FWHM}/1.7627$ and $T$ is co- propagating frame at group velocity of envelope.

The optical wave breaking distance ($Z_{WB}$) is the point at which different frequency components of an optical pulse first overlap and cause significant distortion. $Z_{WB}$, can be calculated using the following formulae (Saini et al. 2018), with $\beta_2=1.4668\times10^{-27}$ s$^2$/m at pump wavelength of 2.5 µm.

$$Z_{WB}=\sqrt{\frac{3\beta_2}{2\beta_2+2\gamma P_0 T_0^2}}\frac{T_0^2}{\beta_2} \tag{9}$$

We have utilized a single-photon-per mode noise model (a semi classical theory) to explore the coherent properties of mid-infrared supercontinuum generation. This investigation employs the magnitude of first order coherence function $g_{12}$ to determine the degree of coherence, which ranges from 0 to 1 (Saini et al. 2018)

$$\left|g_{12}^{(1)}(\lambda, t_1 - t_2)\right| = \left| \frac{\left\langle E_1^*(\lambda, t_1) E_2(\lambda, t_1) \right\rangle}{\sqrt{\left\langle \left|E_1(\lambda, t_1)\right|^2 \right\rangle \left\langle \left|E_2(\lambda, t_2)\right|^2 \right\rangle}} \right| \tag{10}$$

## 4 Results and observations

The dispersion characteristics of the proposed rib waveguide structure is controlled by varying the geometric parameters like core width ($a$), rib height ($b$) and thickness of the CdTe layer ($t$). We have optimized these parameters to obtain a flat dispersion profile and low values of dispersion at pump wavelength. The effect of variation of core width ($a$) on dispersion characteristics of waveguide by keeping other parameters like rib height and thickness constant, is shown in Fig. 2. The zero dispersion wavelength (ZDW) is the wavelength where the dispersion value exhibited by the waveguide transitions from normal to anomalous dispersion regime i.e. passes zero. It can be observed that with the alteration in the value of $a$, from 6.8 to 7.2 µm, while keeping other parameters constant, there is no significant change in zero dispersion wavelength which means that the dispersion curve near ZDW almost overlaps. However, the dispersion values in anomalous region are less for $a = 7.2$ µm, which will help in the smooth SC generation in long wavelength region. Hence, the chosen optimized value for the core width '$a$' is 7.2 µm. The effect of variation in rib height ($b$) on dispersion characteristics is shown in Fig. 3. As $b$ increases from 7.8 to 8.2 µm the dispersion values at pump wavelength of 2.5 µm changes from − 1.130 ps/nm/km to − 1.417 ps/nm/km. It can be seen from Fig. 3 that with the increase in the value of core rib height ($b$) the dispersion curves almost overlaps and in the whole anomalous dispersion region the dispersion values for $b = 7.8$ µm are lesser as compared to other values of $b$. Hence, for the broad and efficient SC we have chosen $b = 7.8$ µm as the optimized value of rib height. Similarly, the effect of altering the values of thickness layer of CdTe ($t$) on dispersion profile is also studied and is shown in Fig. 4. With the increase in the value of $t$, the dispersion values in the anomalous region decreases and that is the reason for us to

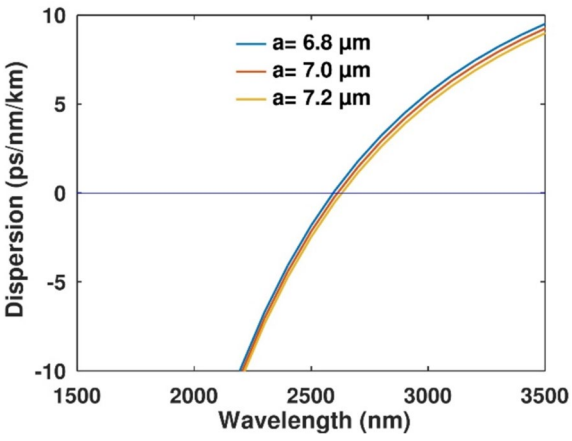



**Fig. 2** Dispersion profile with the alteration in core width '$a$'

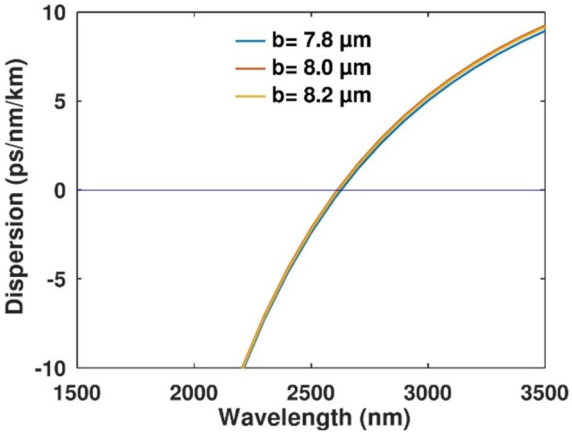



**Fig. 3** Dispersion profile with the alteration in core rib height 'b'

choose the optimized value of $t$ as 3.2 µm. It is worth mentioning the as with the increase in all the geometric parameters, the dispersion values in anomalous region is decreasing but taking the fabrication feasibility and the achieved results, we have finalized the optimized parameters.

The optimized parameters for the final designed structure are shown in Table 2 and from the foregoing study it can be deduced that for the chosen optimized parameters $a = 7.2$ µm, $b = 8.2$ µm and $t = 3.2$ µm, dispersion prolife offered by the waveguide is flat and optimized. The dispersion profile of the waveguide structure with optimized parameters is illustrated in Fig. 5 and it can be seen that the waveguide offers a flat dispersion curve in the wavelength range of 4–15 µm which is within the dispersion value of 10.436 ps/nm/km in anomalous dispersion region. It offers Zero dispersion wavelength at around 2.6 µm and

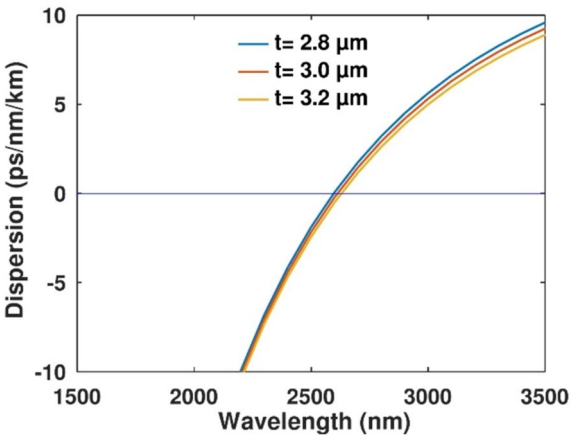



**Fig. 4** Dispersion characteristics with the variation in '$t$' while other parameter are held constant

**Table 2** Optimized Parameters of the proposed rib waveguide structure

| Geometrical Parameters | Core Width ($a$) | Rib height ($b$) | CdTe thickness ($t$) |
|---|---|---|---|
| Values | 7.2 µm | 7.8 µm | 3.2 µm |

the value of dispersion at the chosen pump wavelength of 2.5 μm is − 1.627 ps/nm/km. On taking the practical fabrication errors into account, we have also simulated our optimized final waveguide design with the uncertainty of ± 2% in the parameters and we observed that there is no significant change in the dispersion values, for 2% change in optimized parameters, ~ 0.6% change in dispersion value at pump wavelength is observed.

The variation of $A_{\text{eff}}$ and non-linear coefficient with wavelength is illustrated by Fig. 6. At the pump wavelength of 2.5 μm, the values of $A_{\text{eff}}$ and non-linear coefficient are 4.02 $\mu m^2$ and 3123.407 $W^{-1}$ $km^{-1}$. With the uncertainty of ± 2% in the optimized parameters, there is only 0.04% variation in the values of effective mode area and 0.03% change in the value of material non-linearity. Hence, it can be concluded that our proposed design is robust and stable with minor fabrication errors.

Supercontinuum generation is studied by solving the Eq. (10) with the help of split step Fourier transform method. This method efficiently solves the GNLSE by solving the linear dispersion effects in frequency domain and non-linear effects in time domain. The dispersion coefficients upto 9th order are taken into consideration while simulation SC spectra as the coefficients greater than 9th order does not have a significant effect in the SC generation. The value of non-linear length $L_{NL}(= 1/\gamma P_0)$ is calculated to be $2 \times 10^{-2}$ mm. Now, the optical wave breaking distance $Z_{WB}$ using Eq. (9) comes out to be ~ 4 mm. Now, supercontinuum spectrum is simulated using different propagation lengths of waveguide and different laser operating conditions like by varying pulse width and peak power of input ultrashort pulse. The effect on supercontinuum spectrum with the variation in length is shown in Fig. 7. A conclusion can be drawn from Fig. 7 that the SC broadening increases with the increase in waveguide length until it reaches the maximum limit of transmission window of CdTe. Hence, with 50 fs pulses having peak power 16 kW our proposed waveguide is expected to produce SC broadening of

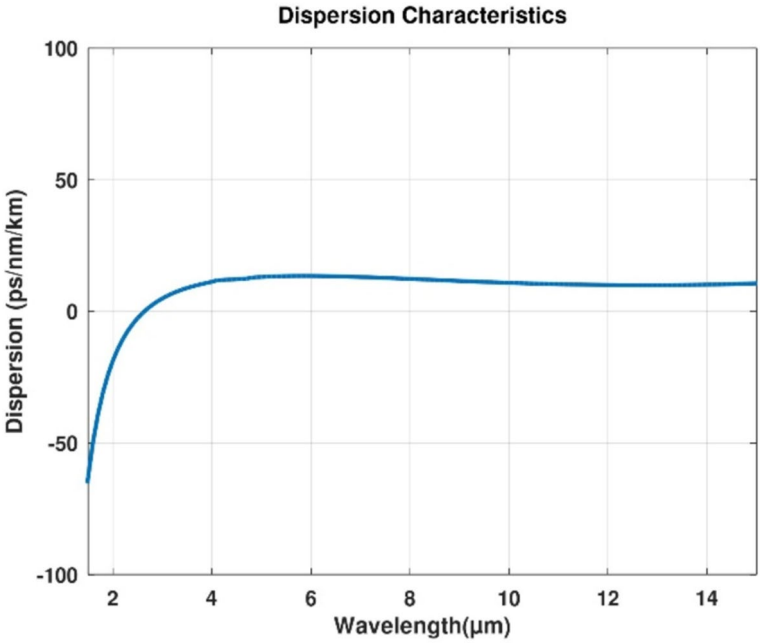



**Fig. 5** Dispersion profile of optimized structural design

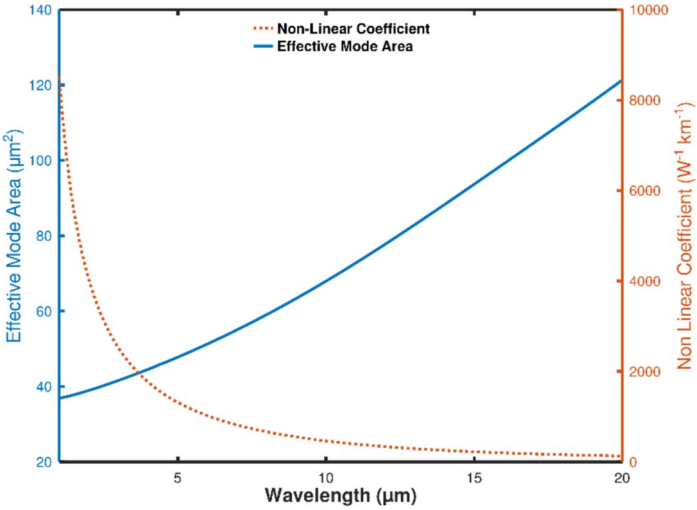



**Fig. 6** Spectral Variation of $A_{\text{eff}}$ and Non-linear coefficient

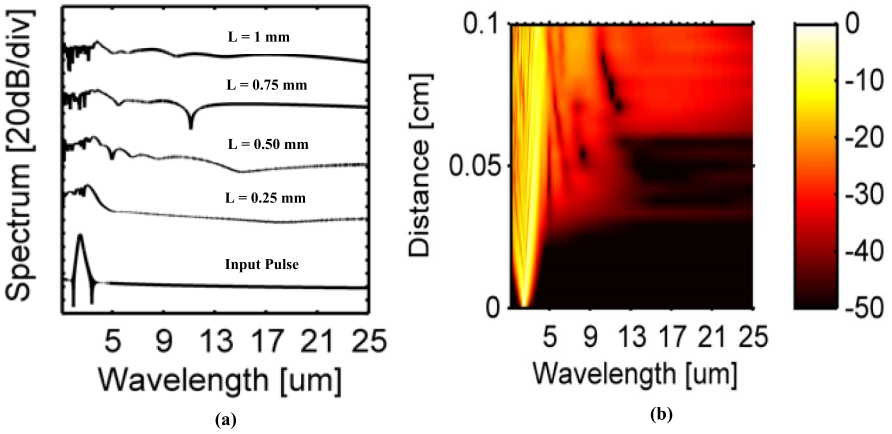



**Fig. 7** **a** Line plot of SC spectra with length variation. **b** Spectral evolution plot of SC spectra

1.5–25 μm in only 1 mm length of fiber which is also clear from the spectral density plot shown in Fig. 7b. It can also be inferred from this that optical wave breaking does not contribute to SC broadening. In the presented waveguide the coherent supercontinuum broadening is due to the dominance of self-phase modulation. The coherent SC spectra can be attributed to SPM as self-phase modulation is a self-generated phenomenon in which it increases the spectral content of the pulse from intensity dependent phase shift [34]. Thus, the SC spectra which is generated because of SPM effect is highly coherent.

The variation in spectral evolution with the change in pulse width of input pulses is illustrated in Fig. 8. It can be noted that with the increase in the pulse width, the spectral evolution of input pulses starts to decrease. When narrowing the temporal width from

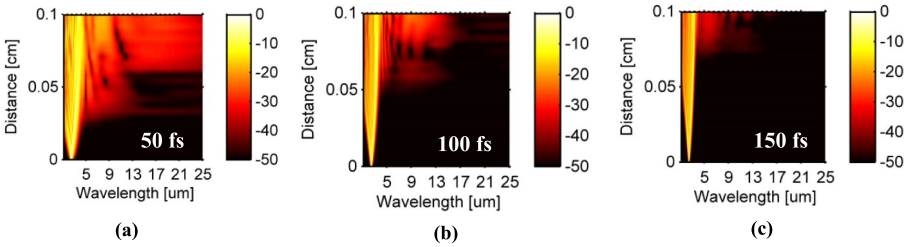



**Fig. 8** Spectral evolution of input pulse with variation in pulse width

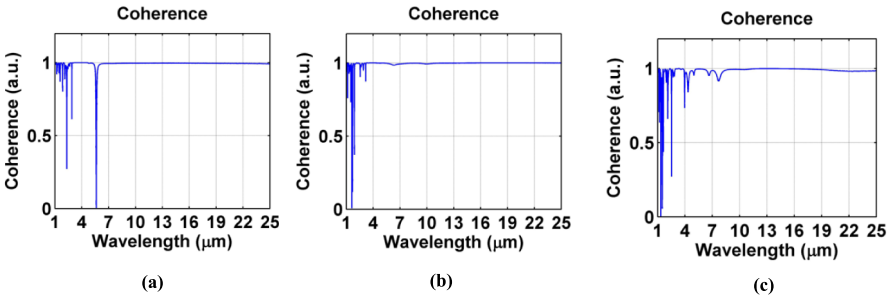



**Fig. 9** Coherence plots with variation in peak power. **a** 12 kW, **b** 16 kW, **c** 20 kW

150 to 50 fs, the resultant SPM effect becomes more dominant, which is the reason for spectral broadening. Here, a simple conclusion can be drawn that the shorter pulse widths favors the spectral broadening process if the propagating length of the waveguide structure is short.

Furthermore, the effect of another fiber parameter i.e. peak power of input pulse on SC spectra coherence is also studied and is illustrated in Fig. 9. It can be noted that as we increase the peak power of the input ultrashort pulse from 12 to 16 kW, coherence offered by presented waveguide structure is high and almost unity over the whole range. This increase can be attributed to the fact that as SPM dominates the broadening at lower peak powers and also as we are working in normal dispersion regime, dispersive wave generation is minimized leading to more coherent spectra at 16 kW. But after 16 kW, upto 20 kW, there is sudden decrease in the coherence value offered which is due to the introduction of noise after a certain threshold peak power.

After examining the impact of various laser operating conditions like pulse width and peak power, we confirm our decision to propose a waveguide structure described above for on-chip mid-infrared supercontinuum generation. Our simulation results in highly coherent SC spectrum broadening covering 1.5–25 μm in only 1 mm long waveguide when pumped with 2.5 μm laser pulses having pulse width 50 fs and peak power 16 kW. Additionally, we have compared our simulation results with previous works on CdTe and other chalcogenide glasses for on-chip supercontinuum generation in waveguide and the comparison is tabulated in Table 3. The final results achieved from the proposed waveguide are shown in Fig. 10 with enlarged view of line plot of spectral broadening at 1 mm distance, spectral evolution and coherence.

**Table 3** Comparison between simulated waveguide results and previous works

| Core material | Structure | SC broadening (μm) | Pump wavelength (μm) | References |
|---|---|---|---|---|
| CdTe/CdS | Waveguide | 3.5–20 | 9 | Long et al. (2022) |
| $Si_3N_4$ | Waveguide | 1.15–2.425 | 2.2 | Zhang et al. (2023) |
| $As_2Se_3$ | Waveguide | 2–15 | 2.5 | Saini et al. (2017) |
| $Si_3N_4/TeO_2$ | Waveguide | 0.88–5.22 | 1.5 | Saini and Supradeepa (2021) |
| Ge–As–Se–Te | SI fiber | 1.5–14 | 4.5 | Zhao et al. (2016) |
| SiGe/Si | Waveguide | 3–8.5 | 4 | Sinobad et al. (2018) |
| CdTe | Waveguide | 1.5–25 | 2.5 | This work |

It is worth to mention here that while this study provides a comprehensive computational and theoretical analysis of the SC generation phenomena in non-linear waveguides, understanding the practical aspects of experimental validation is also important. In practice, SC generation is typically realized using laser source, precision optics including lenses, monochromators, waveguides as a nonlinear medium and spectral analysis equipment. Common experimental setups include configurations with tunable femtosecond lasers, which are often paired with choppers for the modulation of beam and monochromators for the detailed spectral characterization (Ballester et al. 2020). Waveguide, such as our designed waveguide plays a crucial role in enhancing the nonlinear interactions and extending the SC spectrum. To ensure accurate spectral measurements, environmental control measures like nitrogen purging are often employed to reduce absorption effects from atmospheric water vapor (Saini and Sinha 2021). These experimental techniques and setups provide a robust framework for studying and validating the theoretical predictions of SC generation in various nonlinear optical media.

## 5 Conclusion

In conclusion, we have investigated, via computational simulation, highly coherent mid-infrared supercontinuum generation in a rib waveguide structure composed of cadmium telluride core. The selection of this material is driven by the extensive transparency it

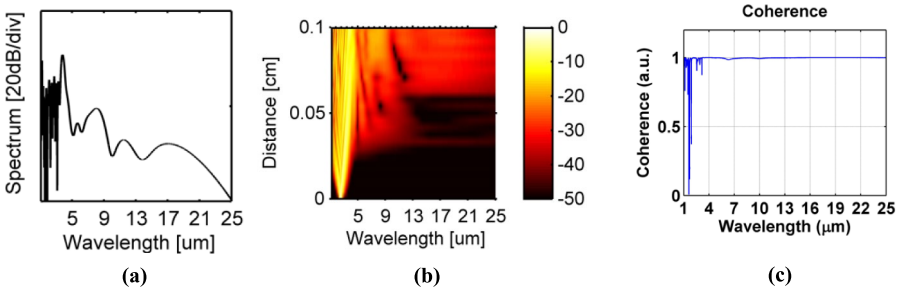



**Fig. 10** **a** Enlarged line plot of SC spectra. **b** Spectral evolution, and **c** Coherence, in 1 mm length injected with pulses having pulse width 50 fs and 16 kW peak power

offers, significant non-linearity and high thermal stability making it a good candidate for SC generation. We analyzed the effect of geometric parameters of waveguide on dispersion characteristics and optimized them to obtain a flat dispersion curve offering a ZDW at around 2.6 μm. Simulation of supercontinuum generation has been done by solving the GNLSE using split step Fourier method with different fiber operating conditions. Finally, the optimized waveguide structure is expected to achieve a highly coherent ultra-broad SC spectra covering 1.5–2.5 μm when pumped with 2.5 μm hyperbolic secant pulses, having pulse width of 50 fs and peak power of 16 kW in only 1 mm long waveguide. The robustness of the waveguide structure is studied by considering the error of $\pm 2\%$ during the fabrication process and from the results it is clear that the proposed waveguide is robust and efficient. Finally, the high coherence and ultra-broadened spectra positions our design as a promising candidate for a diverse range of non-linear applications like environmental monitoring, non-linear microscopy, and biomedical sensing.

**Acknowledgements** Deepak Garg highly acknowledges the support from the Delhi Technological University, Delhi, for providing the research fellowship and necessary support to conduct this research work

**Author Contribution** D Garg has done the simulation work, analysed the results and written the original manuscript and A Kumar has given ideas and reviewed the manuscript.

**Funding** The authors have not disclosed any funding.

**Data availability** No datasets were generated or analysed during the current study.

## Declarations

**Competing Interests** The authors declare no competing interests.

## References

Ballester, M.M., Lafforgue, C., Frigerio, J., Ballabio, A., Vakarin, V., Liu, Q., Ramirez, J.M., et al.: On-chip mid-infrared supercontinuum generation from 3 to 13 μm wavelength. ACS Photonics **7**, 3423–3429 (2020)

Bres, C.S., Della Torre, A., Grassani, D., Brasch, V., Grillet, C., Monat, C.: Supercontinuum in integrated photonics: generation, applications, challenges, and perspectives. Nanophotonics **12**, 1199–1244 (2023)

Chauhan, P., Kumar, A., Kalra, Y.: A dispersion engineered silica-based photonic crystal fiber for supercontinuum generation in near-infrared wavelength region. Optik **187**, 230–237 (2019)

Cheng, Y., Yuan, J., Mei, C., Li, F., Zhou, X., Wu, Q., Yan, B., et al.: Mid-infrared supercontinuum and frequency comb generations by different optical modes in a multimode chalcogenide strip waveguide. IEEE Access **8**, 202022–202031 (2020a)

Cheng, Y., Yuan, J., Mei, C., Li, F., Kang, Z., Yan, B., Zhou, X., et al.: Self-similar picosecond pulse compression for supercontinuum generation at mid-infrared wavelength in silicon strip waveguides. Opt. Commun. **454**, 124380 (2020b)

Chowdury, A., Chang, W., Battiato, M.: Higher-order rogue-wave fission in the presence of self-steepening and Raman self-frequency shift. Phys. Rev. A **107**, 053507 (2023)

Dai, S., Wang, Y., Peng, X., Zhang, P., Wang, X., Xu, Y.: A review of mid-infrared supercontinuum generation in chalcogenide glass fibers. Appl. Sci. **8**, 707 (2018)

Du, Q., Luo, Z., Zhong, H., Zhang, Y., Huang, Y., Du, T., Zhang, W., Gu, T., Hu, J.: Chip-scale broadband spectroscopic chemical sensing using an integrated supercontinuum source in a chalcogenide glass waveguide. Photonics Res. **6**, 506–510 (2018)

Fiore, T., Pellerito C.: Infrared absorption spectroscopy. In: Spectroscopy for Materials Characterization, pp. 129–167 (2021)

Garg, D., Singh, J., Kumar, A.: Design and theoretical study of rectangular photonic crystal fiber based sensor for chemical sensing in terahertz regime. Indian J. Phys. (2024). https://doi.org/10.1007/s12648-024-03343-9

Gauthier, J.C., Robichaud, L., Fortin, V., Vallée, R., Bernier, M.: Mid-infrared supercontinuum generation in fluoride fiber amplifiers: current status and future perspectives. Appl. Phys. B **124**, 1–14 (2018)

Hiscocks, M.P., Ganesan, K., Gibson, B.C., Huntington, S.T., Ladouceur, F., Prawer, S.: Diamond waveguides fabricated by reactive ion etching. Opt. Express **16**, 19512–19519 (2008)

Hu, S., Liu, L., Wang, Y., Ren, Y., Tao, H., Feng, Y., Huang, C.: Supercontinuum generation in aqueous core/shell structured nanoparticles. J. Nanophotonics **9**, 093065–093065 (2015)

Hunsperger, R. G.: Waveguide fabrication techniques. In: Integrated Optics: Theory and Technology, pp. 53–84 (2009)

Kaur, B., Kumar, S., Kaushik, B.K.: Advances in photonic crystal fiber: sensing and supercontinuum generation applications. Opt. Fiber Technol.fiber Technol. **72**, 102982 (2022)

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 (2024)

Liao, M., Chaudhari, C., Qin, G., Yan, X., Suzuki, T., Ohishi, Y.: Tellurite microstructure fibers with small hexagonal core for supercontinuum generation. Opt. Express **17**, 12174–12182 (2009)

Long, Z., Yang, H., Li, Y., Wu, H., Liang, H.: Cadmium telluride waveguide for coherent MIR supercontinuum generation covering 3.5–20 µm. Opt. Express **30**, 2265–2277 (2022)

Matsuura, Y., Harrington, J.A.: Hollow glass waveguides with three-layer dielectric coating fabricated by chemical vapor deposition. JOSA A **14**, 1255–1259 (1997)

Medjouri, A., Djamel, A.: Theoretical study of coherent supercontinuum generation in chalcohalide glass photonic crystal fiber. Optik **219**, 165178 (2020)

Ming, F.Y., Zhen, Y., Peng, X.P., Lun, Y.K., Yan, S., Ping, W.R.: Dispersion engineered ZnSe rib waveguide for mid-infrared supercontinuum generation. J. Infrared Millim. Waves **40**, 508–515 (2021)

Moon, S., Kim, D.Y.: Ultra-high-speed optical coherence tomography with a stretched pulse supercontinuum source. Opt. Express **14**, 11575–11584 (2006)

Ray, S., Tarafder, K.: Investigation of CdSe and ZnSe as potential back surface field layers for CdTe-based solar cells: a study from first principles calculations. Adv. Theory Simul. **6**, 2200718 (2023)

Saini, T.S., Sinha, R.K.: Mid-infrared supercontinuum generation in soft-glass specialty optical fibers: a review. Prog. Quantum Electron.. Quantum Electron. **78**, 100342 (2021)

Saini, T.S., Supradeepa, V.R.: Tellurium-oxide coated silicon-nitride hybrid waveguide for near-to-mid-IR supercontinuum generation: design and analysis. J. Mod. Opt. **68**, 29–36 (2021)

Saini, T.S., Kumar, A., Sinha, R.K.: Design and modelling of dispersion-engineered rib waveguide for ultra-broadband mid-infrared supercontinuum generation. J. Mod. Opt. **64**, 143–149 (2017)

Saini, T.S., Hoa, N.P.T., Nagasaka, K., Luo, X., Tuan, T.H., Suzuki, T., Ohishi, Y.: Coherent midinfrared supercontinuum generation using a rib waveguide pumped with 200 fs laser pulses at 2.8 µm. Appl. Opt. **57**, 1689–1693 (2018)

Saini, T.S., Tuan, T.H., Suzuki, T., Ohishi, Y.: Coherent mid-IR supercontinuum generation using tapered chalcogenide step-index optical fiber: experiment and modelling. Sci. Rep. **10**(1), 2236 (2020)

Sharma, R., Kaur, S., Chauhan, P., Kumar, A.: Computational design & analysis of $GeSe_2$–$As_2Se_3$–PbSe based rib waveguide for mid-infrared supercontinuum generation. Optik **220**, 165032 (2020)

Sinobad, M., Monat, C., Davies, B.L., Ma, P., Madden, S., Moss, D.J., Mitchell, A., et al.: Mid-infrared octave spanning supercontinuum generation to 8.5 µm in silicon-germanium waveguides. Optica **5**, 360–366 (2018)

Tung, K.K., Wong, W.H., Pun, E.Y.B.: Polymeric optical waveguides using direct ultraviolet photolithography process. Appl. Phys. A **80**, 621–626 (2005)

Wang, H., Yang, L., Yang, Z., Kang, Z., Yang, P., Zhang, W., Wang, R., Xu, P.: Nanophotonic chalcogenide waveguides for supercontinuum generation pumped at 1550 nm. Infrared Phys. Technol. **139**, 105281 (2024)

Yang, L., Zhang, B., He, X., Deng, K., Liu, S., Hou, J.: 20.6 W mid-infrared supercontinuum generation in ZBLAN fiber with spectrum of 1.9–4.3 µm. J. Lightwave Technol. **38**, 5122–5127 (2020)

Yu, Y., Gai, X., Ma, P., Vu, K., Yang, Z., Wang, R., Choi, D.Y., Madden, S., Luther-Davies, B.: Experimental demonstration of linearly polarized 2–10 µm supercontinuum generation in a chalcogenide rib waveguide. Opt. Lett. **41**, 958–961 (2016)

Yupapin, P., Sharafali, A., Sreethu, P., Sridarshini, T., Adhikari, D., Karthikeyan, M.: Modulation instability induced Supercontinuum generation in defective core photonic crystal fiber. Opt. Quant. Electron. **54**, 504 (2022)

Zhang, J., Zhang, S., Jiang, X., Yao, J., Wang, A., Zhan, Q.: Numerical study on flat dispersion in a single-layer $Si_3N_4$ coupled-waveguides with four zero-dispersion wavelengths for supercontinuum generation. Opt. Laser Technol. **160**, 109061 (2023)

Zhao, Z., Wang, X., Dai, S., Pan, Z., Liu, S., Sun, L., Zhang, P., et al.: 1.5–14 μm mid-infrared supercontinuum generation in a low-loss Te-based chalcogenide step-index fiber. Opt. Lett. **41**, 5222–5225 (2016)

**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.

Plasmonics
https://doi.org/10.1007/s11468-024-02723-5

RESEARCH

# Highly Sensitive Black Phosphorus–Layered SPR-PCF Refractometric Sensor

**Plakshi Gupta[1] · Akash Khamaru[1] · Ajeet Kumar[1]**

Received: 7 November 2024 / Accepted: 13 December 2024
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2025

## Abstract

This work presents a D-shaped surface plasmon resonance (SPR)-based photonic crystal fiber (PCF) sensor with a coating of gold-$TiO_2$ as a plasmonic material along with black phosphorus (BP) nano-films which works as a sensitivity enhancer functioning in visible to near-infrared spectrum. The affixed BP layers offer a direct bandgap, consequently advancing the proposed sensor's performance. The designed sensor is employed for the detection of a broad range of refractive indices limiting from 1.3–1.4. Various fabrication procedures are discussed for layering of BP films on the PCF surface such as liquid peeling and pulsed laser precipitation technique. Along with the influence of BP layers, the thickness of gold and $TiO_2$ is also optimized for the extraction of favorable outputs. The dimensions of respective cladding air holes of the PCF are also influenced to assess the effectiveness of the proposed sensor. The dynamic sensitivity of the proposed sensor elevated from 9000 to 23,300 nm/RIU with the incorporation of the BP films in the designed PCF RI detector. Additionally, a high amplitude sensitivity of 7446.56 $RIU^{-1}$ is also achieved. The theoretical structure is numerically analyzed using the finite element method in COMSOL Multiphysics Software. The extracted enhanced sensitivity perfectly elaborates the efficacy of the BP layers in the PCF-SPR RI sensor and builds our suggested sensor as a prospective participant in modern plasmonic sensors.

**Keywords** Black phosphorus · D-shaped · Surface plasmons · Wavelength sensitivity · Amplitude sensitivity

## Introduction

Surface plasmons (SPs) were initially introduced by Ritchie, in the 1950s [1]. By applying the theory of SPs, the prism-coupled SPR configuration was studied by Otto where prism and plasmonic materials were separated by the dielectric medium [2]. The setup's configuration was then upgraded by the Kretschmann where prism and plasmonic material were in direct contact [3]. The latter configuration was used for generating surface plasmon waves, by matching the frequency of incident light and surface electrons (i.e., free electrons in metals). In the 1980s, these methods were used to demonstrate SPR sensors for biological and chemical sensing [4].

Surface plasmon resonance is a phenomenon that occurs at the surfaces of metals (like gold, silver), whenever electrons interact with a light beam that strikes the surface at a particular angle. The interaction is sensitive to the changes in the nearby environment or simply the changes in refractive index. Compared with the traditional sensing methods, the sensing technology adopting the surface plasmon resonance technique has the advantage of real-time analysis; it is much more sensitive to the changes in surroundings and is label-free. Therefore, the SPR sensing approach has proved a remarkable detection technique due to its higher sensitivity and a vast number of applications like biosensing, antibody interaction, clinical trials, food quality analysis, and gas detection. The conventional prism-based Kretschmann setup is being widely used for SPR sensors, in which the prism is coated with the plasmonic material. The deviation in the refractive index changes the propagation constant of the surface plasmon mode. This results in variation in coupling conditions. But the prism-based SPR sensor requires

✉ Ajeet Kumar
ajeetdph@dtu.ac.in

Plakshi Gupta
plakshigupta_23mscphy74@dtu.ac.in

Akash Khamaru
akashkhamaru_23phdap02@dtu.ac.in

[1] Advanced Photonics Simulation Research Laboratory, Department of Applied Physics, Delhi Technological University, Shahbad Daulatpur, Bawana Road, Delhi 110042, India

Published online: 21 January 2025

a bulky configuration with optical and mechanical components and is also not cost-friendly. As a result, the discussed configurations are not suitable for field-based applications [5]. These limitations of prism-based SPRs gave rise to the optical fiber-based SPR sensors in the 1990s [6]. The fiber-based SPR sensors were being used to provide a wider operating range and better resolution. But the major disadvantage of traditional PCF-SPR sensors is that they require the direct incidence of light at a narrow angle. To overcome these problems, photonic crystal fibers (PCFs) were introduced in the 2000s for employing them for SPR-based detection [7].

PCFs are special types of optical fibers which have tiny air holes running along their length. These air holes control how the light moves throughout the fiber. It is possible to manipulate the light propagation more precisely with PCFs than with traditional fibers. PCFs have the advantages of high birefringence, higher nonlinearity, and flexible structural design. These properties give rise to a large number of applications like terahertz communication [8], terahertz-based sensors [9, 10], refractive index sensors [11, 12], pressure sensors [13], electric and magnetic field sensors [14], supercontinuum generation [15–17], high-power fiber lasers [18], and fiber components for metrology and spectroscopy [19]. Photonic crystal fibers are currently being produced in a large number of laboratories by using different techniques like stacking of capillaries and rod extrusion, sol-gel casting, and drilling. Materials such as silica, fluoride glasses, polymers, and chalcogenide glasses are being used for the fabrication of PCFs [20].

PCF-based SPR sensors arise as types of sensors that are used to detect tiny amounts of substances in various domains such as biology, environment, and medicine. Looking at its advantages, such as smaller size, single-mode propagation, effortless light launching, and ability to control evanescent field penetration, have made it a promising candidate for sensing. The light enters and travels through the core of the PCF, and as the light strikes the plasmonic material (like silver, gold), it causes electrons in the metals to oscillate. A strong interaction happens when light and electron oscillations match up, and the phenomenon is known as resonance. If the environment around the fiber changes or the refractive index of the analyte changes, the resonance condition shifts. This shift changes the incident light's wavelength, which a PCF-SPR sensor can detect. Even a tiny amount of change in the surrounding environment will be noticed by a very sensitive PCF-SPR sensor. They are mainly investigated numerically or analytically using a simulation technique, known as the finite element method via COMSOL software in which first the structure is divided into smaller parts which is known as meshing and each part has its own set of equations. All the equations are combined at the final stage and are calculated, which is considered simpler to solve. In the current scenario, numerous SPR-PCF sensors have been proposed altering the prism-based technique [21, 22]. There are many types of SPR-PCF designs using metal-coated holes [23], U-grooved [24], etc. already proposed in previous literatures. But they inherit some disadvantages like fabrication complexity and difficulty in the deposition of metals in holes [25]. Several D-shaped PCFs which are flat from the top are also being reported in the present literature, and the same is also proposed in this work [11]. The top portion of PCF with air holes is polished. This provides a larger area for interaction between core-guided light and surface plasmons, effectively improving the overall sensitivity.

Graphene, silicene, germanene, phosphorene, and black phosphorus are reported to possess plasmonic properties [26–28]. Black phosphorus (BP) is a new type of 2-D plasmonic material that is attracting the attention of researchers due to its electrical and optical properties which include direct band gap, smaller work function, high charge carrier mobility, and better binding of molecules on the surface of sensors [29]. Direct band gap allows BP to interact with light at specific wavelengths which enhances the sensitivity of the sensor. Smaller work function allows the transfer of electrons between the surrounding environment and BP layers. High charge carrier mobility makes the sensor transmit signals fast and efficient which results in rapid response and higher sensitivity. The ability to bind molecules tightly on the surface allows BP to be used in biological and chemical sensing. So BP enhances the sensor's ability by providing molecules (like proteins, DNA, or chemical analytes) an active site to attach to. Black phosphorus is integrated into the SPR sensors by adding multiple thin layers of 0.53 nm thickness each on the surface. The layering of BP can be done through various techniques like mechanical peeling, liquid peeling, and pulse layer precipitation [30, 31]. Mechanical peeling involves physically peeling a thin layer of BP using special tools from a larger piece. It is the most commonly used method, as it does not require complex equipment and is cost-efficient. On the other hand, the liquid peeling technique involves placing the bulk piece in a liquid and then subjecting it to ultrasound waves, which produce a large number of thin BP layers. Similarly, the pulsed laser precipitation technique involves the firing of a powerful laser at the bulk material. This method allows to control the size and thickness of the layer. The application of black phosphorus in SPR sensors includes better biomolecule detection, environment monitoring, and clinical applications, and it helps in the prevention of oxidation of plasmonic layers. These are some of the reasons that led to the attraction of researchers toward black phosphorus. Tong et al. [30] proposed an ultrashort side polished structure, with black phosphorus layers, which has a dynamic sensitivity of 8200 nm/RIU, for different values of RI starting from 1.320 to 1.375.

A $TiO_2$/gold/black phosphorus–layered configuration-based, D-shaped refractive index sensor is proposed in this

work. $TiO_2$ film works as an adhesive agent between the PCF body and plasmonic material [32]. Gold is the most favorable plasmonic material because of its chemical stability and corrosion resistance against different environments [33]. It also allows a large shift in peak resonance wavelength. BP nano-films are used to elevate the sensor's sensitivity. Numerous smaller air holes are kept near the fiber core, which helps to squeeze the maximum energy of the core mode to couple with the plasmonic metal and hence enhances the sensitivity of the sensor [34]. Moreover, the structural parameters, like the thickness of $TiO_2$ and gold layer, and the number of BP layers are optimized to achieve better sensing performance of the proposed sensor.

## Structural Parameters

The proposed D-shaped SPR-PCF biosensor's cross-section view and magnified view are shown in Fig. 1a. The PCF structure consists of two types of air holes distributed in the cladding in a triangular lattice. The bigger air holes having a radius of $r_1 = 1\mu$m are distributed in the cladding. The smaller air holes with a radius of $r_2 = 0.3\mu$m are placed near the fiber core in such a way that the distribution forms an inverted triangular core. The plasmonic metal layer, that is, gold of thickness, $t_{Au} = 35$nm, is layered on the flat surface of the D-shape PCF sensor. Along with gold, $TiO_2$ of thickness, $t_{Ti} = 14$nm, is used as an adhesive between the PCF surface and plasmonic metal film and also works as a wavelength sensitivity enhancer. To elevate the sensitivity to a greater extent, a definite number of layers of black phosphorus (BP) each of thickness $t_{BP} = 0.53$nm is layered above the plasmonic materials. For design convenience, the perfectly matched layer (PML) is used in cylindrical geometry and is of width 1 $\mu m$ to prevent the back reflection effect. The refractive index of the measured medium, that is, the analyte of thickness $t_a = 1.5\mu m$, is defined as $n_a$.

Figure 1b represents the schematic diagram for the experimental setup. The light is transmitted to the proposed D-shaped PCF from a broadband source and through a single-mode fiber (SMF). The light then passes through the optical spectrum analyzer (OSA) which is connected to the PCF from the other end. Further, the data is collected using a computer.

The plasmonic metal consists of two kinds of electrons that are free electrons and bound electrons. Paul Drude gave the Drude model for free electrons, and Hendrik Lorentz gave the Lorentz oscillator model for bound electrons, which

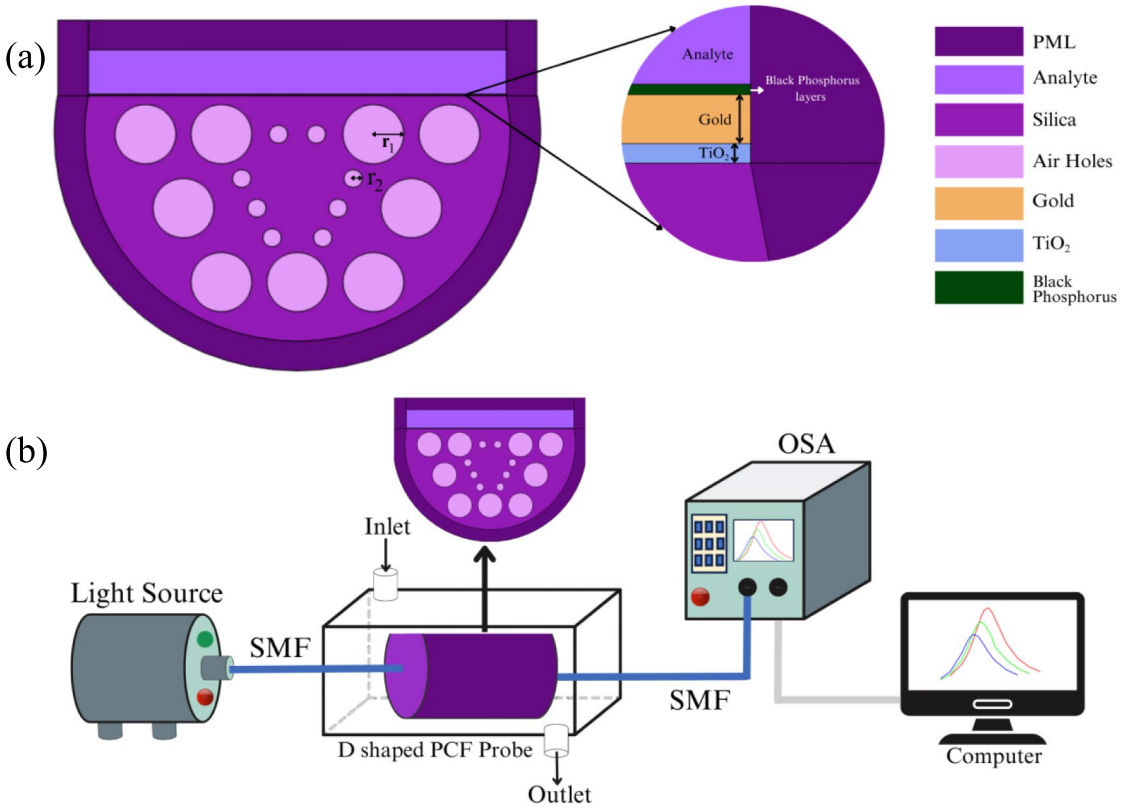



**Fig. 1** **a** Cross-section of the proposed D-shaped PCF-SPR sensor. **b** Schematic diagram for experimental setup

are used to find the dielectric constant of the metal, i.e., gold. The combined model is known as the Drude Lorentz model, which is then used to find the refractive index of the metal [32].

$$\epsilon_{Au} = \epsilon_{\infty} - \frac{\omega_D^2}{\omega(\omega + j\gamma_D)} - \frac{\Delta\epsilon\Omega_L^2}{\omega^2 - \Omega_L^2 + j\Gamma_L\omega} \tag{1}$$

Here, $\epsilon_{\infty}$ is the permittivity at a high frequency having the magnitude 5.9673 that occurs when the electric field of the light, i.e., an electromagnetic wave, is very high, and the electrons cannot respond to that. $\gamma_D$ is the damping factor as free electrons cannot keep moving forever and attain the magnitude ~ $2\pi \times 15.92$ THz, $\omega_D$ (~ $2\pi \times 2113.6$ THz) is the frequency by which the free electrons in the metal oscillate when disturbed by light. ω is the frequency of light, which, in the case of an RI sensor, depends on the wavelength of the light. $\Gamma_L$ (~ $2\pi \times 104.86$ THz) is the damping factor for bound electrons, $\Omega_L$ (~ $2\pi \times 650.07$ THz) is the frequency of bound electrons that vibrate when light hits them, and Δϵ is the change in permittivity that occurs due to the vibrations of bound electrons.

The material used for the composition of the PCF is pure silica, and as its refractive index is a function of wavelength, therefore, wavelength-dependent Sellmeier equation is used:

$$n^2(\lambda) = 1 + \frac{B_1\lambda^2}{\lambda^2 - C_1} + \frac{B_2\lambda^2}{\lambda^2 - C_2} + \frac{B_3\lambda^2}{\lambda^2 - C_3} \tag{2}$$

where λ is the wavelength of light and $B_i, C_i$ are constants with $B_1$=0.696163, $B_2$=0.4079426, $B_3$ = 0.8974794, $C_1$ =0.0684043, $C_2$=0.1162414, and $C_3$=9.896161.

A thin layer of $TiO_2$ is incorporated just above the gold layer to act as an adhesive agent between the PCF structure and plasmonic film. The RI of $TiO_2$ is calculated using the given wavelength-dependent formula, where λ is in the *Å* units [35]:

$$n^2 = 5.913 + \frac{0.2441}{\lambda^2 - 0.0803} \tag{3}$$

The refractive index of black phosphorus used in the structure is measured using the Cauchy absorbent model [36], where λ is in nanometer units. The layers of black phosphorus are introduced to enhance the sensitivity.

$$n(\lambda) = A + B\frac{10^4}{\lambda^2} + C\frac{10^9}{\lambda^4} \tag{4}$$

$$k(\lambda) = D \times 10^{-5} + E\frac{10^4}{\lambda^2} + F\frac{10^9}{\lambda^4} \tag{5}$$

Here, $n(\lambda)$ is the real part, $k(\lambda)$ is the imaginary part of the refractive index, and $A = 3.57$, $B = 6.79$, $C = 39.99$, $D = 3206$, $E = -0.521$, and $F = 10.26$ are constants.

The confinement loss signifies the amount of light that gets leaked from the PCF core toward the plasmonic metal layers and can be calculated using the following equation [34]:

$$\alpha_{\text{loss}}(\text{dB/cm}) = 8.686 \times \frac{2\pi}{\lambda} \times \text{Imag}\left(n_{\text{eff}}\right) \times 10^4 \tag{6}$$

where $\alpha_{loss}$ is the confinement loss in the units of dB/cm, $\lambda$ is the wavelength, and $\text{Imag}\left(n_{\text{eff}}\right)$ is the imaginary part of effective RI.

The performance of the sensor can be evaluated using wavelength sensitivity and amplitude sensitivity. Wavelength sensitivity can be evaluated using the following equation [37]:

$$S_w(\text{nm/RIU}) = \frac{\Delta\lambda_{\text{peak}}}{\Delta n_a} \tag{7}$$

where $\triangle\lambda_{peak}$ denotes the change in the peak wavelength when the refractive index of the analyte is changed around the PCF sensor and $\Delta n_a$ is the change in the refractive index of the analyte.

Further, the performance of the sensor can also be calculated by analyzing the amplitude sensitivity of the sensor which eliminates the complex wavelength interpolation used in finding the wavelength sensitivity. It can be expressed as follows [38]:

$$S_A(\lambda) = -\frac{1}{\alpha(\lambda, n_a)}\frac{\partial\alpha(\lambda, n_a)}{\partial n_a} \tag{8}$$

where $\alpha(\lambda, n_a)$ is overall loss at a particular wavelength, $\partial\alpha(\lambda, n_a)$ provides the difference between two consecutive loss spectra, and $\partial n_a$ is the change in the refractive index of the analyte.

## Results and Discussion

In this section, the analysis of simulated results is discussed, with all the possible combinations of structural parameters to increase the sensitivity of the sensor.

### Dispersion Relation

The electric field distribution of *x*- and *y*-polarized core mode is shown in Fig. 2a, b, respectively. The surface polariton mode (SPP), shown in Fig. 2c, is the mode where light waves travel along the interface of the metal and the PCF structure. When the coupling between core mode and SPP mode takes place, it enables the respective sensor to detect sensitive changes in the refractive index or surrounding environment. This is known as coupling mode and is also defined

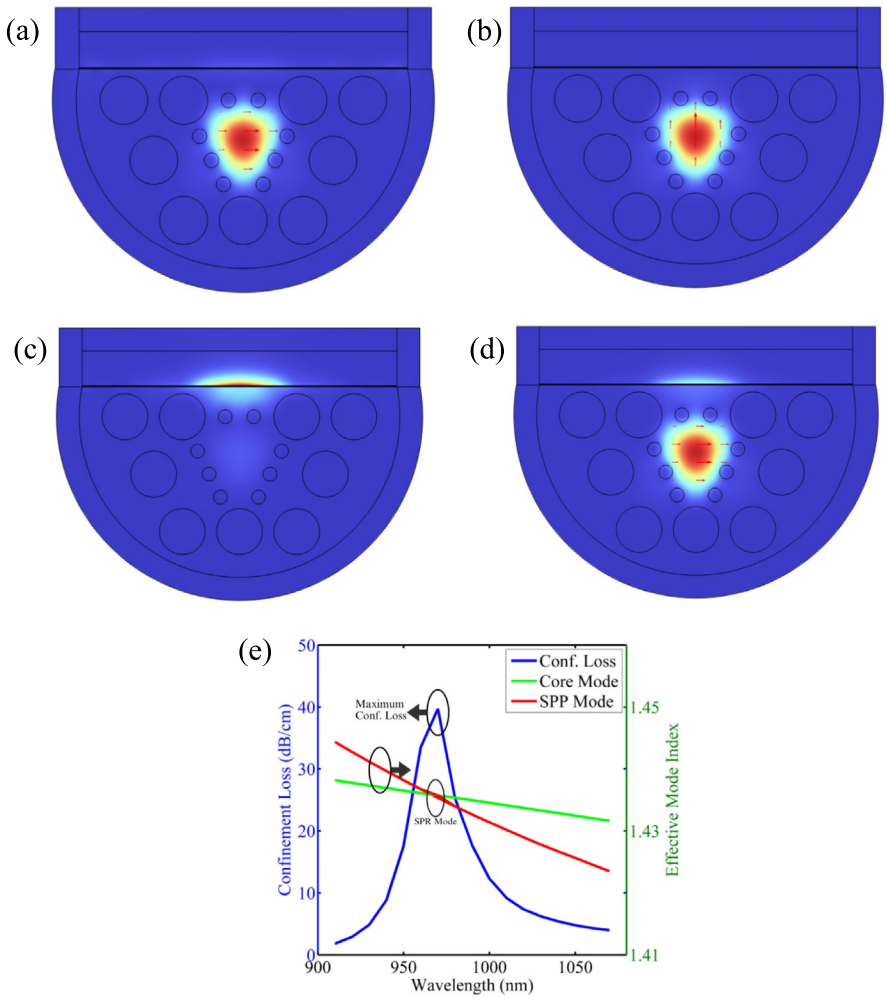



**Fig. 2** Electric field distribution for **a** *x*-polarized core mode, **b** *y*-polarized core mode, **c** SPP mode, **d** coupled mode, and **e** dispersion relation for the optimized structure at 1.38 refractive index

as surface plasmon resonance (SPR) mode (Fig. 2d). Here, the *y*-polarized core mode is taken into consideration as it provides immense visibility of the coupling effect as compared to the *x*-polarized core mode because the plasmonic material is placed at the top surface of the PCF. Figure 2e represents the loss curve and dispersion relation for the proposed sensor at a refractive index of 1.38 for the analyte. As the resonance takes place between the core mode and SPP mode, the loss increases with an increase in wavelength and reaches to maximum value of 39.72 dB/cm at the resonant wavelength of 970 nm. At the resonant point, the RI of core mode and SPP mode intersect, and maximum light gets coupled with the metal interface.

## Investigation with Different Parameters of Proposed Sensor

The thickness of gold, $TiO_2$, number of layers of black phosphorus, and the diameters of big and small air holes can affect the sensor's performance. The following section discusses the optimization of the structural parameters and the simultaneous effect on the sensor's performance which is influenced by optimizing these parameters.

### Performance with Gold Layer Thickness ($t_{Au}$)

The performance of the sensor is initially optimized by affecting the thickness of the gold layer since the thickness of the gold layer hugely influences the loss profile of a certain analyte and ultimately varies the figure of merit. The gold layer thickness is varied from 30 to 40 nm at a rate of 5 nm by keeping other parameters constant as $t_{Ti} = 14$ nm, $r_1 = 1$ µm, $r_1 = 0.3$ µm. The loss peaks are 8.6, 5.09, and 4.38 dB/cm with the gold layer thickness of 30 nm, 35 nm, and 40 nm, respectively, as shown in Fig. 3a. On increasing the gold film thickness, it can give rise to larger damping loss resulting in a broader peak and correspondingly decrease the peak confinement loss also. Therefore, to consider a favorable spectral width of resonance peak and moderate loss value, 35 nm is chosen as the optimized gold layer thickness.

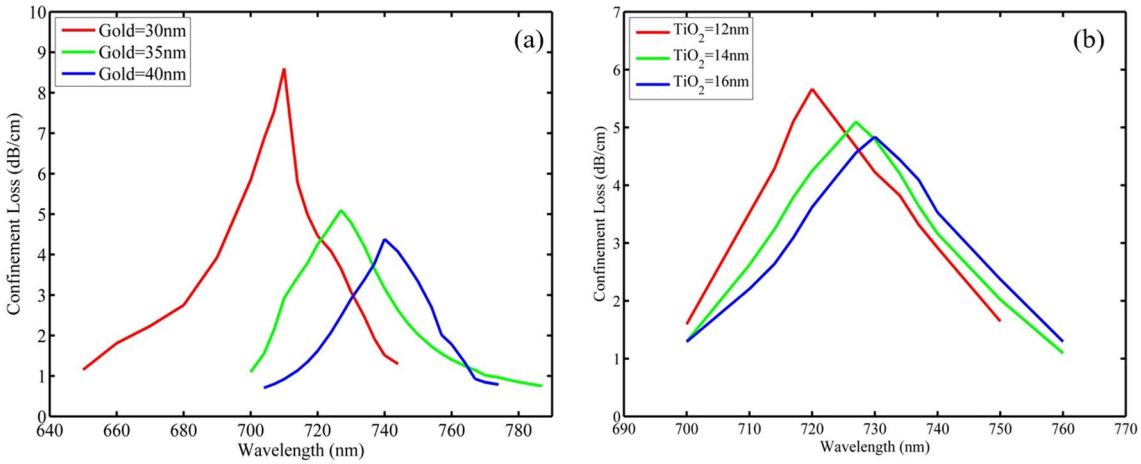



**Fig. 3** Shifts in loss peak at $n_a = 1.38$ **a** with gold layer thickness and **b** with $TiO_2$ layer thickness

### Performance with $TiO_2$ Layer Thickness ($t_{Ti}$)

The second optimization step is performed by optimizing the thickness of $TiO_2$ plasmonic layers. The thickness of $TiO_2$ is assumed from 12 to 16 nm at a changing rate of 2 nm. The overall loss relationship analysis is displayed in Fig. 3b. The plot image shows the change in peak loss values that are 5.66, 5.09, and 4.56 dB/cm with a thickness of $TiO_2$ as 12 nm, 14 nm, and 16 nm, respectively, when the optimized value of gold as 35 nm is taken. It can be seen that the loss value decreases with an increase in the thickness of the $TiO_2$ layer because lesser thickness results in a high penetration depth of core mode toward SPP mode. On comparing the sensitivity with different thicknesses, 14 nm is chosen as the optimized thickness.

### Performance with Different Layers of Black Phosphorus

The number of black phosphorus layers plays an important role in shaping the sensitivity of the SPR sensor. Black phosphorus is a 2-D material that has various advantages as a plasmonic material which are already discussed in the "Introduction" section. Various previous evidences justify that introducing a certain number of BP layers one after another significantly increases the sensitivity but decreases after incorporating more BP layers beyond that certain point. The thickness of a single BP layer is defined as 0.53 nm, which distinguishes the material as a 2-D plasmonic material. The loss profiles for embodied 11, 12, 13, 14, and 15 BP layers above the pre-applied plasmonic material (gold+$TiO_2$) are calculated as 31.25 dB/cm, 35.29 dB/cm,

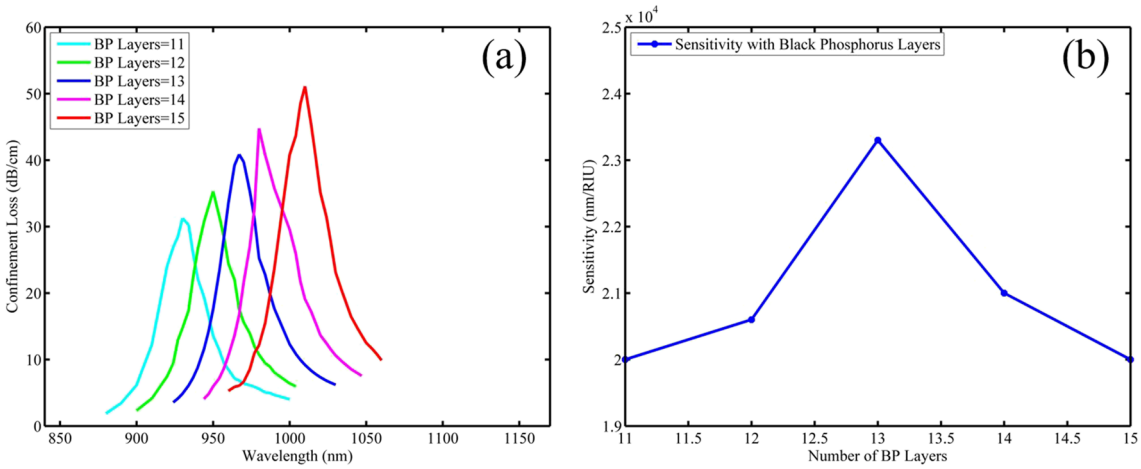



**Fig. 4** **a** Shifts in loss peak at $n_a = 1.38$ with different BP layers and **b** sensitivity variation with BP layers for $n_a = 1.39$

40.86 dB/cm, 44.77 dB/cm, and 51.09 dB/cm, respectively, as shown in Fig. 4a. With the introduction of 13-layered BP film, the loss profile provides a moderate peak loss. Apart from the loss curve, the respective sensitivity of the 1.38 RI analyte is also measured at every embraced BP layer combination (11–15 layers at a period of 1). The observation displays that the sensitivity keeps on increasing and increases up to 13 layers of BP but starts decreasing afterward (beyond 13 layers); hence, any random number of BP layers could not be taken [39]. The variation in the sensitivity with definite layers of BP is exhibited in Fig. 4b. Due to moderate loss peak and enhanced wavelength spectral sensitivity, 13 layers of BP are considered as optimized. The fabrication of BP layers is already covered in the introduction.

### Performance with Different Types of Air Holes ($r_1$ and $r_2$)

After optimizing the plasmonic elements such as the widths of the gold layer, $TiO_2$ layer, and BP layers, two kinds of air holes that are used along the fiber length are chosen for optimizing the suggested D-shaped SPR-PCF sensor. The radius of both the holes is varied to optimize the overall structure's performance. The effect of smaller air holes is more prominent than bigger air holes because the mode field of the core directly interacts with the smaller air holes. To investigate the prominent variation, the smaller air hole radius is changed from 0.25 to 0.35 $\mu$m in the interval of 0.05 $\mu$m, and the corresponding confinement peak losses are calculated as 83.06 dB/cm, 40.86 dB/cm, and 18.63 dB/cm, respectively, shown in Fig. 5a. At 0.25 µm radius of a small air hole, the peak loss is very high because the core's modal field gets enough and maximized room to interact with plasmonic material system, consequently increasing the loss, whereas vice-versa condition occurs at 0.35 µm radius; therefore, the peak loss is very low. On contrasting the peak loss magnitudes at different small air hole radius parameters, the analysis shows that 0.3 $\mu$m radius of small air holes generates modest peak loss; therefore, it is considered as optimized magnitude for small air hole radius. The next observation discusses the influence of a bigger air hole in calculating the confinement losses and respective sensitivities. Changing the diameter of a bigger air hole from 0.9 to 1.1 $\mu$m in the step size of 0.1 $\mu$m shows the confinement loss of 49.52 dB/cm, 40.86 dB/cm, and 33.16 dB/cm, respectively, in Fig. 5b. The reason for peak loss variation at different radii of bigger air holes almost coincides with small air hole radius. The inflated radius blocks the passage for the core modal field to interlink with the plasmonic material, ultimately decreasing the loss. On comparing the peak losses and sensitivities, the investigation shows that 1 $\mu$m is the optimized radius for bigger air holes.

The "Investigation with Different Parameters of Proposed Sensor" section has concluded clearly about the final optimized structural parameters. The parameters are tabulated in Table 1.

## Final Optimized Structures

After optimizing the structural parameters given, the thickness of gold, $TiO_2$, number of layers of black phosphorus, and radius of big and small air holes are 35 nm, 14 nm, 13 layers (each of 0.53 nm), 1 $\mu$m, and 0.3 $\mu$m, respectively.

**Table 1** Optimized structural parameters

| | Thickness of gold ($t_{Au}$) | Thickness of $TiO_2$ ($t_{Ti}$) | Number of BP layers | Radius of small air hole ($r_2$) | Radius of big air hole ($r_1$) |
|---|---|---|---|---|---|
| Optimized parameter | 35 nm | 14 nm | 13 layers | 0.3 $\mu$m | 1$\mu$m |

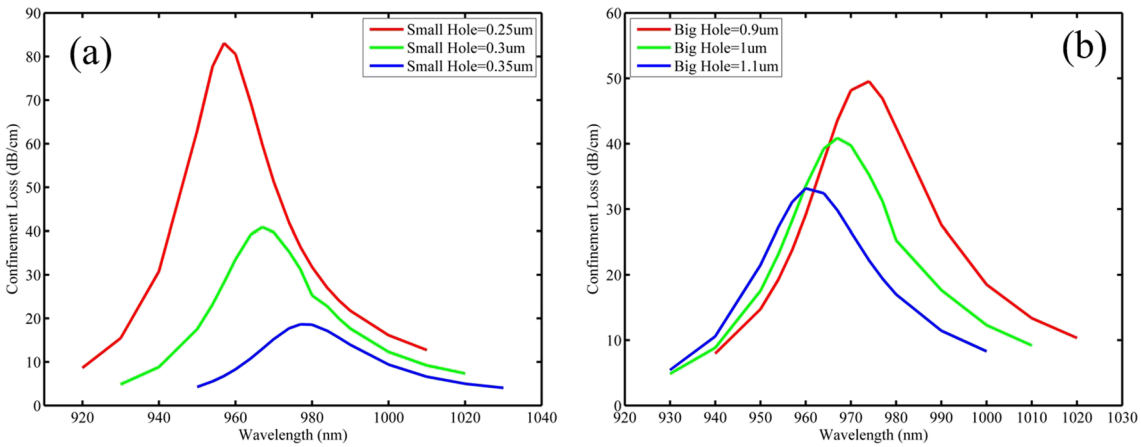



**Fig. 5** Shifts in loss peak at $n_a = 1.38$ **a** with different small air hole radius ($r_2$) and **b** with different big air hole radius ($r_1$)

At the final stage of the study, the confinement loss is calculated by setting the parameters at their optimized magnitudes for a range of refractive index–based analytes ranging from 1.3 to 1.4 RI. The study provides the spectral confinement loss profiles as shown in Fig. 6, for different RIs of analyte ranging from 1.3 to 1.4 with a step size of 0.01. Black phosphorus is proven to increase the sensitivity of the sensor for the proposed structure, and the dynamic sensitivity without depositing black phosphorus layers came as 9000 nm/RIU and, after incorporating BP layers, is 23,300 RIU/nm which is higher than the results reported in previous studies [30]. Figure 6a, b shows the loss curves and amplitude sensitivity respectively for the proposed PCF sensor without black phosphorus and Fig. 6c, d for the sensor with black phosphorus.

The performance of the proposed sensor is tabulated in Table 2, which shows the clear difference in the wavelength and amplitude sensitivity in the refractive index range from 1.3 to 1.4, when there is no black phosphorus and when the BP layer is incorporated into the structure. After the addition of black phosphorus, the sensitivities are constantly increased by an average of 50%. The variation of resonant wavelength and analyte's refractive index are plotted in a graph (Fig. 7) in which a fourth-order polynomial is fitted with coefficients that are tabulated in Table 3. The *R*-square value of the fitted curve approximately equals unity for both analyses (with BP and without BP).

## Possibilities of Fabrication and Tolerance of Proposed Sensor

Recently, PCF structures have become convenient as fabrications of these have been discovered. The experimental realization of PCF-SPR is briefly discussed in various literatures [40]. The gold layer fabrication on the flat surface of the sensor can be achieved by numerous methods such as pulse laser deposition, RF spluttering, and splash technology [40], but here we have used $TiO_2$ as an adhesive between the structure and gold [32]. The layers of black phosphorus are placed above the plasmonic materials. The layers can be prepared by various techniques such as mechanical peeling, liquid peeling, and pulse layer precipitation [30]. Therefore, it is believed that the fabrication of the proposed sensor is possible using current fabrication techniques.

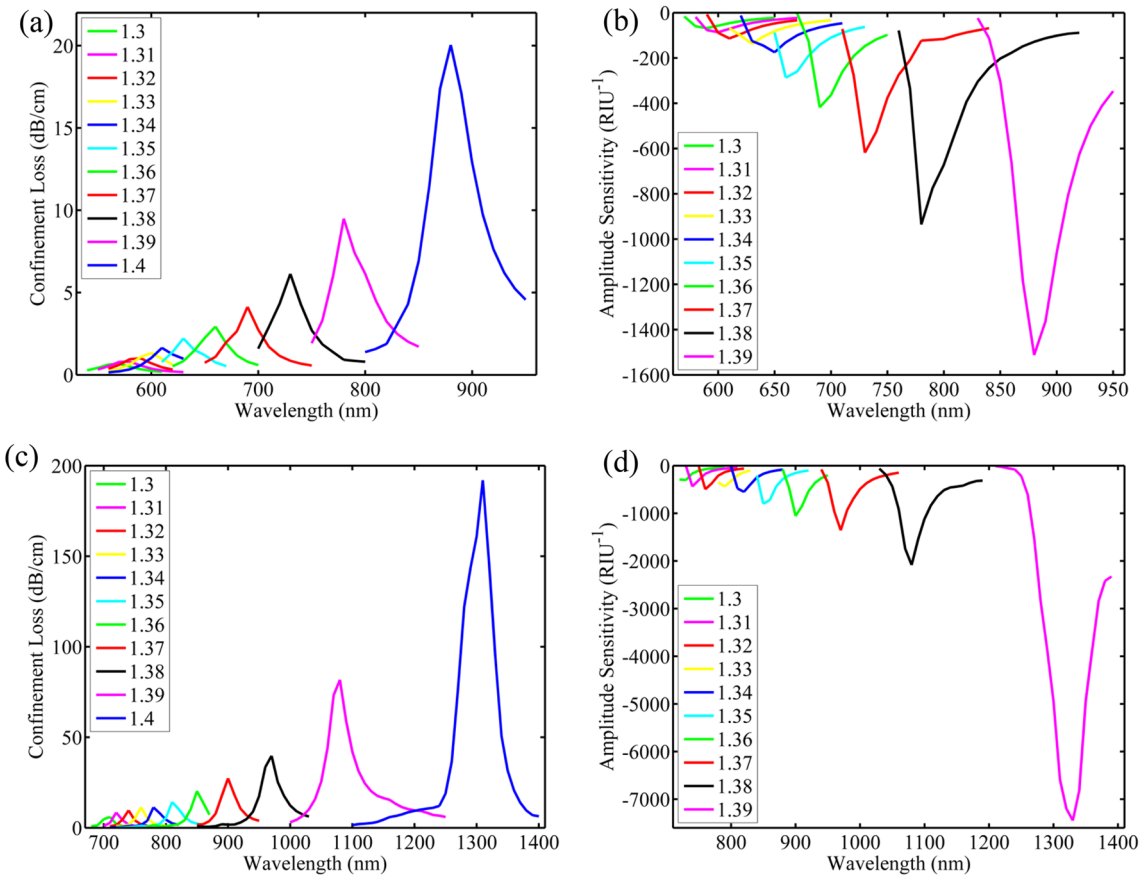



**Fig. 6** **a** Loss curve and **b** amplitude sensitivity for the optimized parameters without black phosphorus. **c** Loss curve and **d** amplitude sensitivity for optimized parameters with 13 layers of black phosphorus

**Table 2** Changes in wavelength and amplitude sensitivity when BP is introduced

| Analyte's RI | Proposed structure without BP | | | Proposed structure with BP | | |
|---|---|---|---|---|---|---|
| | Resonant wavelength (nm) | Wavelength sensitivity (nm/RIU) | Amplitude sensitivity ($RIU^{-1}$) | Resonant wavelength (nm) | Wavelength sensitivity (nm/RIU) | Amplitude sensitivity ($RIU^{-1}$) |
| 1.3 | 570 | 1000 | 68.0059 | 710 | 1000 | 301.294 |
| 1.31 | 580 | 1000 | 86.1046 | 720 | 2000 | 436.857 |
| 1.32 | 590 | 1000 | 113.738 | 740 | 2000 | 495.455 |
| 1.33 | 600 | 1000 | 137.331 | 760 | 2000 | 437.518 |
| 1.34 | 610 | 2000 | 174.83 | 780 | 3000 | 550.098 |
| 1.35 | 630 | 3000 | 289.916 | 810 | 4000 | 799.881 |
| 1.36 | 660 | 3000 | 418.148 | 850 | 5000 | 1055.95 |
| 1.37 | 690 | 4000 | 618.471 | 900 | 7000 | 1357 |
| 1.38 | 730 | 5000 | 935.404 | 970 | 11,000 | 2083.29 |
| 1.39 | 780 | 9000 | 1512.3 | 1077 | 23,300 | 7446.56 |
| 1.4 | 870 | – | – | 1310 | – | – |

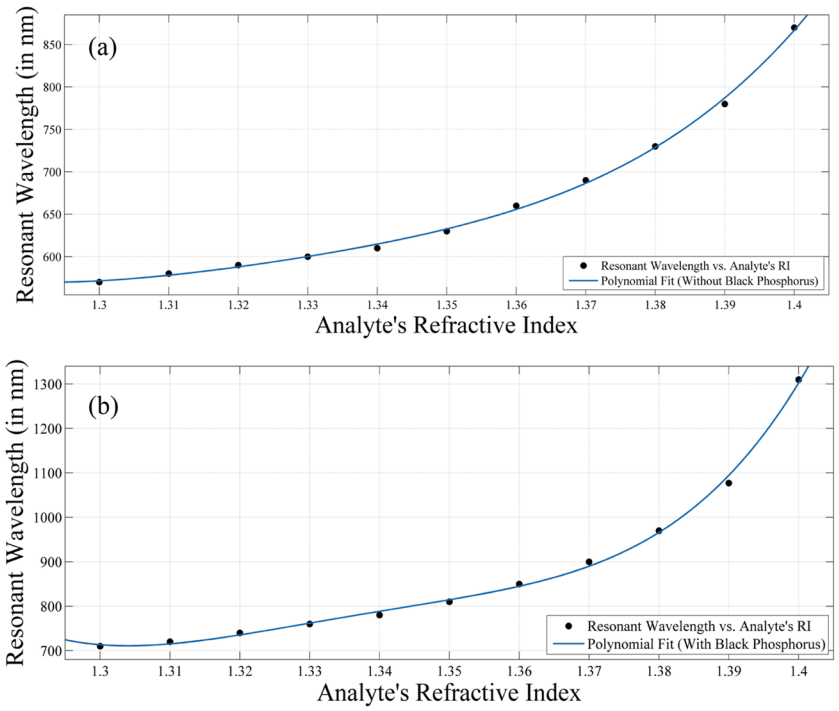



**Fig. 7** Polynomial fitting of resonant wavelength with corresponding RI. **a** Without black phosphorus. **b** With black phosphorus

**Table 3** Coefficients of the fitted polynomial for the curve of resonant wavelength vs analyte's RI

| 4th order polynomial | $p1\times x^4+p2\times x^3+p3\times x^2+p4\times x+p5$ | | |
|---|---|---|---|
| | Without BP | With BP | |
| p1 | 3.788e+06 | 2.325e+07 | Polynomial coefficients |
| p2 | −2.008e+07 | −1.243e+08 | |
| p3 | 3.991e+07 | 2.491e+08 | |
| p4 | −3.527e+07 | −2.219e+08 | |
| p5 | 1.169e+07 | 7.411e+07 | |
| $R^2$ | 0.9985 | 0.9981 | Goodness of fit |

To observe the tolerance in dimensions of structural parameters, they are altered by ± 5% to analyze the change (if any). Initially, the thickness of the gold layer was by ± 5%, and the results show that WS is varied by ~ ± 2% and AS is varied by ~ ± 3%. While changing the thickness of $TiO_2$, there is no change visible in WS or AS, and when $r_1$ is varied WS remains the same but AS changes by ~ ± 1%. And at last, when $r_2$ is altered by ± 5%, WS and AS remain unvaried. This concludes that ± 5% variation in structural parameters does not change WS or AS at larger scales which makes the sensor tolerant to minor

**Table 4** Sensitivity comparison for the proposed SPR-PCF sensor

| Design | Coupling wavelength (nm) | RI range | Dynamic sensitivity (nm/RIU) |
|---|---|---|---|
| Gold and $TiO_2$-based SPR-PCF [22] | 750 | 1.35–1.4 | 10,000 |
| Graphene-based PCF Sensor [41] | 590 | 1.33–1.35 | 2520 |
| AZO coated PCF [42] | 1820 | 1.32–1.34 | 5000 |
| Ultrapolished PCF with BP [10] | 778 | 1.320–1.375 | 8200 |
| Proposed D-shaped PCF (without BP) | **870** | **1.30–1.40** | **9000** |
| Proposed D-shaped PCF (with BP) | **1310** | **1.30–1.40** | **23,300** |

fabrication errors. The literature comparison of the proposed sensor with previous works is presented in Table 4. The comparison shows a very high increase in dynamic sensitivity for the proposed photonic crystal fiber-surface plasmon resonant sensor. It also includes the comparison of the sensor with and without plasmonic material, i.e., black phosphorus, and shows that the sensor becomes more than twice sensitive to the changes in the surrounding environment.

## Conclusion

In summary, the paper presents a D-shaped SPR-PCF refractometric sensor with $TiO_2$, gold, and black phosphorus. Black phosphorus is a 2-D material which is used to increase the sensitivity of the sensor from 9000 to 23,000 nm/RIU, and an amplitude sensitivity of 7446.56 $RIU^{-1}$ is also achieved. The simulations are performed in COMSOL Multiphysics Software which is based on the finite element method. All the parameters are optimized to receive such high sensitivity such as big and small air holes which are optimized to be 1 $\mu$m and 0.3 $\mu$m, respectively. Similarly, 35 nm of the gold layer, 14 nm of $TiO_2$, and 13 layers of black phosphorus are applied to achieve a higher sensitivity. These enhancements make the sensor more appealing to detect even smaller changes in the surrounding environment and are even able to detect a longer refractive index sensing range, i.e., from 1.3 to 1.4. It can therefore be used in various applications such as medical diagnostics and chemical sensing.

**Acknowledgements** Plakshi Gupta, a student of Masters of Science in Physics in the Department of Applied Physics at Delhi Technological University, acknowledges for providing excellent research facilities for conducting research.

**Author Contribution** P.G. 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.

**Data Availability** No datasets were generated or analysed during the current study.

## Declarations

**Competing Interests** The authors declare no competing interests.

## References

1. Ritchie RH (1957) Plasma losses by fast electrons in thin films. Phys Rev 106(5):874
2. Otto A (1968) Excitation of nonradiative surface plasma waves in silver by the method of frustrated total reflection. Zeitschrift für Physik A Hadrons and nuclei 216(4):398–410
3. Kretschmann E, Raether H (1968) Radiative decay of non radiative surface plasmons excited by light. Zeitschrift für Naturforschung A 23(12):2135–2136
4. Liedberg Bo, Nylander C, Lunström I (1983) Surface plasmon resonance for gas detection and biosensing. Sensors Actuators 4:299–304
5. Rifat AA et al (2017) Photonic crystal fiber based plasmonic sensors. Sens Actuators, B Chem 243:311–325
6. Jorgenson RC, Yee SS (1993) A fiber-optic chemical sensor based on surface plasmon resonance. Sens Actuators, B Chem 12(3):213–220
7. Skorobogatiy, Maksim, and Andrei V. Kabashin (2006) Photonic crystal waveguide-based surface plasmon resonance biosensor. Appl Phys Lett 89(14)
8. Singh G, Sharma S, Kumar A (2024) Design and analysis of spider-web photonic crystal fiber in the terahertz regime. Indian J Phys 98(8):2917–2922
9. Singh, J., Khamaru, A., & Kumar, A (2024) A refractive index based cancer cells sensor in terahertz spectrum: design and analysis. J Opt 1–9
10. Singh N, Khamaru A, Kumar A (2024) Design and analysis of a rectangular core refractive index-based PCF sensor for biosensing application. Opt Quant Electron 56(7):1133
11. Meng F, Wang H, Fang D (2022) Research on D-shape open-loop PCF temperature refractive index sensor based on SPR effect. IEEE Photonics J 14(3):1–5
12. Wang G, Lu Y, Duan L, Yao J (2020) A refractive index sensor based on PCF with ultra-wide detection range. IEEE J Sel Top Quantum Electron 27(4):1–8
13. Lu, Y., Tian, F., Chen, Y., Han, Z., Zeng, Z., Liu, C., ... & Zhang, J. (2023). Characteristics of a capillary single core fiber based on SPR for hydraulic pressure sensing. Opt Commun 530: 129125

14. Cennamo N, Arcadio F, Marletta V, Baglio S, Zeni L, Andò B (2020) A magnetic field sensor based on SPR-POF platforms and ferrofluids. IEEE Trans Instrum Meas 70:1–10
15. Tomer, D. S., & Kumar, A. (2024). A graded index hybrid photonic crystal fiber for supercontinuum generation using $AsSe_2$-$As_2S_5$. J Opt: 1–10
16. Khamaru A, Kumar A (2024) $As_{38}Se_{62}$ 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
17. Garg D, Khamaru A, Kumar A (2024) Ge–As–Se–Te chalcogenide based rib-waveguide for highly coherent on-chip mid-infrared supercontinuum generation: design and analysis. Opt Quant Electron 56(10):1643
18. Moghadam MM, Shahshahani F, Yavari MH (2024) Analysis of generation of high-energy dissipative solitons in all-normal-dispersion mode-locked photonic crystal fiber laser. Optik 299:171616
19. Benazza A, Beffara F, Auguste JL, Olivo M, Dinish US, Humbert G (2024) Reliable and easy-to-use SERS spectroscopy probe using a tapered opto-fluidic photonic crystal fiber. Opt Express 32(3):3440–3450
20. Russell, Philip St John (2008) Photonic crystal fibers: basics and applications. Optical fiber telecommunications VA. Academic Press: 485–522
21. Danlard I, Akowuah EK (2020) Assaying with PCF-based SPR refractive index biosensors: from recent configurations to outstanding detection limits. Opt Fiber Technol 54:102083
22. Jain S, Choudhary K, Kumar S (2022) Photonic crystal fiber-based SPR sensor for broad range of refractive index sensing applications. Opt Fiber Technol 73:103030
23. Kumar S, Kumar D (2022) Sensitivity measurement based on the refractive index detection of dual-coated PCF SPR sensor. Mapan 37(2):435–441
24. Hoque AMT, Al-Tabatabaie KF, Ali ME, Butt AM, Mitu SSI, Qureshi KK (2023) U-grooved selectively coated and highly sensitive PCF-SPR sensor for broad range analyte RI detection. IEEE Access 11:74486–74499
25. Dash JN, Jha R (2016) Highly sensitive D shaped PCF sensor based on SPR for near IR". Opt Quant Electron 48(2016):1–7
26. Yuan Y, Yu X, Ouyang Q, Shao Y, Song J, Qu J, Yong KT (2018) Highly anisotropic black phosphorous-graphene hybrid architecture for ultrassensitive plasmonic biosensing: theoretical insight. 2D Materials 5(2):025015
27. Vacacela Gomez C, Pisarra M, Gravina M, Riccardi P, Sindona A (2017) Plasmon properties and hybridization effects in silicene. Phys Rev B 95(8):085419
28. Dávila ME, Xian L, Cahangirov S, Rubio A, Le Lay G (2014) Germanene: a novel two-dimensional germanium allotrope akin to graphene and silicene. New J Phys 16(9):095002
29. Qiao J, Kong X, Hu ZX, Yang F, Ji W (2014) High-mobility transport anisotropy and linear dichroism in few-layer black phosphorus. Nat Commun 5(1):4475
30. Tong K et al (2021) Study on increasing measurement range and enhancing sensitivity of PCF surface-plasmon-resonance biosensor using black phosphorus. J Russ Laser Res 42:283–291
31. Xing C, Zhang J, Jing J, Li J, Shi F (2019) Preparations, properties and applications of low-dimensional black phosphorus. Chem Eng J 370:120–135
32. U. Ramani, H. Kumar, R. Kumar, B. K. Singh, and P. C. Pandey (2024) Dual- core photonic crystal fiber–based plasmonic sensor for a broad range of refractive index sensing. Plasmonics 1–11
33. Zuhayer A, Abd-Elnaby M, Ahammad SH, Eid MM, Sorathiya V, Rashed ANZ (2022) A gold-plated twin core d-formed photonic crystal fiber (PCF) for ultrahigh sensitive applications based on surface plasmon resonance (SPR) approach. Plasmonics 17(5):2089–2101
34. Singh S, Prajapati YK (2020) Dual-polarized ultrahigh sensitive gold/$MoS_2$/graphene based D-shaped PCF refractive index sensor in visible to near-IR region. Opt Quant Electron 52(1):17
35. Singh Y, Raghuwanshi SK (2021) Titanium dioxide ($TiO_2$) coated optical fiber-based SPR sensor in near-infrared region with bimetallic structure for enhanced sensitivity. Optik 226:165842
36 Lee Seong-Yeon, Yee Ki-Ju (2021) Black phosphorus phase retarder based on anisotropic refractive index dispersion. 2D Materials 9(1):015020
37. Liu C et al (2018) Symmetrical dual D-shape photonic crystal fibers for surface plasmon resonance sensing. Opt Express 26(7):9039–9049
38. Singh S, Prajapati YK (2020) TiO2/gold-graphene hybrid solid core SPR based PCF RI sensor for sensitivity enhancement. Optik 224:165525
39. Singh Y, Paswan MK, Raghuwanshi SK (2021) Sensitivity enhancement of SPR sensor with the black phosphorus and graphene with bi-layer of gold for chemical sensing. Plasmonics 16:1781–1790
40. Chen Y et al (2016) Experimental realization of D-shaped photonic crystal fiber SPR sensor. J Phys D Appl Phys 50(2):025101
41. Yang X, Lu Y, Liu B, Yao J (2017) Analysis of graphene-based photonic crystal fiber sensor using birefringence and surface plasmon resonance. Plasmonics 12:489–496
42. Dash JN, Das R, Jha R (2018) AZO coated microchannel incorporated PCF-based SPR sensor: a numerical analysis. IEEE Photonics Technol Lett 30(11):1032–1035


**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.

# Inclusive and supportive aging: mental health and social care strategies for India's elderly

Neha Narula

Neha Narula is based at the Department of USME, Delhi Technological University, Delhi, India.

**Abstract**

**Purpose** – *As population aging accelerates globally, understanding the multifaceted needs of elderly individuals has become increasingly crucial. This paper aims to bridge significant research gaps in elderly well-being, particularly in the Indian context, by conducting a comprehensive bibliometric analysis and systematic literature review. While existing studies on aging primarily emphasize financial aspects like pensions and health insurance, this review identifies other essential yet underexplored dimensions impacting elderly well-being, such as emotional support, social connections, financial security beyond traditional income sources and spiritual well-being.*

**Design/methodology/approach** – *The data methodology for this paper involves a comprehensive bibliometric analysis and systematic literature review conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. The authors identified recurring themes, research gaps and emerging areas within the literature through thematic analysis. The combined insights from the bibliometric analysis and systematic literature review inform the development of a research agenda.*

**Findings** – *The findings of this study suggest that family and social support play an integral role in reducing loneliness, anxiety and depression among the elderly. However, urbanization and migration trends are diminishing traditional support structures. Financial security remains a concern, as many elderly individuals lack access to financial literacy resources and emergency funds, impacting their confidence and autonomy. Additionally, spiritual practices, which have shown positive correlations with mental health and cognitive function, are rarely incorporated into existing health-care programs for the elderly.*

**Originality/value** – *This paper provides several policy recommendations, advocating for integrating community-based social programs, financial literacy initiatives and culturally sensitive spiritual support in elderly care frameworks. These measures foster the aging population's emotional resilience, financial independence and holistic health. Moreover, this study highlights the importance of future research focusing on specific subgroups within the elderly population, including differences by geographic location, educational background and socio-economic status. Ultimately, this research supports the development of inclusive and sustainable policies that promote health, independence and quality of life for the aging population.*

**Keywords** *Population aging, Mental health, Health of elderly, Research agenda, Systematic literature review*

**Paper type** *Conceptual paper*

## Introduction

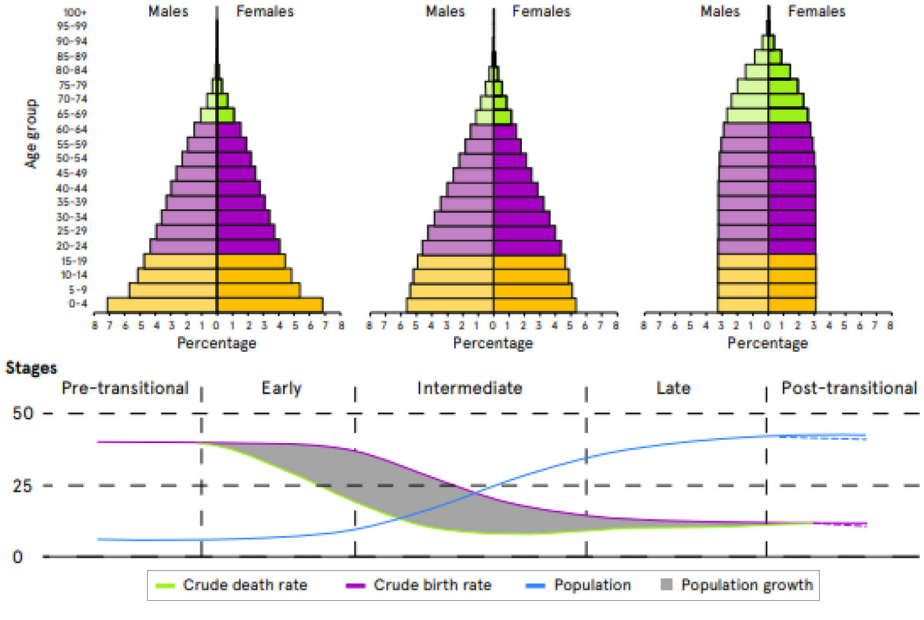
Population aging, marked by increased life expectancy and declining birth rates, is transforming societies worldwide (Norman, 2020). As the proportion of elderly individuals rises, there are profound implications across economic, social and health-care systems (Ogugua *et al.*, 2024). This demographic shift affects labor markets, strains pension systems and drives demand for healthcare services tailored to age-related needs (Fang *et al.*, 2023). Addressing

these challenges requires a multidimensional approach, combining insights from economics, sociology and health-care disciplines (Schneider and Blyth, 2017). By 2030, one in six people globally will be 60 years or older, with the population in this age group expected to grow from 1 billion in 2020 to 1.4 billion. This rapid demographic shift is projected to continue, with the world's population aged 60 years and above doubling to 2.1 billion by 2050. Additionally, the number of individuals aged 80 years or older is set to triple, reaching 426 million by 2050 (WHO, 2024).

In Figure 1, before the demographic transition, age and sex distribution resembles a pyramid, with a wide base representing a high proportion of young people and a narrow top with fewer elderly individuals. Initially, as mortality rates drop – especially among infants and children because of improved nutrition and better control of infectious diseases like pneumonia, diarrhea and malaria, this leads to "population rejuvenation," with an expanding base as the share of children and youth rises. In the intermediate stage, fertility rates decline, and the pyramid's base narrows as the population ages. Adult survival rates improve, expanding the share of people in working and childbearing ages and marking the onset of population aging. Eventually, the older population grows significantly. If fertility stabilizes at the replacement level, then the age distribution shifts from a pyramid shape to a more rectangular or bell-shaped structure, with relatively similar numbers across age groups, resembling a bell tower or rocket.

With research increasingly focusing on population aging, especially since the World Health Organization designated this decade as the "UN Decade of Healthy Ageing," there is a growing need to understand global research trends. Recognizing this shift, the author examined research output across different countries to gain an initial outlook on current efforts and gaps. By analyzing research trends in leading and emerging countries, the aim was to identify areas where further studies are essential, particularly in regions where population aging poses significant socioeconomic challenges especially on there mental health. This initial assessment serves as a foundation to highlight global research needs

**Figure 1** Population distribution by age and sex, crude birth and death rates and total population size at different stages of the demographic transition




**Source:** World social report 2023; leaving no one behind in an ageing world

and guide future work towards developing effective policies and practices that support aging populations.

Figure 2 highlights significant research activity across various countries, with China, the USA and England leading in output, indicating a strong focus on the given research field. These countries are followed by Germany, The Netherlands, Sweden and Australia, which show moderate research activity, while Norway, Canada and Switzerland display relatively lower levels. The high research output in China and the USA reflects substantial investment and institutional support, allowing for significant advancements. For India, the chart reveals a research gap in this field, suggesting the need to strengthen its research capacity to align with international efforts. By increasing funding, fostering academic collaborations and investing in infrastructure, India could enhance its research presence and contribute to regional and global progress.

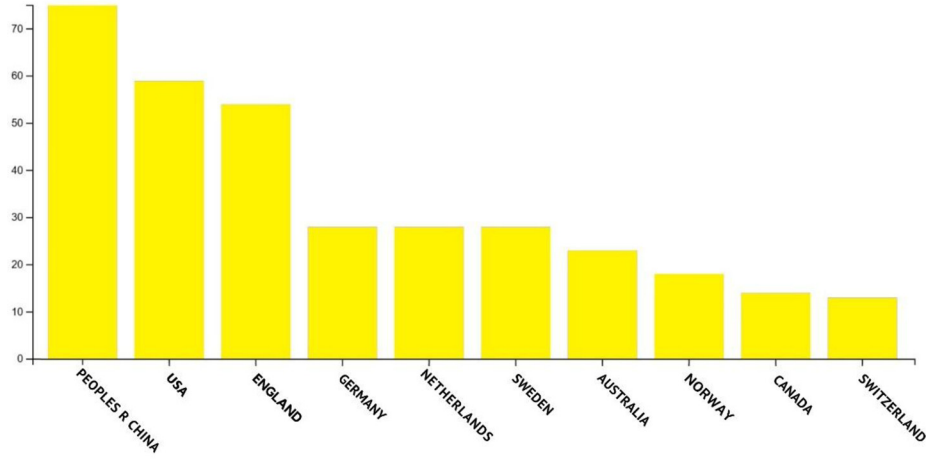
## Data methodology

The data methodology for this paper involves a comprehensive bibliometric analysis and systematic literature review (SLR) conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. This structured approach ensures a rigorous, transparent and replicable methodology, enabling us to analyze and synthesize existing research on population aging.

We used databases such as Web of Science to identify relevant peer-reviewed publications on population aging for the bibliometric analysis. Search terms were carefully selected to capture various studies addressing aging populations' economic, mental health, social inclusion and health-care implications. Data were extracted on publication trends, influential authors, citation networks and geographical distribution to map the landscape of research activity across different countries.

PRISMA guidelines were applied in the SLR phase to screen and select studies systematically. This process involved removing duplicates, assessing eligibility and critically appraising selected articles. We identified recurring themes, research gaps and emerging areas within the literature through thematic analysis. The combined insights from the bibliometric analysis and SLR inform the development of a research agenda,

**Figure 2** Significant research activity across various countries




**Source:** Author's own calculations

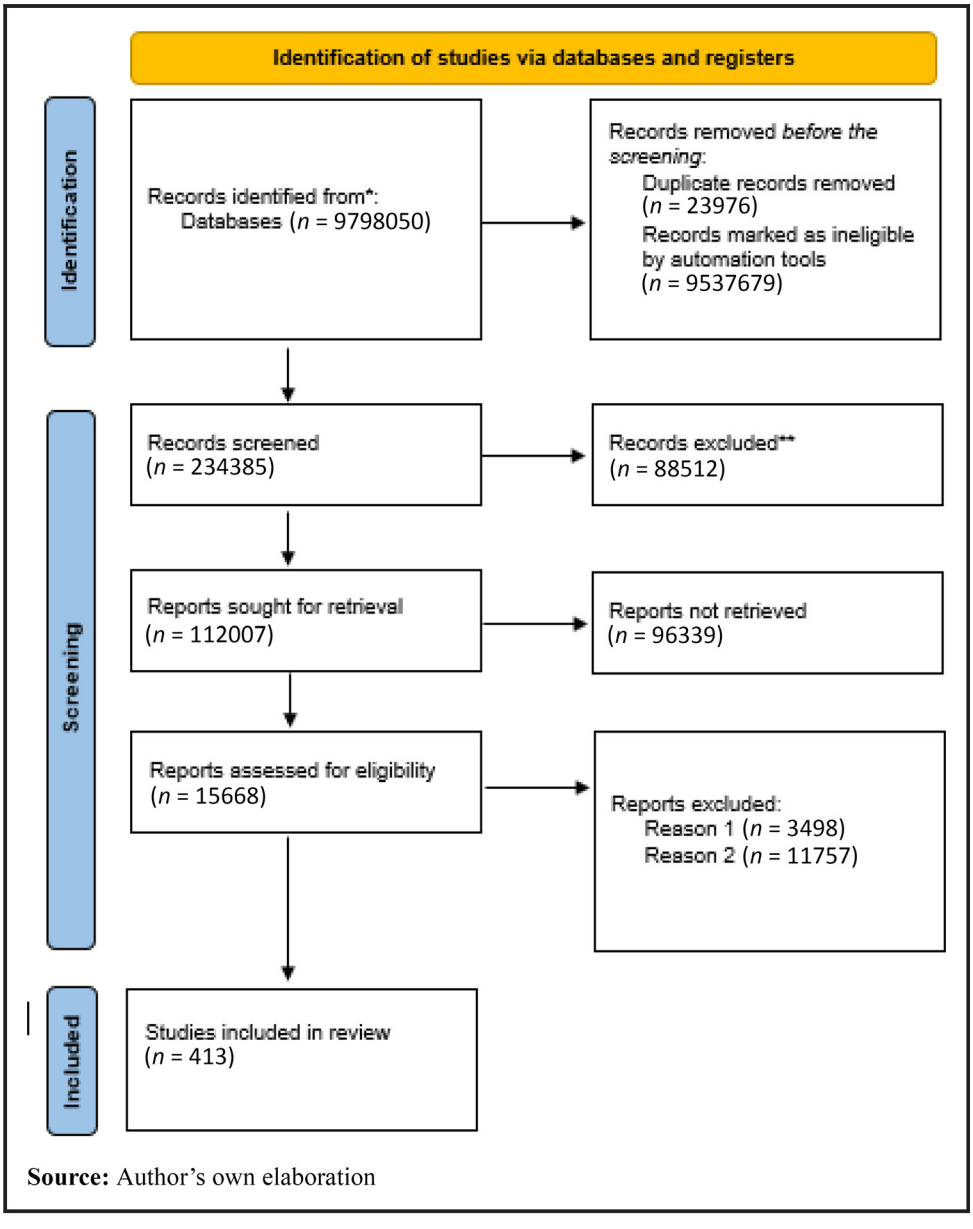
highlighting key areas where further investigation is needed to address challenges associated with population aging globally, with a particular focus on regions with limited existing research.

## Results

### *Current research trends*

This review aimed to gather insights on population research trends in India, identifying the current focus areas and potential directions for future studies. Our initial search yielded 9,796,050 research articles across multiple databases (Figure 3). To refine this pool, we applied several filters using the Web of Science database: We included only open-access articles published in English within the past five years, narrowing the results

**Figure 3** Selection of articles based on PRISMA guidelines




**Source:** Author's own elaboration

to 48,017 studies, ensuring recent and accessible research relevant to India. Further screening and eligibility assessments reduced this collection to 15,668 articles, with each study evaluated for its contribution to the field of population research in the Indian context. After applying our final selection criteria, 413 articles were chosen for the analysis, representing the most pertinent and high-quality research on this topic. The final selection of studies highlights critical themes in population research, including health care, mental health, aging, economic factors and demographic trends specific to India.

### *Top publishers*

Based on Figure 4, the top publishers actively researching population aging include Sage, Springer Nature, Wiley, Taylor & Francis and Elsevier. Sage and Springer Nature lead the group, publishing the most articles on population aging, likely because of their extensive focus on social sciences and health care, which are crucial domains for aging studies. Wiley follows closely, with a similar commitment to disseminating research across multidisciplinary areas, including medicine, public health and demographics. Taylor & Francis and Elsevier also contribute significantly to the field, emphasizing research on demographic changes, health policy and gerontology. The focus of these publishers aligns with the broad implications of an aging population on health-care systems and economic stability, topics of particular relevance in population aging studies.

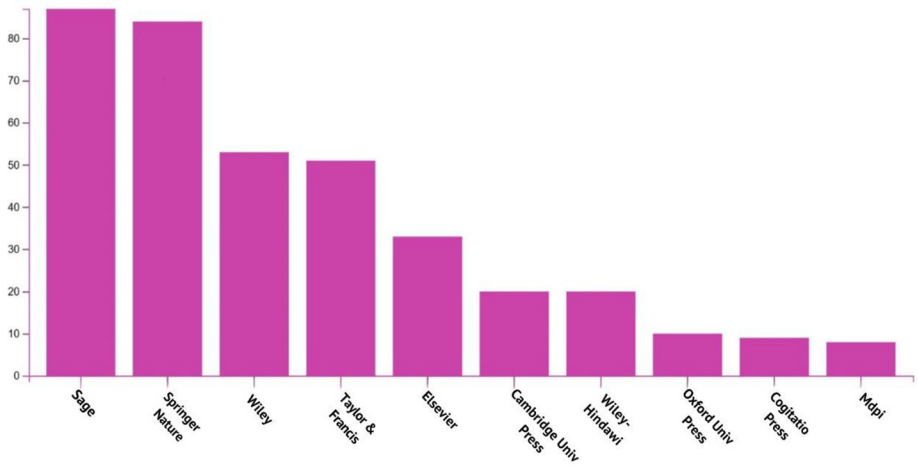
### *Key themes*

Figure 5 represents the network of keywords related to population aging, with nodes representing key terms and the connections indicating co-occurrence and thematic relationships. The size of each node reflects the frequency of keyword usage, while color-coded clusters signify thematic groups.

Key insights include:

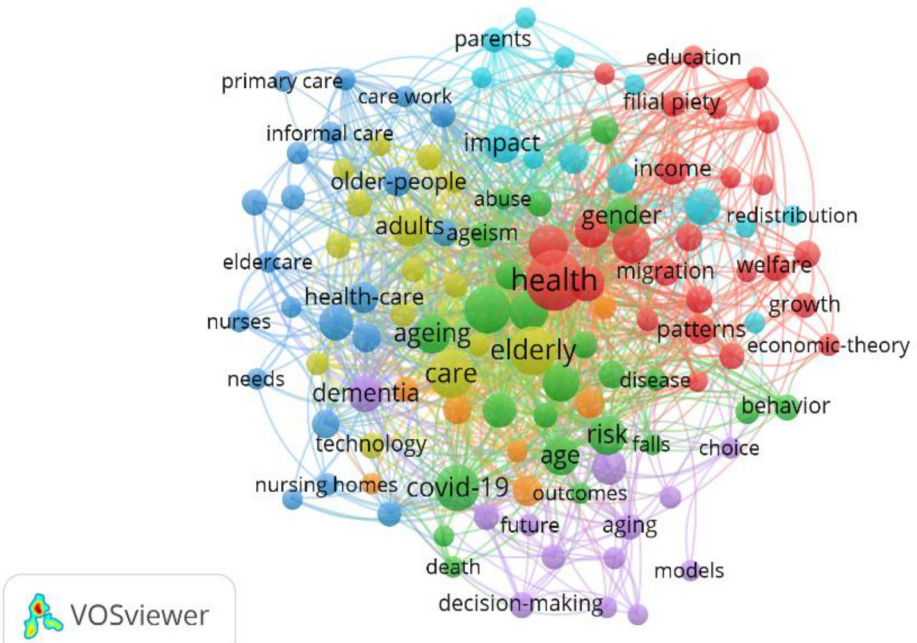
1. *Central themes*: Terms like “health,” “elderly,” “care,” “aging,” “mental health” and “dementia” are prominently displayed in the center with larger nodes, indicating a significant research focus on health and care needs for the elderly population.

**Figure 4** Top publishers




**Source:** Author’s own elaboration

**Figure 5** Recurring keywords




**Source:** Author's own elaboration

2. *Clusters*:
   - *Red cluster*: It focuses on socio-economic themes, including "income," "redistribution," "gender," "migration," "education" and "welfare." This cluster suggests research on the economic and social impacts of aging.
   - Green cluster: It centers on health risks, with keywords like "risk," "outcomes," "aging," "behaviour" and "falls." This area emphasizes the physical and behavioral health outcomes associated with aging.
   - *Blue cluster*: It includes "primary care," "informal care," "elders" and "nurses," highlighting the roles of health-care services and informal caregiving for elderly populations.

### *Existing research gaps*

In recent years, population aging has been extensively studied, primarily focusing on pensions, health insurance and other financial aspects. However, less attention has been given to various other factors that significantly influence the well-being of elderly individuals, including emotional, social and financial confidence and the role of spiritual well-being. This research aims to conduct an in-depth literature review to identify gaps and highlight potential directions for future research in these underexplored areas. Through a systematic literature review of numerous relevant studies, we analyze major findings and pinpoint emerging research gaps, offering insights that could guide future studies to address the comprehensive well-being of aging populations better.

Table 1 summarizes the findings from our extended literature review, detailing the objectives, key findings and primary research gaps identified in various studies. This table

**Table 1** The findings from our extended literature review, detailing the objectives, key findings and primary research gaps identified in various studies

| *Title* | *Objectives* | *Mazor findings* | *Research gaps* |
|---|---|---|---|
| Elderly and health: Role of spirituality in Indian context (Naik *et al.*, 2019) | To examine the association between the frequency of religious and spiritual practices and mental health outcomes, such as depression and cognitive function, while controlling for socioeconomic and health factors | The paper discusses the role of spirituality in the health and well-being of the elderly population in India | The role of spirituality in elderly health care has not been adequately explored or included in existing programs and policies |
| Population Ageing in India: Studies on Ageing in India: A Review (Raju, 2014) | Analyze the economic, social, psychological and health aspects of the elderly population<br>Identify issues and factors that influence models of care for the elderly | The paper reviews the status of research on aging in India, highlighting gaps in understanding the psychological, financial and health aspects of the elderly, including the role of spirituality | The national program for health care of the elderly (NPHCE) primarily focuses on physical, mental and social well-being but does not explicitly address the role of spirituality |
| Does the intensity of religiosity and spirituality in later life improve mental well-being? Evidence from India (Barman *et al.*, 2022) | Highlight gaps in research, particularly on the psychological aspects of aging in India | Those who practiced spirituality almost daily were 20% less likely to be depressed compared to their counterparts<br>The likelihood of reported low cognitive functional health was lower for those who almost daily practiced spirituality (OR = 0.82) and religiosity (OR = 0.55) than those who never practiced | Various social, financial and access-related challenges in the health care of the elderly could be potential areas for further research |
| Spirituality as a Factor in Positive Ageing (P and S, 2017) | To establish spirituality as a critical factor in developing a positive outlook on aging | Spirituality is critical in developing a positive outlook on aging in the Indian context | Lack of comprehensive, large-scale studies on aging in India, with many existing studies being limited in scope and sample size |
| High Time to Observe on Physical and Emotional Wellbeing of Elderlies: A Narrative Review (Kaur *et al.*, 2021) | To assess the overall physical and emotional well-being of the elderly population | The review highlights the need for evidence-based interventions to address the physical and emotional well-being challenges faced by the elderly population in India | Lack of research on the psychological, social and health-related aspects of aging in India, particularly in understanding the needs and experiences of different elderly subgroups |
| Educational Level and Financial Status: Decisive Factors for Level of Spirituality and Psychological Well-being among Elderly (Singh and Mehra, 2020) | To assess the psychological well-being among institutionalized and non-institutionalized elderly | Educational level and financial status are decisive factors for spirituality and psychological well-being among the elderly in India | Need for more holistic, multidisciplinary research on aging in India that views the elderly as active participants in society rather than just passive recipients of care |
| Impact of Spirituality on Well-Being among Old Age People (Saleem and Khan, 2015) | To discuss the link between spirituality, positive attitude and positive aging | Both a sense of purposelessness and maintenance of discipline, as dimensions of spirituality, were positively correlated with overall well-being | The need for evidence-based interventions to support the physical and emotional well-being of the elderly population |

(*continued*)

**Table 1**

| *Title* | *Objectives* | *Mazor findings* | *Research gaps* |
|---|---|---|---|
| Regional Variation and Determinants of Well-being of the Elderly in India (Das *et al.*, 2018) | To construct indices for physical, social and emotional well-being using various data sources | The study emphasizes the need for comprehensive caregiver education and the incorporation of culturally aligned religious-spiritual education to improve the well-being of the elderly | The challenges faced by elderly individuals, both those living in old age homes and those living independently, as well as those with medical conditions, in maintaining their physical and emotional well-being |
| The Effect of Religious Spiritual Education and Care on the Quality of Life among the Elderly: A Review (Yahya *et al.*, 2024) | Evaluate the impact of religious-spiritual education on quality of life for Iranian elders | The growing elderly population in India will lead to new medical, social and economic challenges that need to be addressed by policymakers and program managers | Need for more clearly defined research concepts and a better understanding of how cultural variations impact the role of spirituality in elderly care |
| Geriatric Health in India: Concerns and Solutions (Nath and Ingle, 2008) | Discuss the significance of emotional stability for organizational success in Malaysia's evolving industrial context | The paper discusses the need to address the medical and socio-economic problems faced by the growing elderly population in India | Ineffectiveness of current educational programs in addressing the obstacles to providing spiritual care in elderly care settings |

**Note:** This table offers a structured overview of each study's focus, highlighting the main conclusions and the specific areas where further research is needed, as noted by the authors
**Source**: Author's own work

offers a structured overview of each study's focus, highlighting the main conclusions and the specific areas where further research is needed, as noted by the authors.

### *Identified research questions*

- *Research focus on spirituality and well-being*: Given the limited exploration of spirituality in elderly health programs, future research should investigate how spirituality contributes to mental and emotional well-being:

  *RQ1*. How does regular engagement in spiritual practices affect mental health outcomes such as depression and cognitive function among the elderly?

- *Socioeconomic determinants of elderly care*: Financial and educational factors are crucial to elderly well-being. Expanding research into these determinants can offer insights into holistic elderly support systems:

  *RQ2*. In what ways do socioeconomic and educational disparities impact the mental health and inclusion of institutionalized versus non-institutionalized elderly individuals?

- *Development of comprehensive elderly care programs*: Many studies call for the integration of culturally tailored spiritual education and caregiver training into elderly care. This focus could support the physical, social and emotional aspects of elderly health:

  *RQ3*. How do the psychological and social needs of elderly individuals differ across various subgroups (e.g. location, health status and living arrangements)?

- *Social support and elderly mental well-being*: Social connections with family and friends are crucial for the psychological health of elderly individuals. However, the specific ways these relationships impact emotional and mental well-being require further exploration.

*RQ4*. How do regular interactions with family and friends contribute to reducing loneliness, anxiety and depression among elderly individuals in urban and rural settings?

- *Beyond pension and insurance*: Financial security for the elderly: Financial confidence extends beyond pension and insurance, encompassing day-to-day financial autonomy and planning for unexpected expenses. Exploring alternative aspects of financial security can offer valuable insights for policy and program development:

  *RQ5*. What are the key factors influencing financial confidence among elderly individuals, particularly those without access to pensions or insurance?

## Discussion

Our findings align with numerous global studies that report similar results. These studies corroborate our conclusions and propose comparable policy recommendations and solutions, reinforcing the broader applicability and relevance of the observed patterns and suggested interventions. Sporting memories work is an effective and flexible means of engaging older people to improve their social inclusion and mental well-being. Sporting memory work can be readily deployed in various care and community settings (Clark *et al.*, 2015). In a similar literature review, four key social factors were identified as relevant for determining outcome measures for effective and sustainable integrated community-based care: neighborhood social cohesion, social capital, neighborhood environment and social support. The identified social factors were found to be significantly associated with the mental health of elderly community-dwelling individuals, suggesting that focusing on these social aspects may be important for promoting mental health in the development of effective community-based care (Imaiso, 2020). In a cross-sectional survey of 380 older adults (aged 60+) living in Hamadan, Iran, the GHQ-28 questionnaire was used to assess mental health. Additional social and demographic data was collected. The elderly participants generally had good mental health. Social participation was significantly associated with better mental health overall. The positive relationship between social participation and mental health was found explicitly for the somatic and social performance subscales. Social inclusion is important for improving mental health outcomes for people with mental health conditions (Boardman *et al.*, 2022). In a similar study, the review found a positive association between factors and interventions in the built environment and mental health, well-being, social inclusion and participation in older people. The strength of this association varied depending on the study design, sample size, measurement methods and quality of the primary studies included. The review concludes that acting on the built environment positively impacts mental health and social inclusion in older people (Bonaccorsi *et al.*, 2022). Being socially included was associated with having social roles and responsibilities, such as employment, participating in social activities, environmental work and voting. Many papers provided only simplistic accounts or definitions of social inclusion, despite some engaging in critical discussion. The review identified the relationship between social exclusion, inequality and injustice and questioned the relevance of the concept of social inclusion to current and future mental health policy (Wright and Stickley, 2012). A similar literature review suggests that music and singing, intergenerational interventions, art and culture and multi-activity programs were associated with positive impacts on the health of older adults, including improvements in depression, well-being, subjective health, quality of life, perceived stress, mental health and physical health. Qualitative studies provided insights into potential mediating factors, such as improved self-esteem that may explain how these interventions led to the observed health benefits (Ronzi *et al.*, 2018). A study conducted in India using LASI data suggests that higher levels of social exclusion, mainly exclusion from civic/social activities and services, were associated with a higher prevalence of depressive symptoms among older adults in India. Older adults with higher levels of

overall social exclusion were also more likely to have depressive symptoms (Ronzi *et al.*, 2018).

## Conclusion

This paper examined the multifaceted dimensions of population aging, identifying key research gaps and providing a roadmap for future studies to address the well-being of elderly populations, particularly in the Indian context. Through an extensive literature review and bibliometric analysis, we highlighted that while research on population aging has made strides in areas like pension and health insurance, significant gaps remain in understanding other critical factors that contribute to elderly well-being. These include emotional and social support, financial confidence beyond traditional income sources and the role of spirituality in fostering positive mental health among the elderly.

One primary insight from this research is the crucial role that social support from family and friends plays in the psychological and emotional well-being of elderly individuals. While India and similar societies traditionally have strong family structures, rapid urbanization and migration are gradually eroding these support systems. Elderly individuals often experience isolation, loneliness and anxiety because of weakened social connections. Future research should further examine how social networks can be reinforced or reimagined to support elderly well-being, particularly in regions where families may be less available for regular support.

Additionally, financial confidence among the elderly population goes beyond pensions and health insurance. Many elderly individuals may lack sufficient financial literacy or resources for handling unforeseen expenses, which can impact their sense of security and autonomy. Research findings suggest that financial literacy and planning resources should be expanded to empower the elderly to manage their finances with confidence. Policies aimed at promoting financial education for seniors, incentivizing alternative savings plans and establishing emergency funds for non-pensioned elderly could help address this need. Despite its importance, spirituality is often overlooked in health-care programs aimed at elderly populations. Integrating culturally sensitive spiritual support into elder care policies and programs could be beneficial. For instance, incorporating spiritual activities in care homes or offering counselling that includes spiritual aspects can provide a more holistic approach to elderly care.

This paper identifies several areas for further exploration in terms of future research. First, there is a need for longitudinal studies that track how social support networks, financial preparedness and spiritual practices impact elderly well-being over time. Second, more research is needed on the unique needs of elderly subgroups, such as those living in urban versus rural areas or those with differing educational backgrounds.

### *Policy recommendations*

Based on the findings, this paper offers several policy recommendations to enhance elderly well-being. Policymakers should prioritize community-based social programs, financial literacy initiatives and the inclusion of spirituality in elderly care frameworks. Social programs can include accessible senior centers and digital platforms that facilitate interactions with family and friends, while financial programs should focus on training elderly individuals in managing finances beyond pensions. Finally, introducing spiritual support services within health-care and social programs can address psychological needs that are often neglected.

In summary, while research on aging has made significant progress, there is a need for a more comprehensive approach that encompasses social, financial and spiritual dimensions of elderly well-being. This paper lays the groundwork for future research

and policy development aimed at creating a more inclusive, supportive environment for the aging population, ultimately contributing to healthier and more fulfilling lives for elderly individuals.

## References


Barman, P., Saha, S., Mukhopadhyay, S., Misra, R., Chakraborty, A. and Bhattacharya, A. (2022), "Does the intensity of religiosity and spirituality in later life improve mental well-being? Evidence from India", *Journal of Religion, Spirituality & Aging*, Vol. 35 No. 4, pp. 455-475, doi: 10.1080/15528030.2022.2100560.

Boardman, J., Killaspy, H. and Mezey, G. (2022), *Social Inclusion and Mental Health*, Cambridge University Press, doi: 10.1017/9781911623601.

Bonaccorsi, G., Martini, A., Giorgi, S., Fossi, M.C. and Bartalucci, C. (2022), "Impact of built environment and neighborhood on promoting mental health, well-being, and social participation in older people: an umbrella review", *Annali di Igiene: medicina Preventiva e di Comunita*, Vol. 35 No. 2, pp. 213-239, doi: 10.7416/ai.2022.2534.

Clark, M., Kearns, D. and Abbott, S. (2015), "Sporting memories and the social inclusion of older people experiencing mental health problems", *Mental Health and Social Inclusion*, Vol. 19 No. 4, pp. 202-211, doi: 10.1108/MHSI-06-2015-0024.

Das, B., Sengupta, R. and Paul, K. (2018), "Regional variation and determinants of well-being of the elderly in India", *Journal of Population and Social Studies*, Vol. 26 No. 3, pp. 219-234, doi: 10.25133/JPSSV26N3.016.

Fang, J., Li, X., Zhao, W. and Chen, S. (2023), "The macroeconomic impact of demographic shifts: aging populations and their socioeconomic consequences", *Law and Economy*, Vol. 2 No. 11, pp. 37-43, doi: 10.56397/LE.2023.11.05.

Imaiso, J. (2020), "Consideration of social aspects and mental health of community-dwelling elderly people: a literature review", *Health*, Vol. 12 No. 5, pp. 486-498, doi: 10.4236/health.2020.125038.

Kaur, L., Singh, M. and Dhillon, R.K. (2021), "High time to observe on physical and emotional wellbeing of elderlies: a narrative review", *Medico Legal Update*, Vol. 21 No. 3, pp. 28-34, doi: 10.37506/MLU.V21I3.2960.

Naik, B., Reddy, M. and Kanungo, S. (2019), "Elderly and health: Role of spirituality in Indian context", *Yoga Mimamsa*, Vol. 51 No. 1, p. 38, doi: 10.4103/YM.YM_21_18.

Nath, A. and Ingle, G. (2008), "Geriatric health in India: Concerns and solutions", *Indian Journal of Community Medicine*, Vol. 33 No. 4, p. 214, doi: 10.4103/0970-0218.43225.

Norman, R.A. (2020), "The aging of the world's population", *CRC Press eBooks*, pp. 1-3, doi: 10.1201/9781003076544-1.

Ogugua, N.J.O., Adeyemi, K. and Balogun, T. (2024), "Demographic shifts and healthcare: a review of aging populations and systemic challenges", *International Journal of Science and Research Archive*, Vol. 11 No. 1, pp. 383-395, doi: 10.30574/IJSRA.2024.11.1.0067.

Raju, S.S. (2014), "Studies on ageing in India: a review", *Cambridge University Press eBooks*, pp. 180-214, doi: 10.1017/CBO9781139683456.009.

Ronzi, S., Orton, L., Pope, D., Valtorta, N.K. and Bruce, N.G. (2018), "What is the impact on health and wellbeing of interventions that foster respect and social inclusion in community-residing older adults? A systematic review of quantitative and qualitative studies", *Systematic Reviews*, Vol. 7 No. 1, doi: 10.1186/s13643-018-0680-2.

Saleem, N.R. and Khan, K.S.A. (2015), "Impact of spirituality on well-being among old age people", *International Journal of Indian Psychology*, Vol. 2 No. 3, doi: 10.25215/0203.039.

Schneider, C.H. and Blyth, F. (2017), "Challenges of integrating evidence into health policy and planning: linking multiple disciplinary approaches", *Public Health Research & Practice*, Vol. 27 No. 2, doi: 10.17061/PHRP2721719.

Singh, R. and Mehra, M. (2020), "Educational level and financial status: decisive factors for level of spirituality and psychological well-being among elderly", *Current Journal of Applied Science and Technology*, pp. 1-10, doi: 10.9734/CJAST/2020/V39I4631181.

Wright, N. and Stickley, T. (2012), "Concepts of social inclusion, exclusion and mental health: a review of the international literature", *Journal of Psychiatric and Mental Health Nursing*, Vol. 20 No. 1, pp. 71-81, doi: 10.1111/j.1365-2850.2012.01889.x.

Yahya, F.B., Ahmed, Z. and Khan, T. (2024), "The effect of religious spiritual education and care on the quality of life among the elderly: a review", *Al-Rafidain Journal of Medical Sciences*, Vol. 6 No. 1, pp. 69-75, doi: 10.54133/AJMS.V6I1.481.

## Further reading

Jain, P. and Mehrotra, S. (2017), "Spirituality as a factor in positive ageing", *International Journal of Indian Psychology*, Vol. 4 No. 2, doi: 10.25215/0402.140.

## Corresponding author

Neha Narula can be contacted at: narulanehaeconomics@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

Trans Indian Inst Met (2025) 78:51
https://doi.org/10.1007/s12666-024-03543-1

ORIGINAL ARTICLE

# Influence of Magnesium Content on Tensile Properties and Deformation Behaviour of Al–Mg–Si Alloy Sheets

**Rakesh Kumar**[1] · **Vijay Gautam**[1]

Received: 16 November 2024 / Accepted: 17 December 2024
© The Indian Institute of Metals - IIM 2024

**Abstract** The present study is based on the measurement and effect of Mg content on the tensile behaviour, formability and microstructural characteristics of two different sheets of Al–0.9 Mg–Si and Al–3.1 Mg–Si alloy. The tensile properties of both the sheets have been evaluated by testing laser-cut tensile specimens according to ASTM E8M standard at room temperature. The experimental results indicate that the Al–3.1 Mg–Si alloy, with higher Mg content, exhibits significantly higher yield and tensile strengths, along with nearly double elongation compared to the Al–0.9 Mg–Si alloy. Also, the stretch forming tests revealed better formability in Al–3.1 Mg–Si alloy sheet. Electron backscatter diffraction analysis depicted that the annealed alloy sheet with higher Mg possessed fine equiaxed grains with uniform distribution of $Mg_2Si$ particles, contributing to its higher yield and tensile strengths. After plastic deformation, both the sheet samples exhibited elongated grains with a slight reduction in grain size. Pole figure maps indicated that the Al–3.1 Mg–Si alloy has a stronger texture than the Al–0.9 Mg–Si alloy. After the tensile deformation, the texture of Al–3.1 Mg–Si alloy transformed into a very sharp and strong suggesting a better deformation behaviour. These findings provide valuable insights into how magnesium content influences the strain distribution during plastic deformation and microstructural changes in Al–Mg–Si alloys.

**Keywords** Al–Mg–Si alloy · Tensile properties · Formability · Forming limit diagram · Electron backscatter diffraction

✉ Vijay Gautam
vijaygautam@dce.ac.in

[1] Department of Mechanical Engineering, Delhi Technological University, Delhi 110042, India

## 1 Introduction

The development of various grades of aluminium alloys by incorporating alloying elements has been analysed by design engineers with respect to the various requirements of properties in the manufacturing of the sheet metal components for automotive and aerospace industries. The incorporation of each constituent in the alloy imparts unique characteristics, and the combination of these constituents has varied effects on its mechanical properties [1]. Al–Si alloy sheets are widely utilized in diverse technical applications due to their favourable attributes, such as low density, high strength and robust corrosion resistance. The addition of Mg into the alloy is an important key to improve its ability to undergo heat treatment and results in a highly respectable technical alloy. The positive impact of incorporating a secondary phase, such as the $Mg_2Si$ precipitate, via a heat treatment, on the mechanical characteristics of the alloy has been extensively reported in many investigations [2–5]. The growing popularity of Al–Mg–Si alloys in the field of lightweight structures has created a higher need for extrusions in complex shapes that exhibit high strength, durable and toughness. Also, these alloys are required to meet precise criteria for tensile properties, while demonstrating advantageous characteristics required for excellent formability [6]. The utilization of the 6xxx-series Al–Mg–Si wrought alloys has experienced a notable rise in popularity within the architectural and structural components, as opposed to the 2xxx and 7xxx Al alloys [7]. Their attractive characteristics, such as combination of ease of welding and forming, cost-effectiveness, and ability to resist stress-corrosion cracking [8–10], are the main factors contributing to this. The incorporation of the $Mg_2Si$ intermetallic compound into the Al alloys yields significant enhancements in both mechanical and corrosion resistance characteristics [11]. Magnesium plays a significant role as a key alloying element in Al–Mg–Si alloys,

Published online: 11 January 2025

impacting the precipitation properties, rate, and hardness of the $Mg_2Si$ strengthening phase that forms during the ageing process [12–14]. Previous research has also indicated that increasing the magnesium concentration has potential in making the material more ductile during pre-ageing treatment [15].

There are many studies focused on investigating the influence of Mg and Si levels on the microstructure, crystallographic texture, and mechanical property of the Al–Mg–Si alloy. An investigation was conducted by Birol [16] to examine the impact of solute Mg on the grain size in Al–Mg alloys. In the case of aluminium alloys, the presence of magnesium has a long-lasting grain refining effect that persists until the eutectic point is reached. Magnesium is observed to significantly enhance the strength and work hardening properties of aluminium alloy [17, 18]. Thirugnanam et al. examined effect of magnesium on the fracture behaviour of cast samples of Al–7Si–Mg alloys [19]. Musa Yıldırım and Dursun Ozyurek examined the influence of magnesium addition to A356 aluminium alloy at varying amounts (0.43, 0.67 and 0.86 wt%) on the microstructure and mechanical characteristics. It was seen that the tensile strength and hardness of the alloy increase with an increasing magnesium content [20]. The effects of magnesium addition on the dispersoid precipitation as well as mechanical properties of 3xxx alloys were investigated by Zhen Li et al. [21]. Qiang Li et al. demonstrated hypoeutectic Al–10Si-2.0 Mg ternary alloy with high strength high ductility [22]. The increase in Mg content enhanced the propensity of continuous dynamic recrystallization and simultaneously decreased the average grain size of the alloy [23, 24].

The objective of the present study is to examine the varying quantity of magnesium on the tensile properties and formability of Al–Mg–Si alloys, which are crucial characteristics for their application in automotive and aerospace industries that demand appreciable formability for sheet metal components requiring intricate shapes, moderate strength and hardness. Nevertheless, there is a scarcity of data about the influence of magnesium content on the formability and microstructural analysis of Al–Mg–Si alloys. In order to achieve the objectives, tensile testing and stretch forming experiments have been done followed by microstructural investigations using the scanning electron microscopy (SEM) and electron backscatter diffraction (EBSD) methods.

## 2 Methodology

### 2.1 Material Selection

The selection of Al–0.9 Mg–Si and Al–3.1 Mg–Si alloys is based on their availability and wider relevance to the automotive industry. These alloys, part of the aluminium alloy 6xxx-Series, represent some common compositions used in commercial applications. The present study provides valuable insights for the automotive industry in selecting the sheet materials based on their performance requirements. The wrought AA 6xxx series Al–Mg–Si (Al–0.9 Mg–Si and Al–3.1 Mg–Si) alloy sheets with 1.5 mm thickness have been procured and annealed to 300° C for the present study. The Al–0.9 Mg–Si alloy contains 0.9% Magnesium and 1.07% Silicon, whereas the Al–3.1 Mg–Si alloy consists of 3.16% Magnesium and 0.98% Silicon. Both the alloys exhibit nearly identical contents of Manganese, Nickel, and Iron. The presence of magnesium and silicon in AA 6xxx series enhances ductility and material strength and makes them suitable for applications in structures, trusses, bridges, cranes and components related to transportation [23]. The chemical composition obtained by spark-based spectroscopy of both the materials is given in Table 1.

### 2.2 Mechanical Properties

Tensile testing samples (subsized) were prepared from sheets of Al–0.9 Mg–Si and Al–3.1 Mg–Si alloys by laser cutting as per ASTM-E8M standard. These samples were then subjected to uniaxial tensile testing to determine the tensile properties. An extensometer (make: Epsilon, Model: 3542-050 M-100-ST) was used during tensile testing to accurately measure the strain in the elastic and plastic region. The experiments were performed at a crosshead speed of 2.5 mm/min using a H50KS Bench-top Universal Testing Machine. In order to measure and evaluate normal anisotropy ($\bar{R}$), the specimens were tested at angles, 0°, 45°, and 90° w.r.t. the rolling direction (RD). The tests were conducted a minimum number of five times to ascertain the repeatability of the data. The Vickers method was used to measure the microhardness of the alloys. The test session includes a dwell period of 10 s and utilizes a hardness scale of HV0.5. A 500 g force was exerted utilizing a Struers DURAMIN 40 apparatus which uses a diamond pyramid indenter. Figure 1 displays the tensile specimens prepared at different orientations, tensile testing and microhardness experimental setups.

**Table 1** Chemical composition of aluminium alloys (in wt%)

| Alloy | Mg | Si | Mn | Sn | Cr | Ni | Fe | Al |
|---|---|---|---|---|---|---|---|---|
| Al–0.9 Mg–Si | 0.90 | 1.07 | 0.56 | 0.05 | 0.20 | 0.05 | 0.25 | Rest |
| Al–3.1 Mg–Si | 3.16 | 0.98 | 0.55 | 0.02 | 0.10 | 0.04 | 0.21 | Rest |

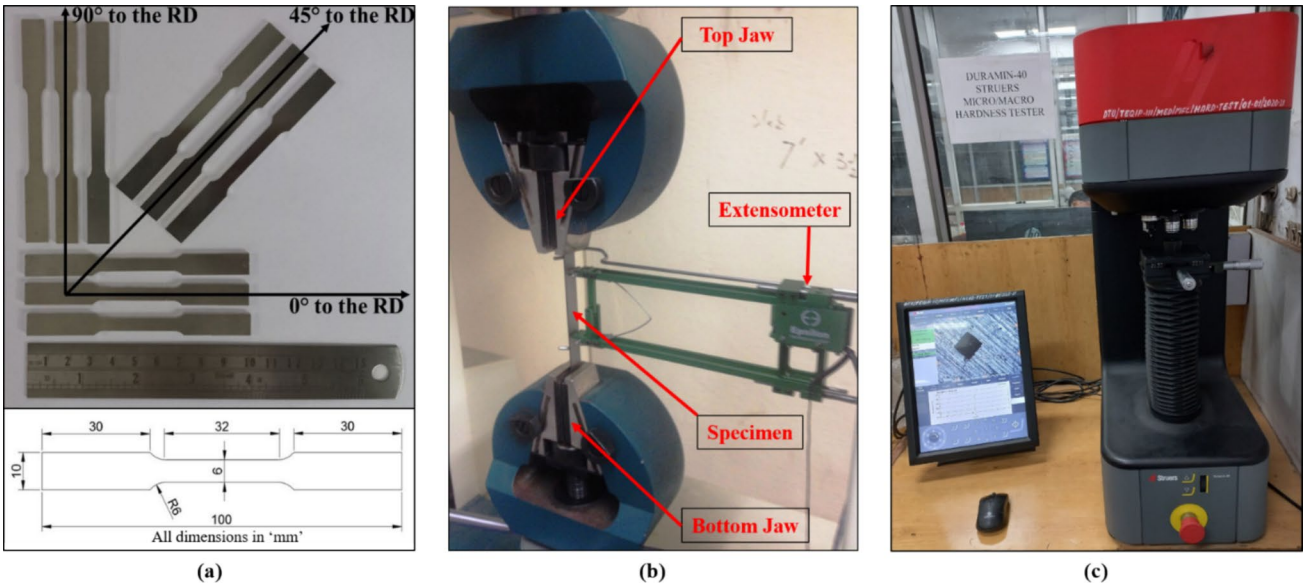



**Fig. 1** **a** Tensile specimens oriented at 0°, 45° and 90° to the RD, **b** Tensile testing in progress and **c** microhardness testing

### 2.3 Formability

To evaluate the formability, stretch forming experiments were carried out on the samples of different widths prepared by laser cutting operation. The stretch forming experiments were conducted on a double action hydraulic press having a capacity of 500 kN each for blank holder and punch. All the specimens with varying widths (10 mm to 100 mm) were laser etched with circular grid of diameter 2.5 mm and are shown in Fig. 2a, b so that the major and minor strains could be evaluated after the test. The laser etching was followed by a stretch forming using a 50-mm diameter hemispherical punch, during which the specimen was clamped between the blank holder and the die (Fig. 2c). The dome height of stretch formed samples was measured with the help of a digital vernier height gauge.

### 2.4 Microstructural Investigations

The crystallographic textures were examined by the EBSD technique applied on the samples obtained from both the original and deformed samples of sheet metal taken from the centre of the specimens subjected to tensile testing. In order to obtain appropriate analysis of samples during EBSD examinations, it is necessary to do electro-polishing. The experiment was conducted using a Struers Lectropol-5 machine with an electrolyte having 20% perchloric acid in methanol. The EDAX/TSL OIM software was utilized to enhance the precision of grain size determination in this procedure. The grain tolerance angle was adjusted to 5°, while the EBSD scan step size was set at 1.5 µm. Further, SEM–EDS of fractured or deformed samples were done to study the fracture and changes in elemental composition.

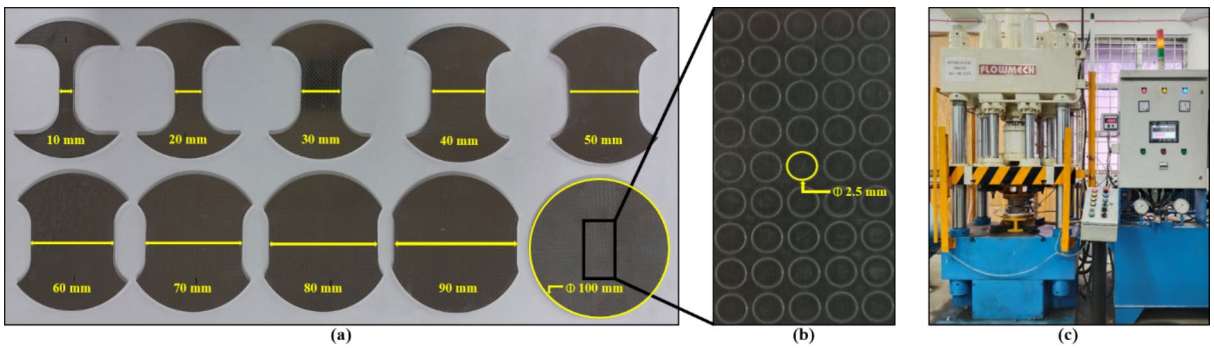



**Fig. 2** **a** Laser cut samples for FLD experiments **b** laser marked grid pattern on specimens and **c** experimental setup

## 3 Results and Discussion

### 3.1 Mechanical Properties

The tensile characteristics of both Al–0.9 Mg–Si and Al–3.1 Mg–Si alloy sheets are detailed in Table 2. The average yield and tensile strengths of the Al–3.1 Mg–Si alloy are observed to be 224 MPa and 289 MPa, respectively. The alloy demonstrates a 19% elongation and a strain hardening exponent of 0.21, indicating a modest level of formability. The Al–0.9 Mg–Si alloy sheet possesses a yield strength of 120 MPa and a tensile strength of 162 MPa on an average basis. This alloy exhibits an elongation of 9% and a strain hardening exponent of 0.18. These findings indicate that the Al–3.1 Mg–Si alloy, which contains more magnesium, has higher yield and tensile strengths and a greater elongation compared to the Al–0.9 Mg–Si alloy. The Al–3.1 Mg–Si alloy has an appreciable level of formability, as seen by its higher ductility and strain hardening exponent indicating higher stretchability. These tensile characteristics make Al–3.1 Mg–Si potentially more appropriate for applications that demand higher ductility and strength. The engineering stress–engineering strain and true stress–true strain plots are shown in Fig. 3a, b, respectively.

### 3.2 Forming Limit Diagram

Both major and minor strains are measured in the safe, necked, and failure zones on the tested samples, and FLD is drawn (shown in Fig. 4) so that strains at necking/fracture are lying just at or above the line. The measured limit strains are most likely the highest safe strains for these materials, and the dome height at failure may be slightly lower at Al–0.9 Mg–Si as compared to the Al–3.1 Mg–Si. The punch stroke at which necking or a crack in the sheet begins is considered as the stretch height.

Table 3 presents the measured limit dome heights of different widths following punch stretching experiments for both Al–0.9 Mg–Si and Al–3.1 Mg–Si alloys. The Al–3.1 Mg–Si alloy sample with a width of 60 mm exhibits the minimal dome height of 7.66 mm, or the $FLD_0$ in plane strain condition, with corresponding major and minor strains of 0.17 and 0.02, respectively. In the case of Al–0.9 Mg–Si alloy samples, width of 60 mm exhibits a dome height of 6.92 mm with corresponding major and minor strains of

**Table 2** Tensile properties of parent sheets

| Alloy | Yield strength (MPa) | Ultimate tensile strength (MPa) | Elongation (%) | Strain hardening exponent, *n* | Strength coefficient, K (MPa) | Hardness (HV0.5) | $\bar{R}$ |
|---|---|---|---|---|---|---|---|
| Al–0.9 Mg–Si | 120 ± 4 | 162 ± 6 | 9 ± 0.77 | 0.18 ± 0.01 | 304 ± 12 | 57 ± 1.14 | 0.91 ± 0.05 |
| Al–3.1 Mg–Si | 224 ± 9 | 289 ± 11 | 19 ± 1.63 | 0.21 ± 0.011 | 508 ± 19 | 107 ± 2.22 | 0.98 ± 0.08 |

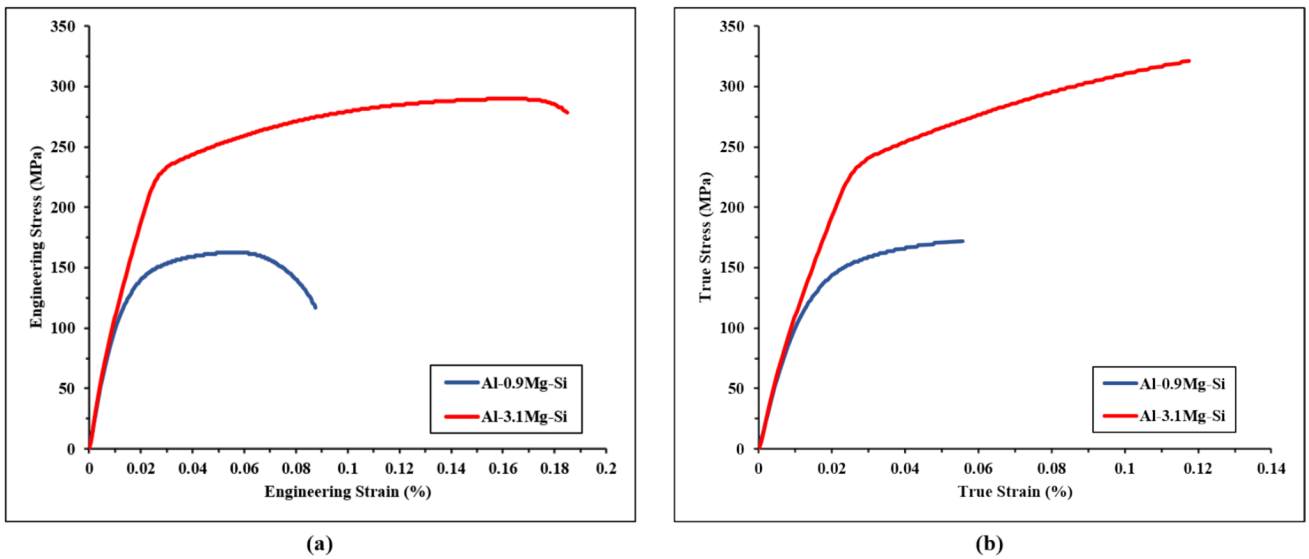



**Fig. 3** **a** Experimental results of engineering stress–engineering strain and **b** True stress–true strain plots

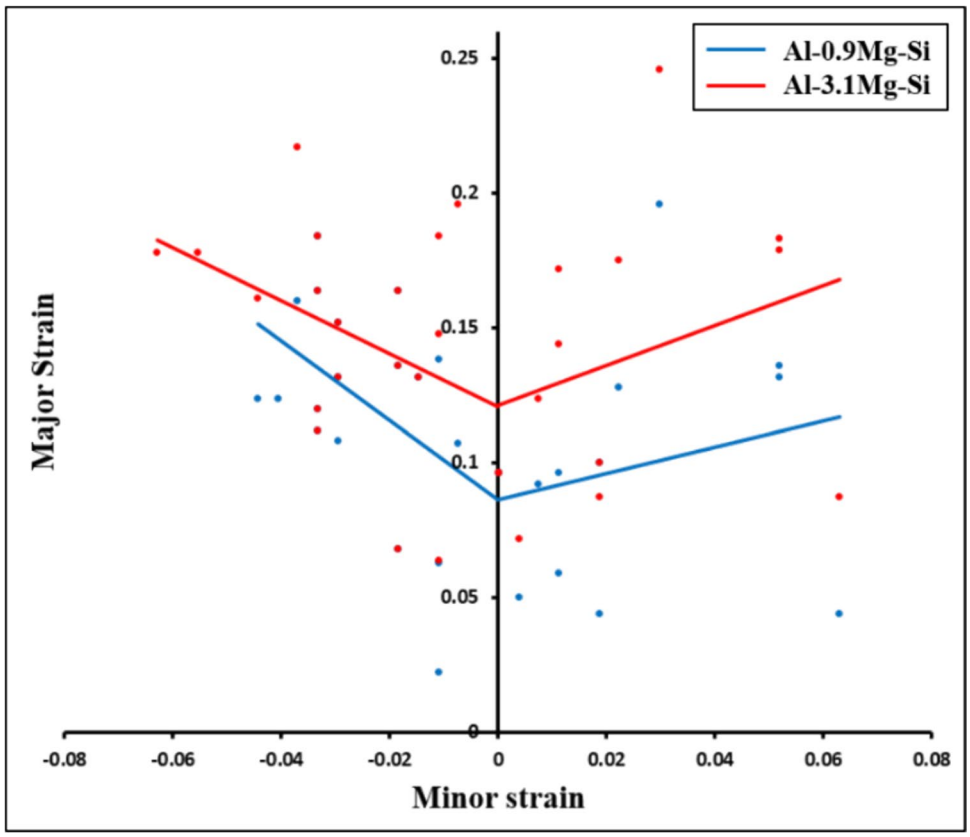



**Fig. 4** Forming limit diagram of Al–0.9 Mg–Si alloy and Al–3.1 Mg–Si alloy sheet

**Table 3** Comparison of dome heights obtained in Al–0.9 Mg–Si and Al–3.1 Mg–Si alloys

| Alloy sheet | Width (mm) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |
| | Dome height (mm) | | | | | | | | | |
| Al–0.9 Mg–Si | 6.85 | 7.51 | 8.78 | 7.08 | 7.52 | 6.92 | 7.66 | 6.86 | 7.22 | 7.14 |
| Al–3.1 Mg–Si | 7.76 | 8.89 | 9.48 | 7.34 | 8.01 | 7.66 | 7.87 | 7.06 | 9.34 | 7.24 |

0.12 and 0.018. As a result, in both the cases, $FLD_0$ falls on the right side of the plane strain condition. The deformed sample exhibits positive major and minor strains, indicating a biaxial stretch (tension–tension) condition. This may have contributed to the material's decreased ductility. It is also observed that the formability of the Al–3.1 Mg–Si alloy sheet is higher as compared to the Al–0.9 Mg–Si alloy sheet. The samples of Al–0.9 Mg–Si and Al–3.1 Mg–Si alloys sheets after forming experiments are shown in Fig. 5.

### 3.3 Microstructural Investigations

The effect of magnesium content on microstructure with reference to the shape, size and distribution of an intermetallic compound, i.e. magnesium silicide ($Mg_2Si$) present in the base material, is explained on the basis of the microstructure obtained by SEM images at a lower magnification as shown in Fig. 6a, b. It is apparently seen that $Mg_2Si$ particles are present in a diffused form of speckles, finer in the case of lower Mg content distributed at random within the microstructure as shown in Fig. 6a. Whereas, in the case of the higher Mg content samples (Fig. 6b), the round shape of the $Mg_2Si$ speckles is well defined and the compound is observed to be distributed uniformly within the microstructure. It is a proven fact that an increase in the content of Mg makes the grains finer [24]. The combination of finer grains and evenly distributed $Mg_2Si$ particles provides well-distributed barriers to dislocation migration during plastic deformation, increasing the material's tensile strength. Also, a homogeneous dislocation density within the grain increases ductility of the sheet with higher Mg content.

Furthermore, in order to explore the influence of Mg on grain size and its distribution, the microstructures of an undeformed and deformed (up to the UTS) sheets were also investigated by the use of inverse pole figure (IPF) map (using EBSD technique) as shown in Fig. 7a–d.

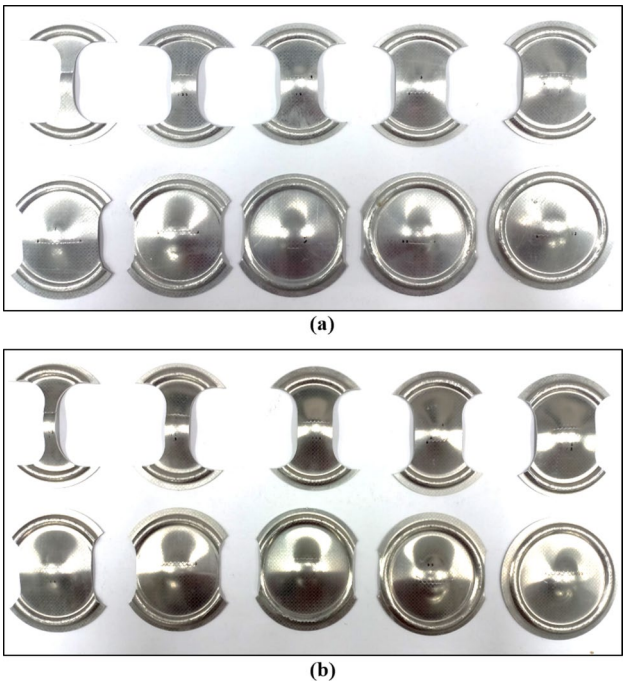



**Fig. 5** **a** Al–0.9 Mg–Si and **b** Al–3.1 Mg–Si alloys sheets after forming operation

Figure 7a exhibits the microstructure of the annealed sheet, and it shows that the Al–0.9 Mg–Si alloy has equiaxed grains with an average grain size of 51.46 µm, whereas the Al–3.1 Mg–Si alloy has equiaxed grains with an average grain size of 35.67 µm as shown in Fig. 7b. After the elongation just before the ultimate tensile strength (UTS), the Al–0.9 Mg–Si alloy grains are further elongated, but the average grain size decreases slightly to 47.58 µm, whereas the Al–3.1 Mg–Si alloy shows elongated grains with finer size of 32.48 µm as shown in Fig. 7c, d, respectively. According to the Hall–Petch relation, the finer grains of Al–3.1 Mg–Si alloy attributed to its higher strength when compared to that of Al–0.9 Mg–Si alloy.

The pole figures of the undeformed samples of Al–0.9 Mg–Si and Al–3.1 Mg–Si alloys are presented in Fig. 8a, b, illustrating the distribution of (001), (110), and (111) planes. For the as-received Al–3.1 Mg–Si alloy, the pole figure of the (111) planes shows a higher peak intensity (4.371 MRD (multiples of random distribution)), indicating a strong crystallographic texture [25, 26] followed by the intensities of the (001) and (110) planes. Similarly, the as-received Al–0.9 Mg–Si alloy exhibits a higher peak intensity (3.751 MRD) in the (111) planes, suggesting a good texture, with subsequent peak intensities in the (110) and (001) planes. The Al–3.1 Mg–Si alloy demonstrates a stronger texture than the Al–0.9 Mg–Si alloy, correlating with the greater deformation observed in the Al–3.1 Mg–Si alloy compared to the Al–0.9 Mg–Si alloy. After elongation to just before the necking at room temperature, the Al–0.9 Mg–Si alloy's (001) planes exhibit a higher peak intensity (5.970 MRD), indicating a very strong and sharp texture, followed by the (111) and (110) planes, as shown in Fig. 8c. The increased texture along the (111) direction significantly influences the peak intensity change in the (001) plane after elongation. Similarly, for the Al–3.1 Mg–Si alloy after elongation to just before UTS at room temperature, the (001) planes show a higher peak intensity (4.852 MRD), indicating a strong texture, followed by the (111) and (110) planes, as shown in Fig. 8d. The change in peak intensity in the (001) plane of the Al–3.1 Mg–Si alloy after deformation is mainly attributed to the increased texture along the (111) direction.

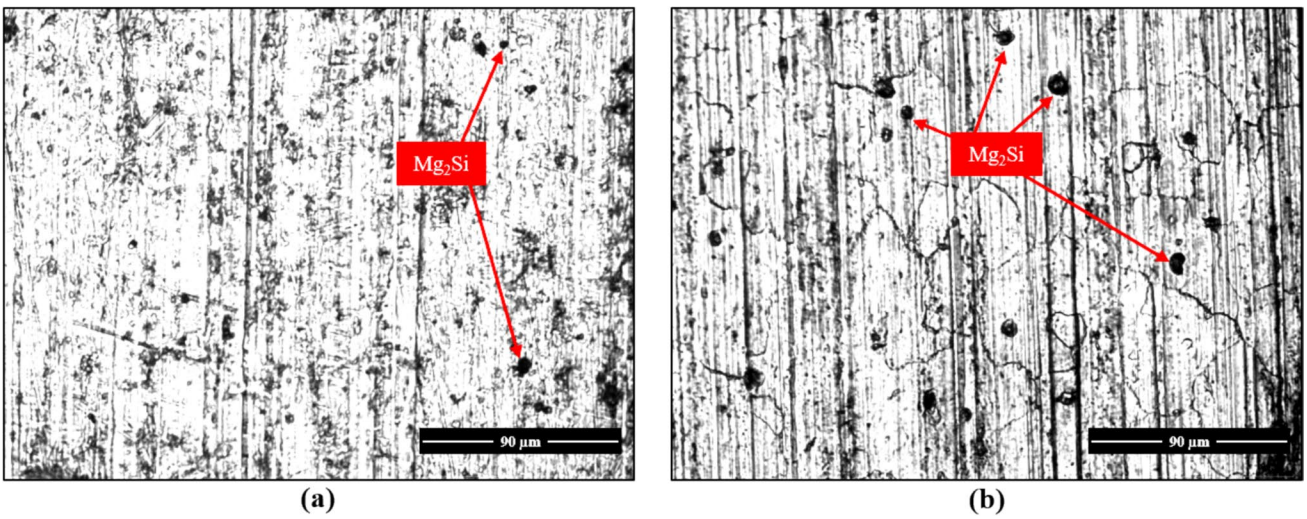



**Fig. 6** SEM image of **a** Al–0.9 Mg–Si and **b** Al–3.1 Mg–Si

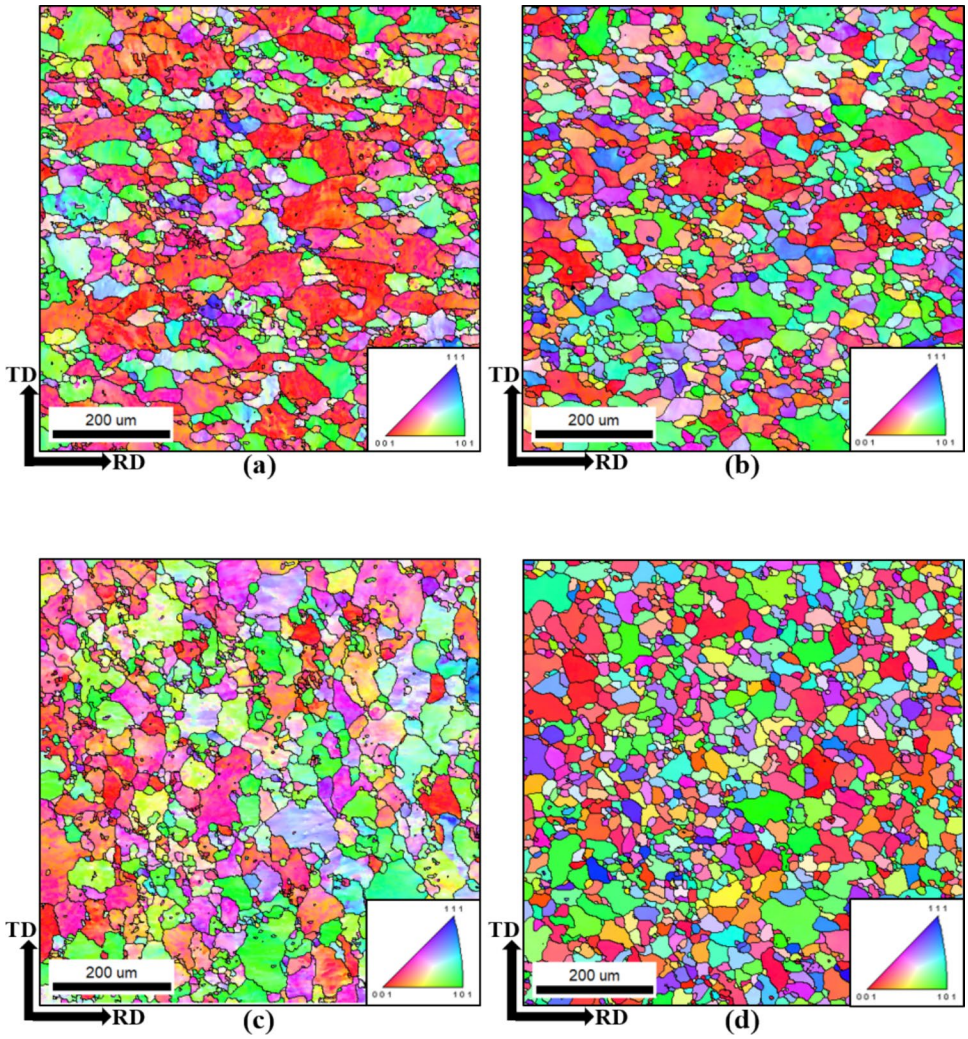



**Fig. 7** IPF map of un-tested samples **a** Al–0.9 Mg–Si, **b** Al–3.1 Mg–Si, and tested samples **c** Al–0.9 Mg–Si and **d** Al–3.1 Mg–Si

The observed texture change after plastic deformation in the samples of Al–3.1 Mg–Si alloy is consistent with the theory of deformation texture evolution. According to this theory, during plastic deformation, grains in polycrystalline materials reorient along specific crystallographic planes and directions. EBSD analysis confirmed that both alloys exhibited a strong {111} texture, which is typical for face-centred cubic materials like aluminium alloys. Post-deformation, the Al–3.1 Mg–Si alloy's (001) planes showed increased peak intensity, indicating a reorientation of grains into stable slip systems. These stable orientations facilitate uniform plastic deformation, supporting the theoretical expectations of texture evolution [27]. The stronger texture in the Al–3.1 Mg–Si alloy aligns with the expected behaviour, enhancing the material's ductility and formability.

The fractured surfaces of the tested samples were analysed using scanning electron microscopy (SEM). The fractured surfaces of the Al–0.9 Mg–Si and Al–3.1 Mg–Si alloy samples, which were subjected to uniaxial stress until failure, reveal that most structural alloys undergo ductile damage through a process involving micro-void formation and cleavage crack surfaces, as illustrated in Fig. 9a–d. In the Al–0.9 Mg–Si alloy, the observed micro-voids are relatively small, suggesting lower ductility. Conversely, the fracture surface of the Al–3.1 Mg–Si alloy exhibits larger and deeper dimples, indicating greater deformation before failure, which corresponds to higher limit strains. This alloy's fracture surface shows a combination of elongated and larger micro-voids cleavage compared to the Al–0.9 Mg–Si alloy. In both alloys, these micro-voids nucleate at local strain discontinuities, such as secondary phase particles, which are frequently located near the cavities in the fractured regions.

## 4 Conclusions

In the present work, the experimental findings provide valuable insights into how magnesium content influences the strength, grain size, strain distribution during plastic deformation and other microstructural changes in Al–Mg–Si

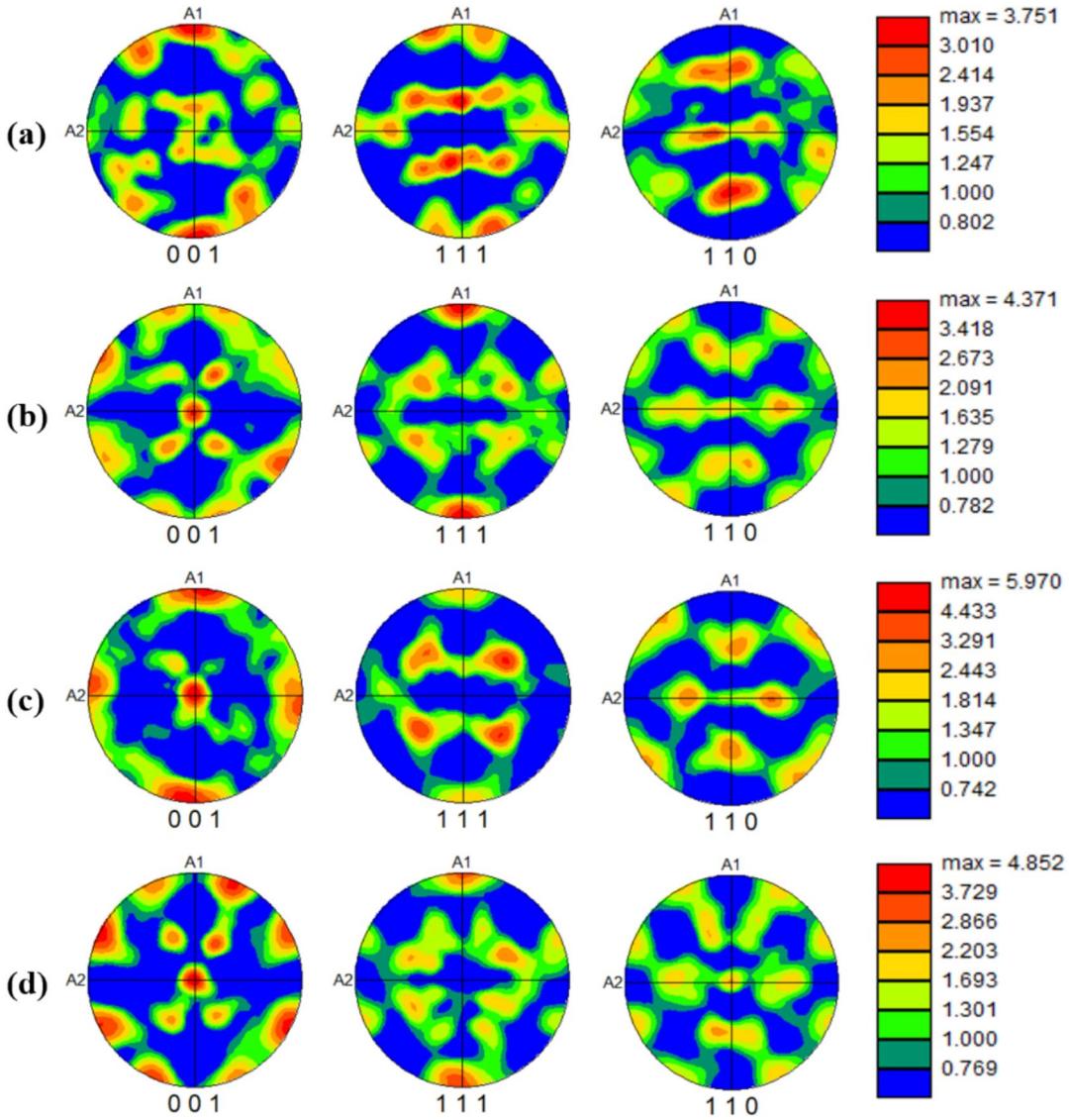



**Fig. 8** Pole figure with scale bar for as-received samples **a** Al-0.9 Mg-Si, **b** Al–3.1 Mg–Si; and deformed samples **c** Al–0.9 Mg–Si, and **d** Al–3.1 Mg–Si

alloys. Based on the experimental and theoretical investigations, the following conclusions are drawn:

1. The Al–3.1 Mg–Si alloy demonstrated superior mechanical properties, with yield and tensile strengths approximately 1.8 times higher than those of the Al–0.9 Mg–Si alloy sheet. Additionally, the higher strain hardening exponent of the Al–3.1 Mg–Si alloy indicates better stretchability during plastic deformation.
2. The better tensile properties may be attributed to the combination of finer grains and evenly distributed $Mg_2Si$ particles which provide well-distributed barriers to dislocation migration during plastic deformation.
3. The forming limit strains and $FLD_0$ are higher in case of Al–3.1 Mg–Si alloy sheets as compared to Al–0.9 Mg–Si alloy sheets suggesting higher formability for the alloy sheet with higher Mg content. The dome height of the Al–3.1 Mg–Si specimen was observed to be greater than that of the Al–0.9 Mg–Si specimen, indicating that the Al–3.1 Mg–Si alloy exhibits greater ductility and stretchability.
4. The pole figure maps indicate that undeformed parent samples of Al–3.1 Mg–Si alloy exhibit a stronger crystallographic texture compared to the Al–0.9 Mg–Si alloy. The deformed samples of Al–3.1 Mg–Si alloy demonstrate a stronger and sharper texture which is an

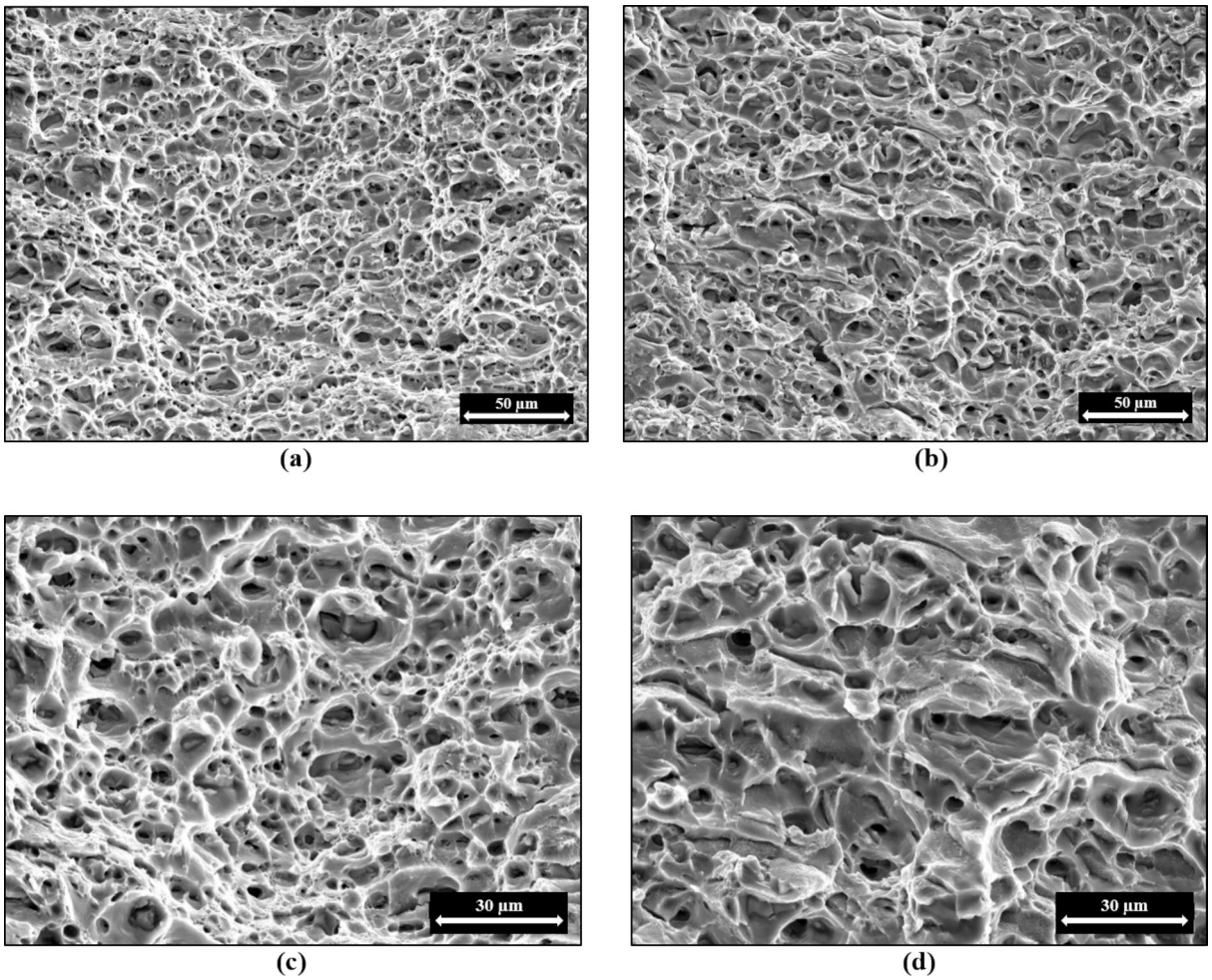



**Fig. 9** Fractographs for **a**, **c** Al–0.9 Mg–Si, and **b**, **d** Al–3.1 Mg–Si alloy

important feature for sheet metals undergoing plastic deformation.

**Acknowledgements** The authors would like to acknowledge the staff of the OIM and Texture Lab at the Indian Institute of Technology Bombay, India, for providing the testing of the EBSD technique and evaluation of the results.

**Authors Contribution** Rakesh Kumar has conceptualized and performed the experimental work. He has also written the original draft. Vijay Gautam has supervised, reviewed and edited the manuscript.

**Declarations**

**Conflict of interest** The authors declare that there is no conflict of interest.

## References

1. Zolotorevsky V S, Belov N A, and Glazoff M V, *Casting Aluminum Alloys*, Elsevier, Amsterdam (2007).
2. Azadi M, and Shirazabad M M, *Mater & Des* **45** (2013) 279.
3. Estey C, Cockcroft S, Maijer D, and Hermesmann C, *Mater Sci Eng: A* **383** (2004) 245.
4. Hekmat-Ardakan A, Liu X, Ajersch F, and Chen X-G, *Wear* **269** (2010) 684.
5. Zhu M, Jian Z, Yang G, and Zhou Y, *Mater & Des* **36** (2012) 243.
6. L N E P. AA6xxx, Mater Tehnol **39** (2005) 101.
7. Maisonnette D, Suery M, Nelias D, Chaudet P, and Epicier T, *Mater Sci Eng: A* **528** (2011) 2718.
8. Panigrahi S K, and Jayaganthan R, *J Alloys Comd* **470** (2009) 285.
9. Shah K, Kumar S, and Dwivedi D K, *Mater & Des* **28** (2007) 1968.
10. Yao J, Graham D, Rinderer B, and Couper M, *Micron* **32** (2001) 865.

11. Mohamed A, Samuel A, Samuel F, and Doty H, *Mater & Des* **30** (2009) 3943.
12. Hirth S, Marshall G, Court S, and Lloyd D, *Mater Sci Eng: A* **319** (2001) 452.
13. Zhen L, Fei W, Kang S, and Kim H, *J Mater Sci* **32** (1997) 1895.
14. Gupta A, Lloyd D, and Court S, *Mater Sci Eng: A* **316** (2001) 11.
15. Gupta A, Marois P, and Lloyd D J, *Mater Sci Forum* **217** (1996) 801.
16. Birol Y, *Mater Sci Technol* **28** (2012) 924.
17. Davis J R, Corrosion of Aluminum and Aluminum alloys, Asm International (1999).
18. Mondolfo L F, Aluminum Alloys: Structure and Properties. Elsevier (2013).
19. Thirugnanam A, Sukumaran K, Pillai U, Raghukandan K, and Pai B, *Mater Sci Eng: A* **445** (2007) 405.
20. Yıldırım M, and Özyürek D, *Mater & Des* **51** (2013) 767.
21. Zhen L, and Zhang Z, *Trans Nonferrous Metals Soc China* **26** (2016) 2793.
22. Li Q, Qiu F, Dong B X, Yang H Y, Shu S L, Zha M, and Jiang Q C, *Mater Sci Eng: A* **798** (2020) 140247.
23. Gautam P C, and Biswas S, *Mater Sci Eng: A* **812** (2021) 141103
24. Li C, Gu H, Wang W, Wang S, Ren L, Wang Z, Ming Z, and Zhai Y, *Mater Technol* **36** (2021) 793.
25. Priyadarshini A, Sancheti C I, and Dwarka S S, Comparative study of forge ability of as extrude AA6082 and as cast AA6082 alloy (2015) 1.
26. Gautam V, and Kumar A, *Procedia Manuf* **29** (2019) 472.
27. Kocks U F, Tomé C N, and Wenk H R, *Texture and Anisotropy: Preferred Orientations in Polycrystals and Their Effect on Materials Properties*, Cambridge University Press, Cambridge (2000).

**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.

This article has been accepted for publication in IEEE Transactions on Nanotechnology. 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/TNANO.2025.3532313

# Insights into Temperature sensitivity analysis of polarity controlled charge plasma based Tunable Arsenide/Antimonide Tunneling Interfaced Junctionless TFET

S. Sharma, J. Madan, and R. Chaujar, Sr. *Member, IEEE*

***Abstract*— This research delves into the temperature-dependent performance of a novel polarity-controlled charge plasma-based InAs/AlGaSb tunneling interfaced junctionless TFET (H-JLTFET). The device leverages the benefits of both charge plasma and heterojunction engineering to enhance device performance. Comprehensive simulations were conducted to assess the impact of temperature on device characteristics. Results indicate that while the device exhibits promising ON-state current and high-frequency metrics, with a peak $f_T$ of 417 GHz and an $f_{max}$ of 4390 GHz, the subthreshold region is significantly influenced by temperature. The observed increase in OFF-state current and degradation in subthreshold swing highlight the need for careful thermal management and circuit design. Furthermore, the study reveals a moderate impact of temperature on intrinsic delay and a slight increase in ambipolar current. Overall, this work provides valuable insights into the thermal behavior of H-JLTFETs, paving the way for optimized device design and reliable operation in various applications.**

***Index Terms*— Analog performance, Charge plasma, H-JLTFET, Junctionless, RF performance, Silvaco, Tunnel FET, Temperature dependence.**

## I. Introduction

Technological evolutions have led to a significant drop in transistor dimensions, resulting in amplified packaging density. This miniaturization of transistors offers both rewards and challenges [1]. privileges consist of higher switching speeds, smaller chip sizes, and lesser power dissipation. Nevertheless, these advantages are accompanied by critical issues such as short channel effects [2]. Downscaling can encompass either full device dimension scaling, where all device features are downscaled by a factor 's', or continual voltage scaling. In the case of dimension downsizing, device productiveness is degraded by the reduction in channel length, leading to short channel effects viz Drain Induced Barrier Lowering (DIBL), subthreshold swing (SS) declination, and hot electron effects [3-5]. Such growing short channel effects and leakage currents in current CMOS technology pose significant challenges at the nanometer scale. Tunnel Field Effect Transistors (TFETs) have arisen as an auspicious alternative owing to their quantum tunneling mechanism. TFETs exceed the fundamental subthreshold swing (SS) constraint of MOSFETs (60 mV/dec), providing an effective solution to reduce power dissipation [6-8].

However, there are a few issues – lower ON state current, poor high frequency characteristics, and ambipolarity [9]. Researchers have introduced various techniques to overcome these issues by utilizing dielectric engineering [10], material engineering [11], work function variation [12], heterojunction [13], and pocket insertion [14] and many more. Furthermore, to resolve some of the issues related to fabrication such as complex thermal budget, random dopant fluctuations, and high temperature annealing techniques, charge plasma method is reported in the literature [15-16]. The charge plasma approach does not rely on the physical doping methods to transform a heavily doped $n^+$- $n^+$-$n^+$ substrate into a conventional $p^+$-i-$n^+$ substrate. Two metal electrodes - polarity gate (PG) and a control gate (CG) with different work functions sited above the source and channel sections of TFET are incorporated to achieve this transformation. When a gate voltage is applied, the PG induces a hole plasma in the source region, while the CG induces an abrupt electron plasma in the channel region [12,16]. A hole plasma is generated in the source area under the PG electrode, while a distinct electron plasma forms beneath the CG electrode when gate voltage is applied. However, conventional Silicon based homojunction TFETs still suffer from low ON-state current. This is because a large energy bandgap indirect semiconductor like silicon, has lower carrier mobility and band-to-band tunneling leading to inferior ON-state current [10-12, 16]. To overcome this issue, hetero-material engineering using direct bandgap III-V compound semiconductors has been effective. In this approach, a TFET structure with lower bandgap source material and a higher bandgap channel material significantly improves electrical performance compared to silicon-based homojunction TFET devices [15-16]. Another important aspect to study the reliability of TFET devices is the temperature dependence because of the following reasons:

(Corresponding author: R. Chaujar)

S. Sharma is with the National Institute of Technology Kurukshetra, Haryana, India (e-mail: sharmasamriti72@gmail.com).

J. Madan is with VLSI Centre of Excellence, Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India (e-mail: jaya.madan@chitkara.edu.in)

R. Chaujar is with Department of Applied Physics, Delhi Technological University, Delhi, India (e-mail: chaujar.rishu@dtu.ac.in).

Copyright © 2025 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:53:21 UTC from IEEE Xplore. Restrictions apply.

© 2025 IEEE. 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 Nanotechnology. 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/TNANO.2025.3532313



- Upsurge in the operating temperature of TFET with increasing the number of transistors on a chip, which further upsurges the heat dissipation [17].
- Industrial usage of transistors based electronic controls under harsh environment conditions – extreme hot and cold weather [18].
- Temperature dependence of energy bandgap of semiconducting materials [19].

The energy bandgap of a semiconducting material depends upon the temperature [20] as defined by Eq. (1).

$$E_g(T) = E_g(0) - \frac{\alpha T^2}{\beta + T} \quad (1)$$

where $E_g$ (0) signifies the energy bandgap at zero kelvin, T denotes the absolute temperature in Kelvin, α and β represent material parameters. Hence, the energy band gap lowers down with the rise in temperature and leads to rise in the ON-State current [20]. Furthermore, variation of drain current under the influence of temperature changes significantly with the gate voltage. In OFF-state having very low gate voltage, the Shockley–Read–Hall (SRH) phenomenon dominates with strong dependence on temperature and hence the leakage current degrades with rise in temperature. On the contrary, Band to Band tunnelling dominates at a high gate voltage (ON-state) having weak temperature dependence. Thus, temperature has a noticeable effect in the OFF state, but its impact at high gate voltage in the ON state is relatively minor. Therefore, temperature sensitivity should be given prior importance to understand the overall device stability. The understanding of device behaviours at a wider temperature range can help to foster various electronic applications – defence, health, wireless communication, space programmes, automobile, and various temperature control applications. The core purpose of this article is to examined the temperature dependent reliability of an innovative integration of advanced compound semiconductor hetero-material InAs/AlGaSb with the charge plasma junctionless TFET in terms of the analog/RF figure of merits (FoMs) for different temperatures, which has been unexplored in the literature otherwise.

## II. II. Device Designing and Simulation Set-up

A 2-D schematic device structure of H-JLTFET is illustrated in Fig. 1. The device design parameters are tabulated in Table 1 by following the International Roadmap for Devices and Systems (IRDS) [21]. A novel amalgamation of charge plasma concept along with III-V compound semiconducting materials InAs and AlGaSb has been merged in the source and channel sections, respectively. The InAs being lower bandgap binary compound semiconducting material, intensify the ON-state current ($I_{ON}$). The $Al_XGa_{1-X}Sb$ is the ternary compound semiconducting material, whose energy band gap can be optimized by changing the aluminum (Al) mole fraction (X). The detailed optimization of the Al mole fraction in AlGaSb used in the channel region has been comprehensively addressed in our previous research [16]. The device performance is examined for Al mole fraction value of 0.15. The conduction and valence band offset at InAs and $Al_{0.15}Ga_{0.85}Sb$ heterojunction interface depicts the enhanced band overlapping. This adjustable bandgap results in a reduction of the tunnel barrier at the source-to-channel (S/C) interface, significantly enhancing the tunneling probability of carriers and consequently leading to an increased $I_{ON}$ as shown in [11]. The charge plasma methodology has been applied by doping the source and channel regions uniformly with a concentration of $1\times10^{19}$ $cm^{-3}$. To prompt the hole plasma in the source region ($p^+$ source) and electron plasma in the channel region, metal gates – PG (5.93 eV) and CG (4.5 eV) are used. The simulations of the H-JLTFET are conducted by the TCAD SILVACO device simulator [22].

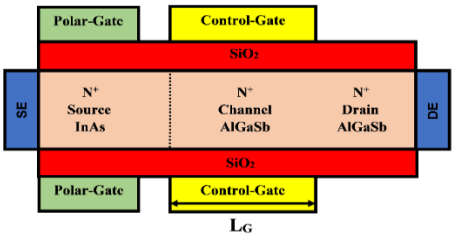



**Fig. 1.** Graphic structure of H-JLTFET.

The assessments regarding the tunneling potential and electric field throughout the S/C interface, a non-local BTBT model is executed by devising a quantum tunneling mesh at the edges. To incorporate calculations of mobility dependent on electric field, temperature, and concentration, CVT (Lombardi) model is administered. To incorporate calculations of mobility dependent on electric field, temperature, and concentration, CVT (Lombardi) model is administered. Quantum confinement effects are well-thought-out by implementing the Hansch model (HANSCHQM) [23]. The SRH and Auger recombination models are employed to weigh up the trap assisted tunneling and the minority carrier recombination effects. To realise the temperature dependent variation in H-JLTFET, TEMP statement is used. Additionally, band gap narrowing (BGN) and Fermi Dirac statistics are applied. Prior to extracting the result parameters from the TCAD simulations, the device physical simulation models are calibrated compared to the experimentally reported data [24] (Fig. 2).

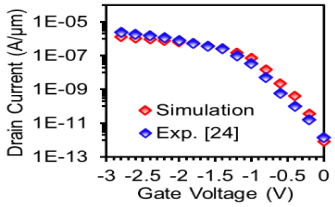



**Fig. 2.** Endorsement of simulation outcomes with experimental results for H-JLTFET [24].

TABLE I Device Design Parameters

| Parameter | *Symbol* | *Value* |
|---|---|---|
| **Channel Length** | $L_G$ | 20nm |
| **Body Thickness** | $T_{Body}$ | 3nm |
| **Doping Concentration** | $N_D$ | $1 \times 10^{19}$ $cm^{-3}$ |
| **Oxide thickness** | $T_{OX}$ | 2nm |
| **Oxide length below CG** | $L_{OX}$ | 20nm |
| **Isolation thickness** | $L_{Gap}$ | 2nm |
| **CG work Function** | $\phi_{CG}$ | 4.5eV |
| **PG Work Function** | $\phi_{PG}$ | 5.93eV |

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:53:21 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Nanotechnology. 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/TNANO.2025.3532313



Since, InAs/AlGaSb hetero-material combination has not yet been utilized in H-JLTFET fabrication, we first calibrated the local band-to-band tunneling model, using experimental data previously reported for hetero-material SiGe, Ge, and Si in HS-TFETs [24]. To ensure accuracy, we matched all device parameters and biasing conditions precisely with those in [24] (Fig. 2). The transfer characteristics from the experimental work in [24] were extracted, and simulations were conducted using Silvaco-ATLAS TCAD software. The resulting data, compared in Fig. 2, demonstrate a close agreement between the experimental and simulated results, thereby validating the model parameters. Once the simulation models were calibrated, we applied them to our proposed device, fine-tuning the hetero-material properties of InAs and AlGaSb within the simulation framework. The fabrication process flow of our designed device is illustrated in Fig. 3. The hetero-material S/C interface is created using a self-aligned optical lithography technique, as described in [25]. Subsequently, electron beam lithography and ultraviolet lithography are employed to deposit the metal gates and define the spacer region [26-27]. The dielectric layer is then deposited through atomic layer deposition [28], followed by the formation of metal contacts using electron beam evaporation [29]. Hence, H-JLTFET is formed and continues processing.

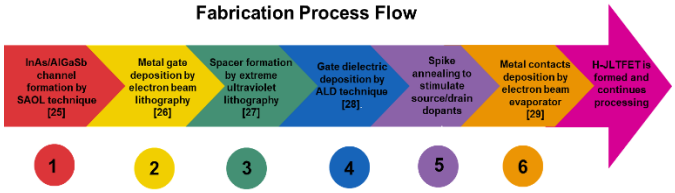



**Fig. 3.** Fabrication process flow of proposed H-JLTFET.

## III. Result Analysis and Discussion

### *A. Analog Analysis*

Understanding the characteristics of H-JLTFET is crucial for the development of robust and thermally stable H-JLTFETs for future nanoelectronic applications. In this respect, a comprehensive temperature analysis for H-JLTFET has been presented in this work. Fig. 4(a) demonstrates the transfer characteristics of H-JLTFET device for various temperatures ranging from 250K to 400K. The drain current is presented on a logarithmic scale, highlighting the subthreshold and superthreshold regions. The temperature dependence of the transfer characteristics in H-JLTFETs involves a complex interplay of various physical mechanisms. It is clearly evident from the plot that the temperature sensitivity is significantly large in the subthreshold regime, while the device is marginally affected by temperature in the superthreshold regime. This is because in the subthreshold regime, the SRH phenomenon dominates, which in turn has an exponential temperature dependence. In contrast, in the superthreshold regime, the band-to-band tunneling phenomenon dominates, which has low temperature dependence. Thus, the off-state current degrades exponentially in the subthreshold regime, while the on-state current is marginally affected by temperature variation. It has been evaluated that with an increase in temperature from 250K to 400K, the off-state current increases from an order of 10-20 A/μm to 10-16 A/μm (see Fig. 4(c)), while the on-state current increases moderately by only 18.24%, i.e., from 13.7 μA/μm to 16.2 μA/μm (see Fig. 4(c).

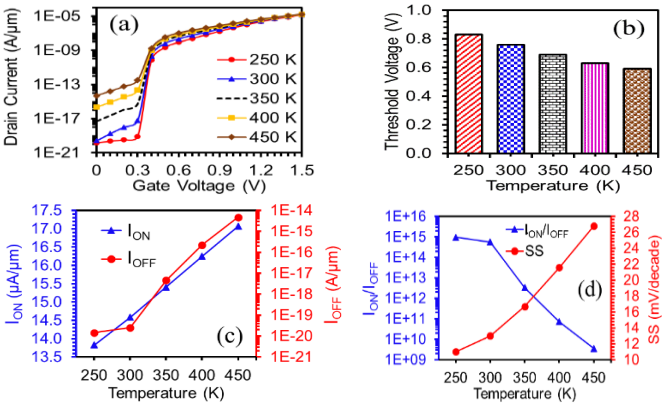



**Fig. 4.** Temperature dependence of (a) transfer characteristics (b) threshold voltage, $V_{th}$, (c) $I_{ON}$ and $I_{OFF}$ (d) $I_{ON}/I_{OFF}$ ratio and subthreshold swing in H-JLTFET.

Further, a clear temperature dependence is observed in the threshold voltage, showing that as the temperature increases from 250K to 400K, $V_{th}$ shifts towards lower values, i.e., from 0.82V to 0.62V. This negative temperature coefficient is attributed to the increased dominance of SRH in the subthreshold regime and the narrowing of the bandgap at higher temperatures. Additionally, the subthreshold swing (SS), a measure of the gate control efficiency, also exhibits a temperature dependence as depicted in Fig. 4(d). SS increases from 11mV/decade to 22mV/decade with rising temperatures from 250K to 400K, indicating a degradation in the steepness of the transition between the off-state and on-state. This can be attributed to the increased leakage current. Furthermore, it must be noted that the proposed H-JLTFET device has a positive temperature coefficient that necessitates careful consideration in circuit design to ensure reliable performance across a wide temperature range. The increase in SS and off-state current with temperature highlights the need to minimize leakage currents and optimize device parameters for high-temperature operation.

To investigate the carrier dynamics and temperature affectability in H-JLTFET, the energy band diagram and carrier concentration in device at OFF state and ON state at temperature range 250K to 400K has been plotted in Fig. 5(a-b) and Fig. 6(a-b). The heterojunction architecture strategically integrates InAs, characterized by a narrow bandgap of 0.35 eV, in the source region, and AlGaSb, with a wider bandgap of 1.03 eV, in the channel region to optimize device performance. To enhance understanding, the band alignment of the InAs-$Al_{0.15}Ga_{0.85}Sb$ heterojunction, previously detailed in our work [11] as Fig. 3(d), is referenced. At the Source/Channel (S/C) interface, the heterojunction forms a staggered configuration within the depletion region upon contact. This results in a distinctive energy band feature under thermal equilibrium, observable as a pronounced kink near the 18 nm mark in Fig. 5(a-b). The presence of this kink

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:53:21 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Nanotechnology. 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/TNANO.2025.3532313



induces local minima in the conduction band at the S/C interface, serving as a critical point for charge carrier dynamics. In the ON state, the application of a gate bias leads to significant band bending at the S/C interface. This bending reduces the width of the potential barrier, triggering a pronounced increase in band-to-band tunneling of carriers across the conduction band minima region. The hetero-material S/C interface (InAs/AlGaSb) and homo-material D/C interface in the energy band diagram shows minor variation in terms of tunneling barrier width reduction and minority carriers tunneling in the OFF state for the entire temperature range (Fig. 5(a)).

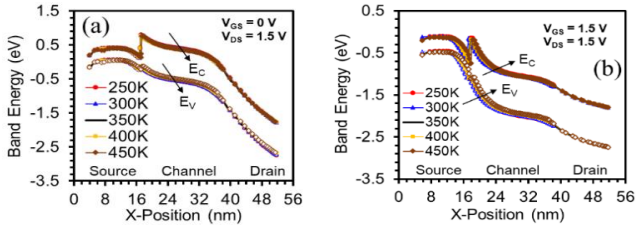



**Fig. 5.** Temperature dependence of energy bands in H-JLTFET in (a) OFF state and (b) ON state.

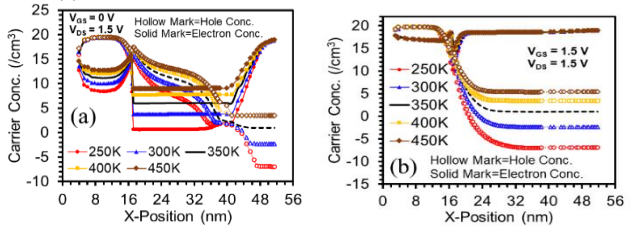



**Fig. 6.** Temperature dependence of carrier concentration (both electrons and holes) in H-JLTFET in (a) OFF state and (b) ON state.

Upon applying gate biasing, the conduction band minima are observed at the source/channel interface of the H-JLTFET for the entire temperature range. Furthermore, the dominance of BTBT at the channel's conduction band and source's valence band remains consistent, with a decrease in the potential barrier width and an increase in tunneling carriers as the temperature rises from 250 K to 400 K (see Fig. 5(b)). This is further supported by the temperature dependence of the charge carrier concentration in the H-JLTFET, both in the OFF state and ON state, from 250 K to 400 K. The electron concentration at the source/channel interface increases with temperature even in the OFF state, as shown in Fig. 6(a), while the ON-state electron and hole concentration is only marginally affected by temperature changes (Fig.6(b)).

As shown in Fig. 7(a), the BTBT rate of H-JLTFET is minimally affected by temperature variations, resulting in only minor changes in the tunneling rate. This provides better immunity against temperature increment from 250 K to 450 K. Another important aspect of H-JLTFET device is the ambipolar conduction resulting from the BTBT at the D/C interface for negative gate biasing conditions. The ambipolar current ($I_{AMB}$) limits the feasibility and applicability of H-JLTFET for circuit-level applications. Therefore, overlooking the quantitative analysis of this aspect may lead to performance degradation at elevated temperatures. Fig. 7(b) shows the elevation in the ambipolar current in H-JLTFET as the temperature rises from 250 K to 450 K, due to the increase in drain current. Further, for digital applications, such as in digital logic circuits or high-speed communication systems, faster switching speeds are desirable. Transconductance ($g_m$) plays key role in determining the switching speed of FETs. Higher $g_m$ means the device can switch on and off more quickly, leading to faster operation.

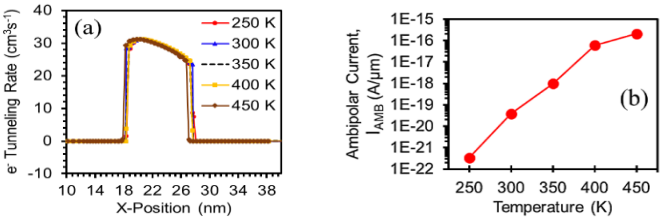



**Fig. 7.** Temperature dependence of (a) electron band to band tunneling rate and (b) ambipolar current conduction in H-JLTFET.

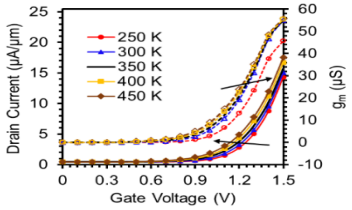



**Fig. 8.** Temperature dependence of drain current and $g_m$ in H-JLTFET.

Fig. 8 illustrates the temperature affectability of H-JLTFET on drain current (linear scale) and $g_m$ with gate voltage for temperature ranging from 250 K to 450 K. The drain current and $g_m$ shows similar variation with temperature increment from 250 K to 450 K. It has been evaluated that with an increase in temperature from 250K to 400K, the on-state current increases moderately (shown in Fig. 4(c)). Since $g_m$ represents the ratio of output current to input voltage while maintaining a constant drain bias. Therefore, $g_m$ of H-JLTFET is only 1.22 times higher in magnitude (from 45.7 μS to 55.7 μS) for temperature increment from 250 K to 450 K (Fig. 8).

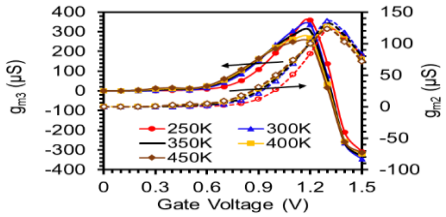



**Fig. 9.** Temperature dependence of $g_{m2}$ and $g_{m3}$ in H-JLTFET.

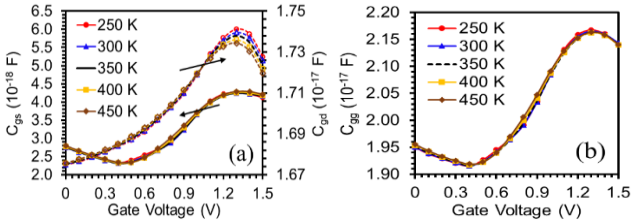



**Fig. 10.** Temperature dependence of parasitic capacitance (a) $C_{gd}$ and $C_{gs}$ (b) $C_{gg}$ in H-JLTFET.

### B. *Linearity*:

In linear applications like amplifiers, linearity is crucial to maintaining fidelity in signal transmission. Among the harmonic distortions due to nonlinearities of higher-order derivatives of $g_m$, the third order transconductance coefficient ($g_{m3}$) plays a more significant role than the second-order

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:53:21 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Nanotechnology. 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/TNANO.2025.3532313



transconductance coefficient ($g_{m2}$), as it primarily contributes to output signal distortions caused by non-linearity and frequency interference. For better linear performance the amplitude of $g_{m3}$ should be as low as possible. Fig. 8 represents the impact of temperature variations on the $g_{m2}$ and $g_{m3}$ of H-JLTFET w.r.t. gate voltage. The peak of both $g_{m2}$ and $g_{m3}$ gets lowered down with rise in temperature from 250 K to 450 K. Thus, H-JLTFET promises better linear performance and lesser signal distortion with the temperature increment from 250 K to 450 K.

### *C. RF Analysis*

In this subsection, the impression of temperature on the parasitic capacitances has been invetigated over a range of 250K to 450K. As depicted in Fig. 10(a), the parasitic capacitances $C_{gd}$ and $C_{gs}$ exhibit distinct temperature dependencies. In the subthreshold region, both $C_{gd}$ and $C_{gs}$ remain remarkably stable with varying temperature. However, upon transitioning to the ON state, a subtle decrease in Cgs is observed as temperature rises, while $C_{gd}$, the dominant contributor to total gate capacitance, remains virtually unaffected. At a reference temperature of 300K, $C_{gd}$ is measured to be 0.17 aF, whereas Cgs is found to be 4.17 aF. Within the ON state, a gradual increase in $C_{gs}$ of approximately 0.2 aF per 50K temperature increment is noted. The total gate capacitance, $C_{gg}$, has also analyzed and is presented in Fig. 10(b). Given the predominant influence of $C_{gd}$ on $C_{gg}$, its near-constant behavior with temperature translates to a similarly stable $C_{gg}$. This temperature insensitivity of parasitic capacitances underscores the robust nature of the proposed H-JLTFET. Notably, the peak total capacitance across the entire temperature range is determined to be 0.21 aF. Plot illustrated in Fig. 11(a) depicts the impression of temperature on the cut-off frequency ($f_T$) in GHz of H-JLTFET device. As the gate voltage increases, the device transitions from the subthreshold region to the on-state. This leads to a significant increase in the channel current and, consequently, a rise in $g_m$. As $f_T$ has direct reliance on $g_m$ thus it increases with $V_{gs}$. Further, the marginal dependence of $g_m$ and $C_{gg}$ on temperature results in marginal increase of $f_T$ with temperature. Similarly the maximum oscillation frequency also increases marginally with temperature for entire $V_{gs}$ range (Fig. 11(b)). Importantly, the H-JLTFET demonstrated exceptional high-frequency performance, achieving a peak $f_T$ of 417 GHz and an $f_{max}$ of 4390 GHz.

$$f_T = \frac{g_m}{2\pi(C_{gs}+C_{gd})} \qquad (2)$$

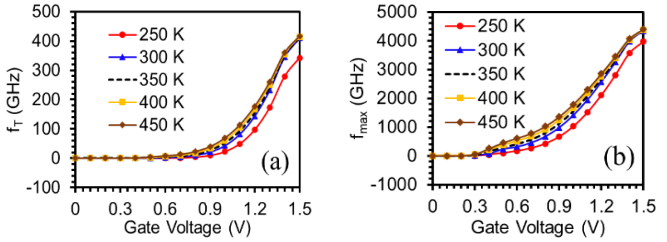



**Fig. 11.** Temperature dependence of (a) cut-off frequency and (b) maximum oscillation frequency in H-JLTFET.

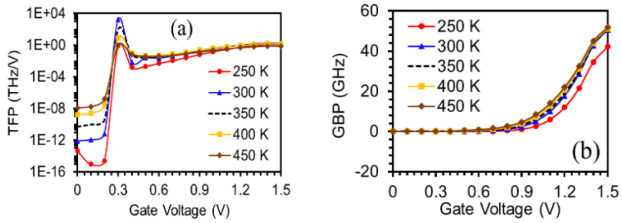



**Fig. 12.** Temperature dependence of (a) transconductance frequency product and (b) gain bandwidth product in H-JLTFET.

Fig. 12(a-b) illustrates the temperature dependence of two critical performance metrics in H-JLTFET devices: the transconductance frequency product (TFP) and the gain bandwidth product (GBP). TFP is a key parameter in FETs, especially for high-frequency applications, as it indicates the FET's bandwidth capability. A higher TFP means the FET can operate effectively at higher frequencies, enhancing performance in RF amplifiers and oscillators. In amplifiers, TFP directly influences the GBP, resulting in higher gain over a broader frequency range. Additionally, a higher TFP often correlates with better linearity, ensuring stable performance over wider frequencies. Overall, TFP is a critical figure of merit for assessing FET performance in various applications, from high-frequency RF systems to low-frequency linear circuits. Moreover, the GBP is crucial in FETs as it indicates the balance between gain and frequency range, essential for designing effective amplifiers. A higher GBP allows FETs to provide greater gain over a wider frequency spectrum, making them ideal for high-frequency applications like RF amplifiers and broadband systems. It also ensures better linearity and stability, maintaining signal integrity and minimizing distortion. Overall, GBP is a key performance metric, aiding in selecting suitable FETs and optimizing amplifier designs across various applications. Fig. 12(b) depicts the GBP against temperature variations for H-JLTFET. The behavior of the curve(s) here reveals the impact of temperature on the H-JLTFET bandwidth capabilities and overall gain is marginal. However, the increase is temperature increases the GBP of the H-JLTFET. Fig. 13 (a-b) displays the impact of temperature on intrinsic delay (evaluated using eq. 2) and intrinsic gain ($g_mR_o$) in H-JLTFET. The slight increase in $I_{ON}$ leads to reduction in intrinsic delay of the device with increase in temperature. It is examined that the intrinsic delay of the device reduces from 2.3ps to 1.9ps as the temperature increases from 250K to 450K. Further the intrinsic gain of the device remains unaltered to any change in temperature, while it increases as the applied gate voltage increases.

$$\tau = \frac{C_{gg}V_{dd}}{I_{ON}} \qquad (3)$$

TABLE II COMPARISON OF PROPOSED DEVICE WITH PUBLISHED TFET DESIGNS

| | UTB-DG-FE-TFET [30] | FE-TFET [30] | H-JLTFET (Present work) |
|---|---|---|---|
| $I_{ON}/I_{OFF}$ | 2.9 x $10^{13}$ | 1.2 x $10^{16}$ | 5.67 x $10^{14}$ |
| SS (mV/decade) | 23 | 27 | 13.2 |
| $g_m$ (μS) | 6.28 | 0.08 | 54.9 |
| $f_T$ (GHz) | 8.7 | 3.8 | 410 |

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:53:21 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Nanotechnology. 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/TNANO.2025.3532313



further the gain of the devices reduces with temperature at high $V_{gs}$ = 1.5V from 221 to 159 as temperature increases from 250K to 450K. The proposed H-JLTFET design is benchmarked against recently published TFET designs from the literature, as summarized in Table II. The comparison highlights the superior performance of our H-JLTFET, particularly under standard operating conditions at a temperature of 300 K. The results demonstrate significant advancements in key performance metrics, underscoring the effectiveness of our design in addressing the limitations of existing TFET architectures.

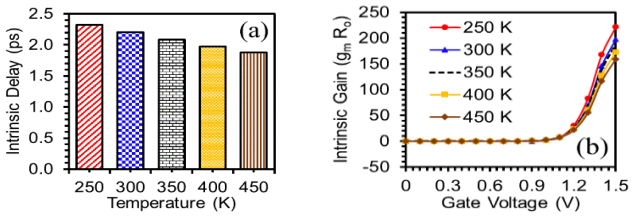



**Fig. 13.** Temperature dependence of (a) intrinsic delay and (b) intrinsic gain in H-JLTFET.

## IV. Conclusion

This research provides a comprehensive analysis of temperature-dependent performance in a novel polarity-controlled charge plasma-based InAs/AlGaSb tunneling interfaced junctionless TFET (H-JLTFET). The device exhibited a promising on-state current of 13.8 µA/µm at 250K, which increased to 17 µA/µm at 450K, demonstrating a moderate temperature sensitivity in the superthreshold region. However, the subthreshold leakage current increased significantly, with an order of magnitude increase from $10^{-20}$ A/µm to $10^{-16}$ A/µm over the same temperature range. The device showcased a peak $f_T$ of 417GHz and an $f_{max}$ of 4390GHz, indicating its potential for high-speed applications. While the intrinsic delay decreased from 2.3 ps to 1.8 ps with increasing temperature, the ambipolar current exhibited a detrimental increase. These findings underscore the need for careful consideration of temperature effects for optimal device performance and circuit design. Future work should focus on developing techniques to mitigate temperature-induced degradation and further explore the potential of H-JLTFETs for specific applications.

## References


[1] M. M. Waldrop, "The chips are down for Moore's law," *Nature News*, vol. 530, no. 7589, pp. 144, 2016.

[2] B. Hoefflinger, "IRDS–international roadmap for devices and systems, rebooting computing, S3S," *Nano-Chips 2030: On-Chip AI for an Efficient Data-Driven World*, pp. 9–17, 2020.

[3] S. Veeraraghavan and J. G. Fossum, "Short-channel effects in SOI MOSFETs," *IEEE Trans. Electron Devices*, vol. 36, no. 3, pp. 522–528, 1989.

[4] G. H. Jessen et al., "Short-channel effect limitations on high-frequency operation of AlGaN/GaN HEMTs for Tgate devices," *IEEE Trans. Electron Devices*, vol. 54, no. 10, pp. 2589–2597, 2007.

[5] K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand, "Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits," *Proc. IEEE*, vol. 91, no. 2, pp. 305–327, 2003.

[6] S. W. Sun and P. G. Tsui, "Limitation of CMOS supply-voltage scaling by MOSFET threshold-voltage variation," *IEEE J. Solid-State Circuits*, vol. 30, no. 8, pp. 947–949, 1995.

[7] M. Shrivastava, "Drain extended tunnel FET–a novel power transistor for RF and switching applications," *IEEE Trans. Electron Devices*, vol. 64, no. 2, pp. 481–487, 2016.

[8] Q. Huang et al., "A novel Si tunnel FET with 36mV/dec subthreshold slope based on junction depleted-modulation through striped gate configuration," in *Proc. IEEE Int. Electron Devices Meeting*, 2012, pp. 8–5.

[9] A. C. Seabaugh and Q. Zhang, "Low-voltage tunnel transistors for beyond CMOS logic," *Proc. IEEE*, vol. 98, no. 12, pp. 2095–2110, 2010.

[10] S. Sharma and R. Chaujar, "RF, linearity and intermodulation distortion analysis with small-signal parameters extraction of tunable bandgap arsenide/antimonide tunneling interfaced JLTFET," *Microsyst. Technol.*, vol. 28, no. 12, pp. 2659-2667, 2022.

[11] S. Sharma, J. Madan, and R. Chaujar, "Interfacial charge associated reliability improvement in arsenide/antimonide tunneling interfaced-junctionless TFET," *Phys. Scr.*, 2024.

[12] S. Sharma and R. Chaujar, "Impact of tunnel gate process variations on analog/radio frequency (microwave) and small signal parameters of hetero-material tunneling interfaced charge plasma junctionless tunnel field effect transistor," *Int. J. Circuit Theory Appl.*, vol. 50, no. 10, pp. 3626-3641, 2022.

[13] M. R. Tripathy et al., "Device and circuit-level assessment of GaSb/Si hetero-junction vertical Tunnel-FET for low-power applications," *IEEE Trans. Electron Devices*, vol. 67, no. 3, 2020.

[14] D. B. Abdi and M. J. Kumar, "In built N+ pocket p-n-p-n tunnel field-effect transistor," *IEEE Electron Device Lett.*, vol. 35, no. 12, pp. 1170–1172, 2014.

[15] S. Sharma and R. Chaujar, "Influence of source electrode metal work function on polar gate prompted source hole plasma in arsenide/antimonide tunneling interfaced junctionless TFET," *J. Micromech. Microeng.*, vol. 32, no. 4, p. 044004, 2022.

[16] S. Sharma and R. Chaujar, "Performance enhancement in a novel amalgamation of arsenide/antimonide tunneling interface with charge plasma junctionless-TFET," *AEU-Int. J. Electron. Commun.*, vol. 133, p. 153669, 2021.

[17] A. E. Islam, "Current status of reliability in extended and beyond CMOS devices," *IEEE Trans. Mater. Device Rel.*, vol. 16, no. 4, pp. 647–666, 2016.

[18] S. Datta, H. Liu, and V. Narayanan, "Tunnel FET technology: A reliability perspective," *Microelectron. Rel.*, vol. 54, no. 5, pp. 861–874, May 2014.

[19] Y. P. Varshni, "Temperature dependence of the energy gap in semiconductors," *Physica*, vol. 34, no. 1, pp. 149–154, 1967.

[20] R. Narang et al., "Impact of temperature variations on the device and circuit performance of tunnel FET: A simulation study," *IEEE Trans. Nanotechnol.*, vol. 12, no. 6, pp. 951–957, 2013.

[21] M. Neisser, "International roadmap for devices and systems lithograph roadmap," *J. Micro/Nanopattern. Mater. Metrol.*, vol. 20, 2021.

[22] ATLAS User's Guide, SILVACO Int., Santa Clara, CA, USA, Version 5.26.1.R, 2018.

[23] W. Hänsch et al., "Carrier transport near the Si/SiO2 interface of a MOSFET," *Solid-State Electron.*, vol. 32, no. 10, pp. 839-849, 1989.

[24] Y. Zhao et al., "A novel tunnel FET design through adaptive bandgap engineering with constant sub-threshold slope over 5 decades of current and high ION/IOFF ratio," *IEEE Electron Device Lett.*, vol. 38, no. 5, pp. 540-543, 2017.

[25] G. Zhou, Y. Lu, R. Li, Q. Zhang, W. Hwang, Q. Liu, T. Vasen, *et al.*, "Self-aligned InAs/Al0.45_{0.45}0.45Ga0.55_{0.55}0.55Sb vertical tunnel FETs," in *Proc. 69th Device Research Conf.*, 2011, pp. 205–206.

[26] K. Y. Na and Y. S. Kim, "Silicon complementary metal–oxide–semiconductor field-effect transistors with dual work function gate," *Jpn. J. Appl. Phys.*, vol. 45, no. 12, p. 9033, 2006.

[27] C. H. Wang, H. Chu, Y. S. Lai, *et al.*, "Dual work-function metal gates," *U.S. Patent 73 81619B2*, Jun. 3, 2008.

[28] R. W. Johnson, A. Hultqvist, and S. F. Bent, "A brief review of atomic layer deposition: from fundamentals to applications," *Mater. Today*, vol. 17, no. 5, pp. 236–246, 2015.

[29] V. Brouzet, B. Salem, P. Periwal, G. Rosaz, T. Baron, F. Bassani, P. Gentile, and G. Ghibaudo, "Fabrication and characterization of silicon nanowire p-i-n MOS gated diode for use as p-type tunnel FET," *Appl. Phys. A*, vol. 121, no. 3, pp. 1285–1290, 2015.

[30] G. Gopal and T. Varma, "Impact of temperature on the reliability of UTB-DG-FE-TFETs and their RF/analog and linearity parameter dependence," *J. Electron. Mater.*, vol. 52, no. 9, pp. 6293–6307, Sep. 2023.


Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:53:21 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

Int J Syst Assur Eng Manag
https://doi.org/10.1007/s13198-024-02702-3

ORIGINAL ARTICLE

# Modeling the enablers for sustainable adoption of water ionizers: an ISM-MICMAC approach

**Saurabh Pandey**[1] · **Deepti Aggrawal**[1]

Received: 6 September 2024 / Revised: 22 December 2024 / Accepted: 30 December 2024
© 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** Literature is evident of research work pertaining to water sustainability. Plenty of ways have been described in the past wherein energy efficient and water quality devices have gained good consideration of researchers. Amongst this theme, water ionizers have fetched significant attention in recent years due to their potential benefits for human health and the environment. The objective of research on water ionization technology, sustainable practices, and consumer behaviour is to empirically analyse and clarify the factors that impact the consumers' perception and adoption of water ionizers. Firstly, this study explores the enablers for the sustainable adoption of water ionizers using an Interpretive Structural Modeling (ISM) approach establishing contextual relationships among these factors, which results in hierarchical structure in the form of digraph. Secondly, MICMAC analysis approach to categorize and identifying factors based on driving power and dependence power aiming to construct a comprehensive framework that highlights the critical variables influencing the adoption process. This dual approach not only provides valuable insights into the dynamics of water ionizer adoption but also facilitates the development of targeted strategies to promote or mitigate specific influences. The findings of this study are expected to contribute significantly to the understanding of sustainable technology adoption and offer practical implications for stakeholders aiming to enhance the acceptance and utilization of water ionizers. Questionnaires are thoughtfully crafted to gather precise responses, ultimately providing valuable insights for manufacturers and policymakers. The research hypothesizes that the ISM-MICMAC model effectively identifies and emphasizes product features and environmental sustainability as critical factors contributing to the health benefits that drive the sustainable adoption of water ionizers.

**Keywords** Interpretive structural modeling · MICMAC analysis · Sustainable adoption · Water ionizers

## 1 Introduction

Humans have consistently evolved in their pursuit of better living conditions, greater technological advancements, and improved health outcomes. Food safety has profound social significance (Lin et al. 2019) for people's health. Drinking plenty of water to maintain a balanced diet is essential for building immunity and overall health. Nutrient supplements and the consumption of healthful foods are also essential. More than 80% of our body is water; as we age, this proportion declines (Lu et al. 2023). Although we may acknowledge the significance of water in the human organism, its intrinsic characteristics to its diverse roles, which encompass hydration, circulation, digestion, thermoregulation, dermal health, and excretion, are critical. Among these, hydration (Seal et al. 2019) plays a vital role in maintaining the retention of water concentration within human organs.

In recent years, after COVID-19, every individual has become more focused on the pursuit of good health and well-being, and this intensified consciousness has sparked an interest in multiple innovative technologies that can provide potential health benefits (Henry and Chambron 2013). Water ionizers alter the pH level of water, producing both

✉ Deepti Aggrawal
deeptiaggrawal@dtu.ac.in

Saurabh Pandey
saurabhr22003@gmail.com

[1] University School of Management and Entrepreneurship (USME), Delhi Technological University, Delhi 110095, India

Published online: 09 January 2025

alkaline and acidic water streams through the process of electrolysis (The Electrolysis Process n.d.). Due to the high pH value and low surface tension of electrolyzed reduced water (ERW), the virus's envelope, made up of proteins and phospholipids, breaks down and exhibits inhibitory activity against the Severe Acute Respiratory Syndrome Coronavirus 2 (Okajima et al. 2021). Alkaline water is believed to possess health-enhancing properties due to its alkaline nature and high antioxidant potential. The creation of hydrogen-rich water is another important characteristic of water ionizers. As a potent antioxidant, hydrogen is a strong antioxidant that might reduce oxidative stress, which could delay aging and increase energy (Pisoschi et al. 2021). The technical and scientific community responds after the innovation by creating standards to facilitate industry and economic activity (Aggrawal et al. 2017) and lead to the adoption process by customers. Water ionizers have become a practical option for supplying ionized, alkaline water with claimed health benefits (Henry and Chambron 2013). They work by electrolyzing water to change its pH level. Also, one critical area of research that must be done is understanding how customers accept and adopt these technologies within the context of sustainability.

In pursuing a healthy lifestyle, society has changed its perseverance towards the environment and sustainability (Vacchi et al. 2021) of any product consumed or adopted. Water ionizers employ high-quality, eco-friendly materials to enhance longevity and mitigate their ecological footprint. Consumer perception toward the sustainable adoption (Silva 2023) of innovative technological products involves a complex interplay of different factors, including technological, economic, social, and sustainability dimensions, and adopting products or technologies is often slow by the end users (Crosno and Cui 2014; Pandey and Aggrawal 2024). Marketing professionals, confronted with constantly shifting market conditions, continually seek to develop new products and make changes of varying degrees, from minor to major, or a series of changes spread out over a predetermined period, which may even lead to the creation of entirely new offerings of goods and services. Organizations frequently must determine that how quickly the consumers will adopt a technologically enhanced product (Anand et al. 2019). Managers, policymakers, and researchers must comprehend what happens when innovations are embraced by individual customers and spread in the market (Kristensson et al. 2020). In terms of the creation, uptake, and dissemination of new products, innovators are suggested as a valuable resource for marketers, exhibit stronger domain-specific innovativeness, and view new technologies as less "complex," which may make them more likely to be adopted (Schreier et al. 2007). The existing investigation intends to confine the components contributing to water ionizer customer acceptance, define interrelationships between them, and use the ISM approach to build a hierarchical theoretical framework, a powerful tool for modeling complex systems. The framework to organize factors for MICMAC analysis, aids in identifying key variables that influence the adoption process, enabling targeted strategies to encourage or mitigate specific influences based on their categorized impact into dependence and driving power.

This combined ISM-MICMAC approach offers a subtle understanding of the dynamic interdependencies among influencing factors, providing insights into their hierarchical structures and mutual influences. The details of the article's structure are explained in the following section. This literature provides insights in details on water ionizers, health benefits and sustainability as well as the concept of ISM-MICMAC, are presented in Sect. 2. Section 3 presents a discussion of the research methodology, the variables affecting the uptake of these technologies, and the ISM-MICMAC technique used in this study. Section 4 explains the details of ISM model development and analysis, and consequently, MICMAC analysis is discussed in Sect. 5. The discussion and managerial implications are presented in Sects. 6 and 7, followed by the conclusion, study limits, and future scope of the work to provide directions on the consumer's views of sustainable adoption of water ionizers.

## 2 Literature review and building blocks of the present study

The existing literature explores scholarly research on the water technology industry and its market potential of penetration to examine individual discernments concerning the embracing of such technologies by community. This existing literature research dissects various reasons of acceptance which include sustainability, overall well-being for health benefits, and perceived cost-effectiveness that how their interconnections impact customers' decisions for acceptance of water ionizers within the context of sustainable living practices. The research study while aligning individual actions toward the acceptance of sustainable practices, considering healthy society, and products innovation in water industry to investigate the key determinants driving the adoption.

### 2.1 Sustainability

US Sustainable directs the development Growth Report 2023 (M 2023), Goal 6 ensures water availability and sustainable management. Various factors and drivers, such as personal, societal, situational, and demographic aspects, environmentally sustainable impact behavior (Nguyen and Johnson 2020). Water ionizers epitomize a promising possibility for promoting healthy water consumption, and this study

identifies the key indicators of water ionizers for the acceptance by individual or community. Human health and well-being frequently align with sustainable practices (Ronen and Kerret 2020). To analyze the adoption of water ionizers, it is necessary to comprehend the relationship between health beliefs and sustainable consumption patterns (Tanner and Wölfing Kast 2003). Also, a vital component of sustainable consumer behavior (Trudel 2019) is belief in the efficiency and dependability of technology. Most water ionizers from good brands are ISO 14001 certified (ISO 14001:2015 2022), which authorizes the product for best practices and strategies for environmental management (EMS) and leads toward environmental sustainability. The use of water from ionizers will reduce carbon footprint of bottled water significantly as plastic bottles are not recommended due to its ionization behaviour and should be consumed directly from tap. The results indicate that tap water has about 300 times lower carbon footprint values than bottled water (Botto 2009). The production and transportation of bottled water require substantial energy. The Pacific Institute estimates that producing bottled water consumes up to 2000 times more energy than tap water (Gleick and Cooley 2009). Adoption of green technologies is prioritized (Usmani et al. 2023), emphasizing lowering energy consumption and negative environmental impact to provide sustainable solutions. EMS is dedicated to innovating best practices for environmentally sustainable products, such as water ionizers which can be assessed through multiple perspectives, including their environmental impact, energy efficiency, and long-term economic viability (Pandey and Aggrawal 2024). With customers' demand for sustainable products and regulatory obligations on organizations, sustainability is an extreme goal regarding raw material, innovation, product, and process design (Kumar et al. 2020). Product awareness among consumers and sustainability labeling impact consumers' perceptions of products, their desire to behave, and their purchasing behaviors (Majer et al. 2022; Joshi and Rahman 2017).

### 2.2 Ionization, health benefits, and usage

Human body parts predominantly consist of water in it e.g. the brain consists of 80–85% water, the heart 70–80%, etc. (water percentage of the-human body 2020). The percentage of water in the human body decreases (Pocock et al. 1981) with aging. Alkalize ionized water (AIW) depicts critical properties such as alkalinity, negative Oxidation–Reduction Potential (ORP), Micro-clustering, and a high content of dissolved hydrogen, which distinguish it from regular tap or reverse osmosis (RO) water (Ignacio et al. 2012). A quantitative analytical study is carried out to explore the distinctive traits associated with the adoption of water ionizers in community and to specify the factors that influence individuals' findings to use these devices.

AIW produced by water ionizers provides various potential health benefits, detoxification, including enhanced hydration, antioxidant effects, augmented alkalinity, and improved digestion. It is shown that the ERW or AIW suppress oxidative stress-related illnesses as arteriosclerosis, diabetes, cancer, neurodegenerative diseases, and the side effects of hemodialysis (Shirahata et al. 1997). After consuming this water for six days, blood glucose and HbA1c levels dramatically dropped, which continued to drop during prolonged consumption (Shirahata et al. 2012). Studies have demonstrated that alkaline ionized water protects against the accumulation of lipids and cholesterol in the body (Jin et al. 2006). Electrolyzed reduced water (ERW) causes telomere shortening in cancer cells (Shirahata et al. 2002). It suppresses tumor angiogenesis by scavenging intracellular reactive oxygen species (ROS) and suppressing secretion of vascular endothelial growth factor and the gene expression and Ye et al. (2008). According to a mathematical analysis, water ionizers are highly effective for the prophylaxis of the development of cancer cells (Ignatov et al. 2016) by influencing the chemical reorganization of H2O molecules. ERW suppresses microorganisms and the growth of cancer cells and (Hamasaki et al. 2005) and induces apoptosis with glutathione in human leukemia HL60 cells (Tsai et al. 2009).

Further studies on ionized water revealed that the bandwidth of ERW (LeBaron et al. 2022) and depicts a type of rearrangement of the ionized water, specifically the presence of a smaller micro-cluster (Hayashi 1995). It is often claimed that ionized water carries fewer micro-clusters than regular tap water, which helps penetrate water molecules inside human organs (Mohr 2016). The higher pH ionizes water molecules, increasing their solubility and subsequent extraction (Kun-Kul 2005), which contributes to the removal of oil-based pesticides from foods and claims that microclustering is responsible for the ability of this high-pH water to emulsify the oil. As a sanitizer to reduce or eliminate bacterial populations on surfaces used in food processing, non-food contact areas, and food items, electrolyzed water (EW) is gaining popularity in the food industry (Hricova et al. 2008). The Health and Welfare Ministryin Japan has approved ERW as a food additive (Yoshida et al. 2004). The high pH alters the properties of normal water and produces a negative oxidation–reduction potential (ORP). ORP measures the antioxidants in water by checking out its electrical conductivity as unsafe radicals due to the valuable antioxidants that resist diseases like cancer, cardiovascular disease, and immune disorders (Ignatov 2017).

### 2.3 ISM-MICMAC approach

ISM combines different or related necessities into an easy-to-understand and well-organized framework. Therefore, it is a methodology that looks for links between the several

factors taken into consideration and gives them a multilevel, hierarchical framework (Lin et al. 2019) of water ionizers. The ISM methodology was developed by Warfield (1974). The mutual interaction of enabler identification and analysis in total productive maintenance was studied by Attri and Grover (2018). Anand and Bansal (2017) discussed an analysis of software reliability and sustainability attributes. The following phases are considered in the implementation of ISM:

- Investigation of business enablers or attributes for analysis to achieve the goal.
- We are ascertaining the relative relationships between enablers.
- Building structural self-interaction matrix (SSIM)
- Conversion of SSIM to the initial reachability matrix (IRM)
- Interpretation of the IRM into the final reachability matrix (FRM)
- A pictorial representation for the analysis of the ISM

Further, MICMAC analysis is considered to obtain enablers' driving and dependence power as the subsequent steps of ISM. Michel Godet and François Bourse developed the MICMAC method. It is a structural prospective analysis used to investigate indirect correlations and is a cross-impact matrix multiplication. This approach aims to determine the critical components that contribute to the system's development into four distinct segments: independent, dependent, linkage, and autonomous variables. ISM-MICMAC is used to investigate how essential enablers interact to remove obstacles in the way of a product's entry into a potential market (Ullah and Narain 2021). The highlights of the literature review are summarised in Table 1.

## 3 Research methodology

The utility of ISM methodology can be understood from Fig. 1, which shows the high-level framework used in this study to assess critical factors. The ISM technique has been used to identify and summarize the relationships between the various variables. Therefore, this methodology provides a clear picture of complicated real-life issues so that researchers, managers, and individuals may evaluate the findings from ISM and MICMAC and take appropriate action (Anand et al. 2019). ISM is primarily an interpretive learning process that portrays variable order and transforms it into an organized framework (Mathiyazhagan et al. 2013). It is a qualitative approach that transforms complex articulated structural models into well-defined conceptual models illustrating interrelationships among variables (Gupta and

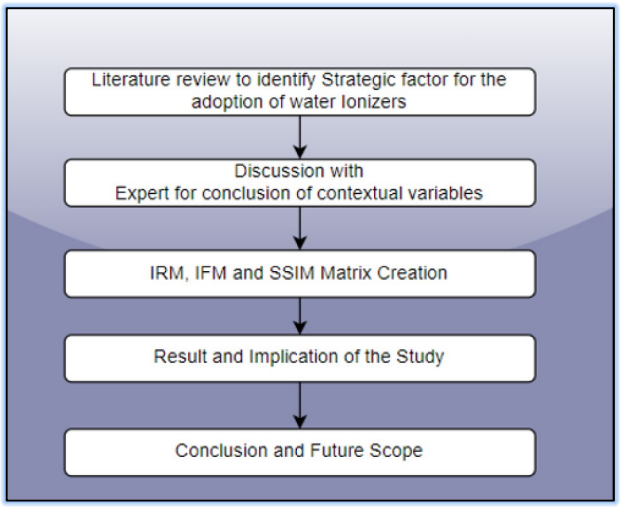



**Fig. 1** High-level methodology. *Source*: authors

**Table 1** Studies on ERW and sustainable practices

| Author(s) | Methodology | Key findings |
|---|---|---|
| Shirahata et al. (2002) | ERW Health Benefits | Cancer cells' telomeres shorten due to ERW, among other health advantages |
| Tanner and Wölfing Kast (2003) | Sustainability practice | Understanding the connection between sustainable consumption patterns and health beliefs is vital to assessing the adoption of water ionizers |
| Nguyen and Johnson (2020) | Environmental sustainability | Various variables and drivers, including environmental sustainability and features of the individual, society, environment, and demography, influence behavior |
| Ronen and Kerret (2020) | Sustainability practice | Sustainable methods promote human health and well-being |
| LeBaron et al. (2022) | Water technology | The ERW has a narrower bandwidth than tap water and is approximately 170 nuclear magnetic resonance sensitive. It also defines the other valuable properties of ERW water |
| Usmani et al. (2023) | Green technology | The adoption of green technologies is prioritized, emphasizing the reduction of energy consumption and negative environmental impact |

*Source*: Authors

Sahu 2013; Sushil 2012). It adheres to a defined methodology, which is outlined below:

- *Creating the structured self-interaction matrix (SSIM) and defining the variables*: ISM begins with identifying variables from the literature relevant to the research agenda or challenge. After defining a set of variables, an expert group prepares a pairwise association matrix (Singh and Samuel 2018; Sindhwani et al. 2023).
- *Reachability Matrix*: To obtain the original reachability matrix, the SSIM findings are further substituted with binary integers. Subsequently, the ultimate reachability matrix is obtained by employing the transitivity technique. Transitivity demonstrates that X is unquestionably associated with Z if X is related to Y and Y is related to Z.
- *Level Partitioning*: The outcomes are after that divided into discrete levels based on the reachability matrix.
- *ISM Model*: A digraph, a structured model of the suggested investigation, is created by transfiguring the outcomes of partitioned levels.

Multiple rounds of discussion with a focused group involved in the in-person or commercial use of water ionizers signify the below enabler to evaluate consumers' adoption intentions. The study determined fifteen factors that influence consumers' adoption behavior toward water ionizers. Table 2 explains the details of the key enables required for consumers' decision-making regarding water ionizer adoption.

### 3.1 Data and sample size

Primary data was collected to assess client opinions based on their unique experiences and accompanying documentation. Fifteen industry professionals from disciplines like marketing, information technology, healthcare, and operation management participated in our study. The questionnaire was discussed in person to get the consumers' insightful answers in the format required to create the ISM matrix. The ISM approach suggests using subject-matter expert opinion based on different management strategies, including brainstorming and interviews, to establish the contextual interaction between enablers and, according to Murry Jr and Hammons (1995), the ideal number of qualified experts for qualitative decision-making research could range from 10 to 30.

### 3.2 Development and analysis of the ISM model

The initial stage of the ISM procedure (Patra et al. 2021) involves identifying variables and establishing contextual connections among them. It requires consultation with experts to elucidate variables and their interrelationships. Figure 2 depicts the step-by-step process used to develop the ISM model. This approach leverages the expertise and practical insights of professionals. Various methodologies are available to identify contextual variables within this process, including workshops, nominal group procedures, idea engineering workshops, open brainstorming sessions, questionnaires, discussions among expert panels, and literature surveys.

*Structured Self-Interaction Matrix (SSIM):* Generating the SSIM involves comparing each row to each column and assigning a code based on the predetermined set (V, A, X, O). The ISM methodology uses Expert judgment to determine the contextual relationship among variables. The following symbolic designations indicate how the variables are interrelated:

V: variable *i* influences variable *j*.
A: variable *j* influences variable *i*.
X: variables *i* and *j* influence each other, and
O: variables *i* and *j* are unrelated.

According to the earlier stated principles, the SSIM matrix is developed from collected data as shown in Fig. 3,

- *Reachability Matrix*: The SSIM is converted into the initial reachability matrix (IRM) as shown in Fig. 4. IRM is described as a binary matrix, with codes replaced by 0 s and 1 s. The conversion from SSIM to IRM follows these rules.
- SSIM cell (*i, j*): The 'V' code is converted to 1, and then IRM value changed to the (*j, i*) to a value becomes "0".
- SSIM cell (*i, j*): The 'A' code is converted to 0 and then IRM value changed to (*j, i*) to a value becomes "1".
- SSIM cell (*i, j*): The 'X' code is converted to 1 and then IRM value changed to (*j, i*) to a value become "1".
- SSIM cell (*i, j*): The 'O' code is converted to 0 and then IRM value change to (*j, i*) to a value become "0".
- *Final Reachability Matrix (FRM):* Using the ISM function in the R studio software (Peng 2015), the IRM is then transformed into the final reachability matrix (FRM) by removing transitivity using predetermined methods. Any values in the resultant matrix greater than 1 are then set back to 1. Until transitivity is attained, this iterative procedure is continued, producing the FRM shown in Fig. 5. All the 1's shown in the cells shaded green in Fig. 5 are the result of the transitivity operation performed. For example, the enablers "Social Acceptance" and "Perceived Value (Cost)" have the value 0 in IRM (Fig. 4), but "Social Acceptance" define the enabler "Brand Value" and "Brand Value" derive the enabler "Perceived Value (Cost)", Hence, "Social Acceptance" is also related "Perceived Value (Cost)" and the value

**Table 2** Enablers description

| Key enablers | Description | References support |
|---|---|---|
| Perceived Value (Cost) | Individual evaluation of the costs and benefits involved in adopting and using a product | Consumers' analysis of product value and cost influences sustainable adoption (Jurison 2000) |
| Health Benefits | The adoption of consumer products can positively affect physical or overall well-being by improvising health of an individual | Significant health benefits to consumers for influence positively in adoption (Rubik 2011) |
| Brand Value | A product or brand's market position, repute, recognition, and customer loyalty are all part of its intangible and financial worth | Brand value positions the product in consumer consensus and influences satisfaction (Agarwal et al. 2017) |
| Maintenance and Ease of Use | Explaining the simple, trouble-free operation of long-term, minimally implied-cost sustainable water ionisers | This makes the consumer aware of the technological product for adoption. (Sugandini et al. 2018) |
| pH Level | The pH scale measures the alkalinity or acidity of water, reflecting the relative concentration of hydrogen ions considering production of water from very strong acidic i.e. 2.5 to acidic 4–6 pH and further to slightly alkaline i.e. 8 or 8.5 to strongly alkaline 9.5 and above, which is based on the applications or use of water at various industrial or domestic level | Product features that improve the overall use of water and reveal multi-facet applications in consumers' daily lives for holistic health benefits (Rubik 2011) |
| Oxidation–Reduction Potential (ORP) | Assessing the ionized potential of water, which shows its ability to donate or accept electrons, determines whether it can act as an oxidizing agent or antioxidant | |
| Micro-Clustering | Phenomena of cluster formation process by which water molecules group together to create smaller clusters, which improve hydration and absorption at the cellular level | |
| Social Acceptance | Social and Environmental factors which impact consumer behavior, reflecting various perspectives on personal choice. Precisely, evaluating acceptance, valued, and used by individuals and groups as a preferred and socially acceptable option for drinking water | Socially accepted product gains tend to increase when recommended by customers, and vice versa (AlSaleh and Thakur 2019) |
| After Sale Service | Installation Support, Product Training, Maintenance Services, Repair Services, Customer Support, Warranty Coverage and Replacement Parts include ASS | New product adoption and sales performance from the adoption perspective (Li and Lin 2015) |
| Environmental Sustainability | Ensuring products and their associated processes are produced, used, and disposed of in a manner that supports long-term ecological balance, promotes resource conservation, and reduces adverse environmental impacts | Environmental sustainability is vital in innovating newly developed technological products (Fu et al. 2018) |
| Customer Recommendation | Consumer experiences and testimonials regarding the health benefits and other uses of water ionizers provide practical discernment for potential buyers considering these devices | Socially accepted product gains tend to increase when recommended by customers, and vice versa (AlSaleh and Thakur 2019) |
| Adaptability or Openness to Change | Establishing values that direct behaviour and influence judgements about the adoption of innovation is a crucial phase in the adoption process | It urges consumers to product adoption (Li et al. 2017) |
| Adoption Intention of Water Ionizer | Subjective probability of user engagement concerning the adoption of an innovative product refers to an individual's personal assessment or belief about the likelihood that users will engage with or adopt the new product | |
| Perceived usefulness | The possible health benefits of drinking, the commercial uses of ionizers for cleaning medical devices, and the usage of ionizers for cleaning and sanitisation in domestic settings are factors that influence how valuable and beneficial an ionizer is seen by its users | This makes the consumer aware of the technological product, and product usefulness influences adoption. (Sugandini et al. 2018) |

**Table 2** (continued)

| Key enablers | Description | References support |
|---|---|---|
| Compliance Standard | Water ionizer systems and products are required to comply with established standards, guidelines, and regulations related to safety, performance, and legal requirements. This includes adherence to quality standards such as ISO 13485, ISO 14001, and ISO 9001, among others | Compliance standards build trust for product adoption (Villegas-Palacio and Coria and Villegas-palacio 2014) |

*Source* authors

is updated to 1 after removing the transitivity in FRM (Fig. 5).

- A8 → A3 and A3 → A1,
  - So, A8 → A1
- *Level Partitions (LP)*: The FRM establishes the antecedent and reachability set for each key recognized factor. Each factor's reachability set includes additional criteria that might make achieving it easier. Similarly, a factor's antecedent set is made up of the factor itself as well as potential contributors. For every factor, intersection sets are also calculated. Figure 6 represents iteration-wise level partitioning.

In the ISM hierarchical architecture, factors with the same intersection and reachability sets are ranked highest. When a factor reaches its highest level, it does not aid in accomplishing any other factor beyond its tier. As a result, these superior components are recognized, set apart, and kept separate from other variables. Until all factors reach a particular level, this process is repeated for progressively higher-tier factors at lesser levels. These designated levels are the foundation for generating the final ISM model and digraph.

## 4 MICMAC analysis

A primary method rooted in the properties of matrix multiplication (Nandal et al. 2019) is employed, known as the MICMAC analysis. The examination of the driving and dependent powers of variables, this method aids in categorizing and identifying key variables, which clarifies the factors influencing the adoption of water ionizers. Using the concepts of driving and dependence power, the factors in this study are classified as follows:

- *Autonomous variables:* Situated in the first quadrant, these variables exhibit low dependence and driving power, signifying complete independence and minimal association within the system.
- *Dependent variables:* Situated in the second quadrant, these variables demonstrate high dependence and low driving power.
- *Linkage variable:* Located in the third quadrant, these variables exhibit both high dependence and high driving power. They influence other variables while also being significantly affected by them.
- *Independent variables:* These variables showcase low dependence and high driving power.

Using these concepts, a MICMAC graph (Fig. 7) is created, with the driving power of the variable indicated by the y-axis and dependence power represented by the x-axis. This

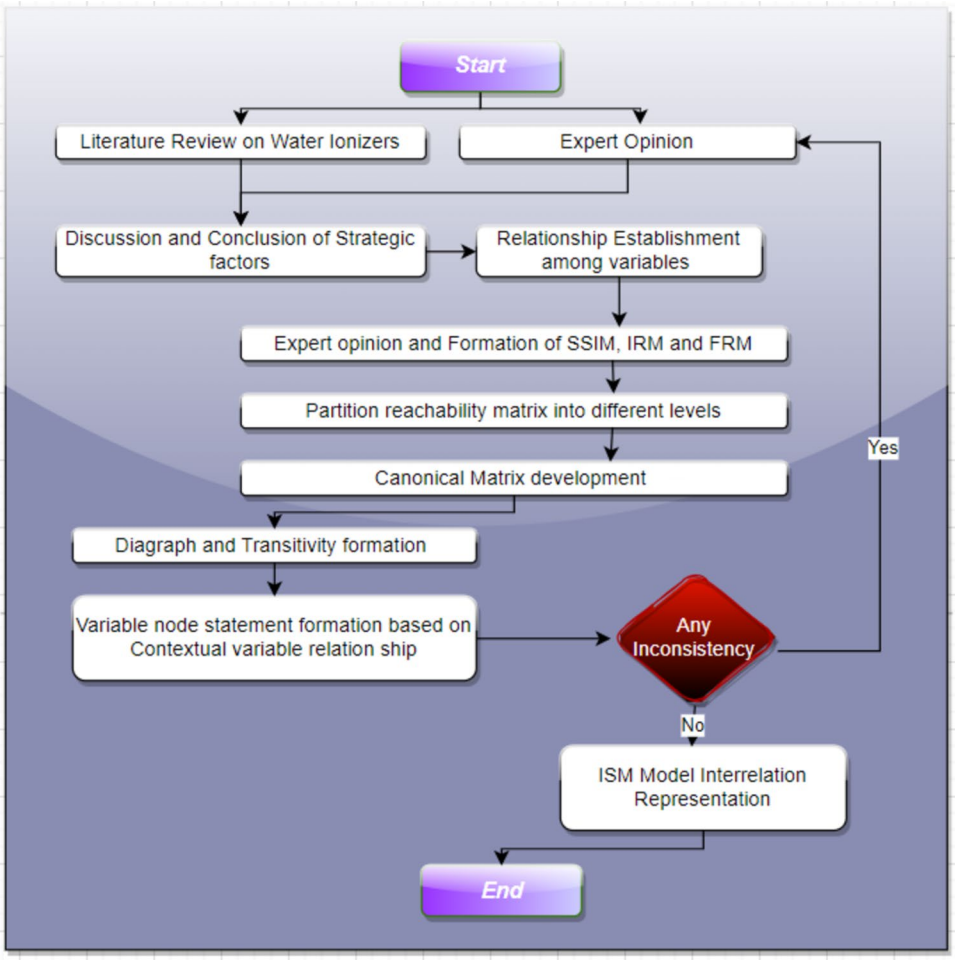



**Fig. 2** Flow diagram for ISM. *Source:* authors

| Enablers | A15 | A14 | A13 | A12 | A11 | A10 | A9 | A8 | A7 | A6 | A5 | A4 | A3 | A2 | A1 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A1.Perceived Value (Cost) | X | A | V | V | V | A | A | V | A | A | A | A | X | X | |
| A2.Health Benefits | 0 | V | V | V | V | O | O | V | A | A | A | O | V | | |
| A3.Brand Value | A | A | V | V | X | A | A | X | A | A | A | A | | | |
| A4.Maintenance and Ease of Use | V | V | V | V | V | O | O | V | O | O | O | | | | |
| A5.pH Level | V | X | V | V | V | O | O | V | X | X | | | | | |
| A6.Oxidation-Reduction Potential (ORP) | V | X | V | V | V | O | O | V | X | | | | | | |
| A7.Micro-Clustering | V | X | V | V | V | O | O | V | | | | | | | |
| A8.Social Acceptance | A | A | A | X | A | A | A | | | | | | | | |
| A9.After Sale service | V | O | V | V | V | O | | | | | | | | | |
| A10.Environmental Sustainability | V | A | V | V | O | | | | | | | | | | |
| A11.Customer Recommendation | A | A | V | V | | | | | | | | | | | |
| A12.Adaptability or Openness to Change | A | A | V | | | | | | | | | | | | |
| A13.Adoption Intention of Water Ionizer | A | A | | | | | | | | | | | | | |
| A14.Perceived usefulness | V | | | | | | | | | | | | | | |
| A15.Compliance Standard | | | | | | | | | | | | | | | |

**Fig. 3** Structural self-interaction matrix (SSIM). *Source*: authors

graph not only forms a conceptual model but also defines the hierarchy of variables.

Plotting the MICMAC analysis reveals how different elements influence consumer perception and aid in decision-making based on their driving and reliance power. The high importance on factors influencing consumers' decisions include pH value, perceived utility, micro-clustering, ORP, and ease of maintenance. Conversely, the factors most

| Enablers | A1 | A2 | A3 | A4 | A5 | A6 | A7 | A8 | A9 | A10 | A11 | A12 | A13 | A14 | A15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A1.Perceived Value (Cost) | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 |
| A2.Health Benefits | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 |
| A3.Brand Value | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| A4.Maintenance and Ease of Use | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| A5.pH Level | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| A6.Oxidation-Reduction Potential (ORP) | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| A7.Micro-Clustering | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 |
| A8.Social Acceptance | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| A9.After Sale service | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 |
| A10.Environmental Sustainability | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| A11.Customer Recommandation | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| A12.Adaptability or Openness to Change | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| A13.Adoption Intention of Water Ionizer | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| A14.Perceived usefulness | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 |
| A15.Compliance Standard | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 |

**Fig. 4** Initial reachability Matrix (IRM). *Source*: authors

| Enablers | A1 | A2 | A3 | A4 | A5 | A6 | A7 | A8 | A9 | A10 | A11 | A12 | A13 | A14 | A15 | Driving Power |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A1.Perceived Value (Cost) | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | *1* | 1 | 9 |
| A2.Health Benefits | 1 | 1 | 1 | 0 | *1* | *1* | *1* | 1 | 0 | *1* | 1 | 1 | 1 | 1 | *1* | 13 |
| A3.Brand Value | 1 | *1* | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | *1* | 8 |
| A4.Maintenance and Ease of Use | 1 | *1* | 1 | 1 | *1* | *1* | *1* | 1 | 0 | *1* | 1 | 1 | 1 | 1 | 1 | 14 |
| A5.pH Level | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | *1* | 1 | 1 | 1 | 1 | 1 | 13 |
| A6.Oxidation-Reduction Potential (ORP) | 1 | 1 | 1 | 0 | *1* | 1 | 1 | 1 | 0 | *1* | 1 | 1 | 1 | 1 | 1 | 13 |
| A7.Micro-Clustering | 1 | 1 | 1 | 0 | *1* | 1 | 1 | 1 | 0 | *1* | 1 | 1 | 1 | 1 | 1 | 13 |
| A8.Social Acceptance | *1* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | *1* | 1 | *1* | 0 | 0 | 6 |
| A9.After Sale service | 1 | *1* | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 9 |
| A10.Environmental Sustainability | 1 | *1* | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | *1* | 1 | 1 | 0 | 1 | 9 |
| A11.Customer Recommandation | *1* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 6 |
| A12.Adaptability or Openness to Change | 0 | 0 | *1* | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 4 |
| A13.Adoption Intention of Water Ionizer | 0 | 0 | *1* | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | *1* | 1 | 0 | 0 | 4 |
| A14.Perceived usefulness | 1 | *1* | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 13 |
| A15.Compliance Standard | 1 | *1* | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 8 |
| Dependence Power | 13 | 11 | 15 | 1 | 6 | 6 | 6 | 15 | 1 | 7 | 13 | 15 | 15 | 7 | 11 | |

**Fig. 5** Final reachability matrix (FRM). *Source*: authors

dependent on water ionizers are brand value, societal acceptance, flexibility, readiness to change, and adoption goals.

Health benefits for customers and their Perceived value along with compliance standard and brand values in the third quadrant signifies the critical linkage and these elements demonstrate both high dependence and driving power, acting as enablers of adoption. Linking factors are those that, by balancing the advantages and disadvantages of the product, consistently influence consumer perceptions. It depicts the consumer perception of stainability and after sale to put high driving power, which signifies the consumer's intention to purchase a sustainable product for maximum benefits.

## 5 Discussion

Innovations in the water technology industry through ERW particularly benefit commercial and household applications. AIW offers health benefits, and acidic water signifies a trade-off for various industrial applications, especially in the healthcare industry. The analysis in this study reveals the impact of adoption through identifying factors that influence the adoption of water ionizers by applying the ISM methodology. Investigating the relationships among different elements that impact prospective users' decisions to adopt water ionizers, is the spirit of this research. In Fig. 7,

| Variable_Names | Reachability_Set | Antecedents_Set | Intersection_Set | Level |
|---|---|---|---|---|
| A1 | A1 A2 A3 A8 A11 A12 A13 A14 A15 | A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 A11 A14 A15 | A1 A2 A3 A8 A11 A14 A15 | 0 |
| A2 | A1 A2 A3 A5 A6 A7 A8 A10 A11 A12 A13 A14 A15 | A1 A2 A3 A4 A5 A6 A7 A9 A10 A14 A15 | A1 A2 A3 A5 A6 A7 A10 A14 A15 | 0 |
| A3 | A1 A2 A3 A8 A11 A12 A13 A15 | A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 A11 A12 A13 A14 A15 | A1 A2 A3 A8 A11 A12 A13 A15 | 1 |
| A4 | A1 A2 A3 A4 A5 A6 A7 A8 A10 A11 A12 A13 A14 A15 | A4 | A4 | 0 |
| A5 | A1 A2 A3 A5 A6 A7 A8 A10 A11 A12 A13 A14 A15 | A2 A4 A5 A6 A7 A14 | A2 A5 A6 A7 A14 | 0 |
| A6 | A1 A2 A3 A5 A6 A7 A8 A10 A11 A12 A13 A14 A15 | A2 A4 A5 A6 A7 A14 | A2 A5 A6 A7 A14 | 0 |
| A7 | A1 A2 A3 A5 A6 A7 A8 A10 A11 A12 A13 A14 A15 | A2 A4 A5 A6 A7 A14 | A2 A5 A6 A7 A14 | 0 |
| A8 | A1 A3 A8 A11 A12 A13 | A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 A11 A12 A13 A14 A15 | A1 A3 A8 A11 A12 A13 | 1 |
| A9 | A1 A2 A3 A8 A9 A11 A12 A13 A15 | A9 | A9 | 0 |
| A10 | A1 A2 A3 A8 A10 A11 A12 A13 A15 | A2 A4 A5 A6 A7 A10 A14 | A2 A10 | 0 |
| A11 | A1 A3 A8 A11 A12 A13 | A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 A11 A14 A15 | A1 A3 A8 A11 | 0 |
| A12 | A3 A8 A12 A13 | A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 A11 A12 A13 A14 A15 | A3 A8 A12 A13 | 1 |
| A13 | A3 A8 A12 A13 | A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 A11 A12 A13 A14 A15 | A3 A8 A12 A13 | 1 |
| A14 | A1 A2 A3 A5 A6 A7 A8 A10 A11 A12 A13 A14 A15 | A1 A2 A4 A5 A6 A7 A14 | A1 A2 A5 A6 A7 A14 | 0 |
| A15 | A1 A2 A3 A8 A11 A12 A13 A15 | A1 A2 A3 A4 A5 A6 A7 A9 A10 A14 A15 | A1 A2 A3 A15 | 0 |
| | | | | |
| Variable_Names | Reachability_Set | Antecedents_Set | Intersection_Set | Level |
| A1 | A1 A2 A11 A14 A15 | A1 A2 A4 A5 A6 A7 A9 A10 A14 A15 | A1 A2 A14 A15 | 0 |
| A2 | A1 A2 A5 A6 A7 A10 A11 A14 A15 | A1 A2 A4 A5 A6 A7 A9 A10 A14 A15 | A1 A2 A5 A6 A7 A10 A14 A15 | 0 |
| A4 | A1 A2 A4 A5 A6 A7 A10 A11 A14 A15 | A4 | A4 | 0 |
| A5 | A1 A2 A5 A6 A7 A10 A11 A14 A15 | A2 A4 A5 A6 A7 A14 | A2 A5 A6 A7 A14 | 0 |
| A6 | A1 A2 A5 A6 A7 A10 A11 A14 A15 | A2 A4 A5 A6 A7 A14 | A2 A5 A6 A7 A14 | 0 |
| A7 | A1 A2 A5 A6 A7 A10 A11 A14 A15 | A2 A4 A5 A6 A7 A14 | A2 A5 A6 A7 A14 | 0 |
| A9 | A1 A2 A9 A11 A15 | A9 | A9 | 0 |
| A10 | A1 A2 A10 A11 A15 | A2 A4 A5 A6 A7 A10 A14 | A2 A10 | 0 |
| A11 | A11 | A1 A2 A4 A5 A6 A7 A9 A10 A11 A14 A15 | A11 | 1 |
| A14 | A1 A2 A5 A6 A7 A10 A11 A14 A15 | A1 A2 A4 A5 A6 A7 A14 | A1 A2 A5 A6 A7 A14 | 0 |
| A15 | A1 A2 A11 A15 | A1 A2 A4 A5 A6 A7 A9 A10 A14 A15 | A1 A2 A15 | 0 |
| | | | | |
| Variable_Names | Reachability_Set | Antecedents_Set | Intersection_Set | Level |
| A1 | A1 A2 A14 A15 | A1 A2 A4 A5 A6 A7 A9 A10 A14 A15 | A1 A2 A14 A15 | 1 |
| A2 | A1 A2 A5 A6 A7 A10 A14 A15 | A1 A2 A4 A5 A6 A7 A9 A10 A14 A15 | A1 A2 A5 A6 A7 A10 A14 A15 | 1 |
| A4 | A1 A2 A4 A5 A6 A7 A10 A14 A15 | A4 | A4 | 0 |
| A5 | A1 A2 A5 A6 A7 A10 A14 A15 | A2 A4 A5 A6 A7 A14 | A2 A5 A6 A7 A14 | 0 |
| A6 | A1 A2 A5 A6 A7 A10 A14 A15 | A2 A4 A5 A6 A7 A14 | A2 A5 A6 A7 A14 | 0 |
| A7 | A1 A2 A5 A6 A7 A10 A14 A15 | A2 A4 A5 A6 A7 A14 | A2 A5 A6 A7 A14 | 0 |
| A9 | A1 A2 A9 A15 | A9 | A9 | 0 |
| A10 | A1 A2 A10 A15 | A2 A4 A5 A6 A7 A10 A14 | A2 A10 | 0 |
| A14 | A1 A2 A5 A6 A7 A10 A14 A15 | A1 A2 A4 A5 A6 A7 A14 | A1 A2 A5 A6 A7 A14 | 0 |
| A15 | A1 A2 A15 | A1 A2 A4 A5 A6 A7 A9 A10 A14 A15 | A1 A2 A15 | 1 |
| | | | | |
| Variable_Names | Reachability_Set | Antecedents_Set | Intersection_Set | Level |
| A4 | A4 A5 A6 A7 A10 A14 | A4 | A4 | 0 |
| A5 | A5 A6 A7 A10 A14 | A4 A5 A6 A7 A14 | A5 A6 A7 A14 | 0 |
| A6 | A5 A6 A7 A10 A14 | A4 A5 A6 A7 A14 | A5 A6 A7 A14 | 0 |
| A7 | A5 A6 A7 A10 A14 | A4 A5 A6 A7 A14 | A5 A6 A7 A14 | 0 |
| A9 | A9 | A9 | A9 | 1 |
| A10 | A10 | A4 A5 A6 A7 A10 A14 | A10 | 1 |
| A14 | A5 A6 A7 A10 A14 | A4 A5 A6 A7 A14 | A5 A6 A7 A14 | 0 |
| | | | | |
| Variable_Names | Reachability_Set | Antecedents_Set | Intersection_Set | Level |
| A4 | A4 A5 A6 A7 A14 | A4 | A4 | 0 |
| A5 | A5 A6 A7 A14 | A4 A5 A6 A7 A14 | A5 A6 A7 A14 | 1 |
| A6 | A5 A6 A7 A14 | A4 A5 A6 A7 A14 | A5 A6 A7 A14 | 1 |
| A7 | A5 A6 A7 A14 | A4 A5 A6 A7 A14 | A5 A6 A7 A14 | 1 |
| A14 | A5 A6 A7 A14 | A4 A5 A6 A7 A14 | A5 A6 A7 A14 | 1 |
| | | | | |
| Variable_Names | Reachability_Set | Antecedents_Set | Intersection_Set | Level |
| A4 | A4 | A4 | A4 | 1 |
| | | | | |

**Fig. 6** Iteration-wise level partitioning. *Source*: authors

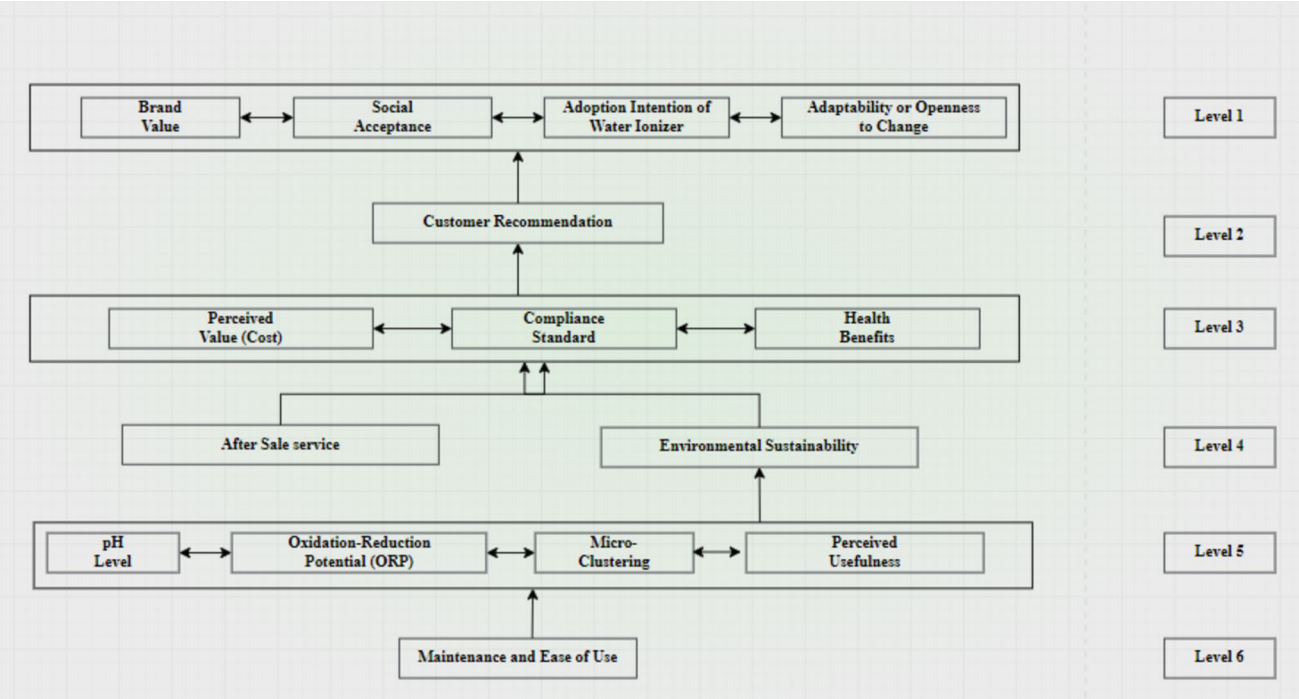



**Fig. 7** Interpretive structure model (ISM). *Source*: authors

maintenance and usability are positioned at the sixth level, which theoretically serves as the base of the hierarchical model. pH level, ORP, Micro-clustering, and usefulness are essential features of water ionizers, which are also interrelated at the same level. At the top, brand value, social acceptance, adoption intention of water ionizers are interrelated at the same level, which is essential to derive the adoption of water ionizers.

Furthermore, the framework containing the MICMAC quadrant confines water ionizer factors based on their driving and dependence power. Figure 8 displays the graph which is the MICMAC outcome, and divides factors into four quadrants according to their driving and its reliance power.

Low driving and dependency powers characterize the variables in the first quadrant. However, the study's analysis reveals that any variable is not located in this quadrant, revealing that any of the determined water ionizer factors does not exhibit autonomous behaviour of dependence and driving power.

Health benefits and perceived value are the key enablers influencing consumer decisions about adopting water ionizers. Based on the above ISM-MICMAC implementation, as summarized in Fig. 9, the following inferences could be drawn:

- Level partitioning and digraphs depict the pH level, ORP, micro-clustering, and potential usefulness as independent after product Maintenance and Use. These are highly interrelated to show the significant critical factors for usefulness and benefits.
- Brand value, Social Acceptance, Adaptability or Openness to Change, and the Intention to adopt a Water Ionizer are key factors that are highly dependent and interrelated and can motivate consumers toward adoption.
- The Oxidation–Reduction Potential (ORP), Micro-Clustering, and most essentially pH Factor are critical factors that drive health benefits and impact consumers' perceptions of adoption.
- It is visible that the compliance standard is a linkage variable along with health benefits, brand value, perceived value, and cost. Compliance standards like ISO 13485, 14001, and 9001 ensure standard practices to ensure environmental sustainability and product features for consumer adoption.
- The MICMAC graph shows that factors falling in the third and fourth quadrants are the critical enablers for adoption.

The subjectivity of hierarchical structure in ISM is interpreted as the outcome of the expert judgement, however, the robustness and validity of the results is justified through several measures carried out for data munging. The qualified experts of selected panel from various relevant fields ensures a balanced perspective, reducing individual biases. Thereafter, the implementation of well-documented ISM process while

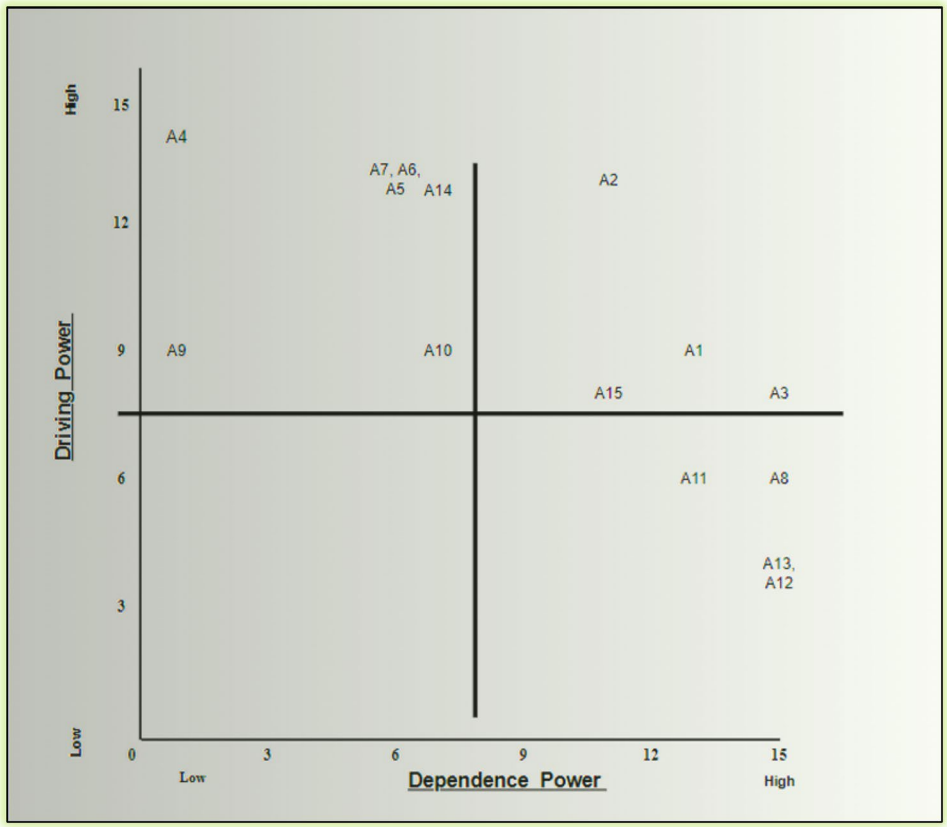



**Fig. 8** MICMAC analysis.
*Source*: authors

| MICMAC Quadrant | Enabler Types | Determinants |
|---|---|---|
| I | Autonomous Enabler | |
| II | Dependent Enabler | Social Acceptance, Customer Recommendation, Adaptability or Openness to Change, Adoption Intention of Water Ionizer |
| III | Linkage Enabler | Perceived Value (Cost), Health Benefits, Brand Value, Compliance Standard |
| IV | Independent Enabler | Maintenance and Ease of Uses, pH Level, Oxidation-Reduction Potential (ORP), Micro-Clustering, After Sale service, Environmental Sustainability, Perceived usefulness |

**Fig. 9** Enablers in MICMAC quadrant

forming the SSIM, IRM, and FRM from the input expert data provides a clear and consistent methodology. Additionally, the use of MICMAC analysis complements the qualitative judgments with quantitative validation, categorizing variables based on their driving and dependence power. Transparency in documentation and clear justification of each relationship based on expert inputs and literature review also contribute to the reliability of the results to ensure the hierarchical structure accurately represents the complex interdependencies among the enablers for the sustainable adoption of water ionizers.

## 6 Managerial implications

The use the ISM-MICMAC approach indicates the enablers of water ionizer assist in adoption to improvise human health

and sustainability in society. An organization which practices for their manufacturing, marketing, and sustainable adoption of product or services, the managerial implications of water ionizers are significant with a particular emphasis on their acceptance and sustainable practices. A marketing strategy targeting potential consumers can emphasize the factors revealed as enablers in above ISM-MICMAC study, including the product sustainability that serves their preferences and beliefs. Manufacturers prioritize developing water ionizers with sustainable attributes that reduce environmental effects throughout the product's lifecycle, such as energy efficiency and recyclable materials. Original equipment manufacturer (OEM) based organization pursue the various ISO standards for better health and sustainability. As per the above analysis, environmental sustainability and product features that follow the ISO standards, influencing consumers' health and recommendations, are vital in adopting water ionizers.

ISM-MICMAC analysis classified critical factors affecting water ionizers' effectiveness and the consumer perception on its adoption. Managers should employ this study to prioritize their efforts and resources on factors with high driving power and the most influential in achieving desired outcomes in product marketing for consumer adoption and can focus more on consumer awareness to penetrate the market more effectively. Also, improving environmental management, sustainable practices, enriching product features for better health results, and effective business planning would result in business expansion, growth, and continuity.

The consumer awareness on ERW or AIW for health benefits, its role for sustainable society through various platforms e.g. educational programs, internet resources for adoption could be encouraged to strategies and additionally, organization collaborating with related industry, healthcare providers, and environmental groups to push laws supporting the use of water ionizers and encouraging sustainable hydration habits in the community.

Organizations can enhance the sustainability of water ionizers, foster customer acceptance, and support the adoption of eco-friendly and improved hydration practices by following the insightful details from the study and incorporating these holistic managerial implications to form their strategies and operational decision.

## 7 Limitations of the study, future work, and conclusion

The awareness on health benefits of ERW or AIW is key to driving up demand for water ionizer among consumers. The technological breakthroughs for society well-being, eco-friendly activities, consumer awareness, and environmental sustainability are pivotal strategic attention of manufacturers to increase significant demand in the market. The strategy and campaigning to inform customers about the devices' optimistic effects on the environment as well as on their health would improve the finances through various channel and to fully realize the industry's potential, challenges as limited consumer awareness and high initial costs attention must be addressed.

This study furnishes valuable perspicuity into the interrelationships among behavioral constructs while acknowledging several limitations. This study illustrates an initial effort to explore the relationship between the prototypes and the influences of ionizer's adoption by consumers. The lack of statistical validation of the proposed model in this study, and consequently, future research should scope the marketing trend and use of digital platform for the market penetration and revenue growth. Additionally, the strategic road map for water technology industry progression will serve as a pioneering approach to predicting future market and revenue growth.

**Declarations**

**Conflict of interest** The authors have no potential conflict of interest to disclose.

**Research involving human participants and/or animals** This research does not contain any human participant and/or animals.

**Informed consent** All the authors of the article are fully aware about this submission.

## References

AlSaleh D, Thakur R (2019) Impact of cognition, affect, and social factors on technology adoption. Int J Technol Market 13(2):178–200

Agarwal M, Aggrawal D, Anand A, Singh O (2017) Modeling multigeneration innovation adoption based on conjoint effect of awareness process. Int J Math Eng Manage Sci 2(2):74

Anand A, Bansal G (2017) Interpretive structural modeling for attributes of software quality. J Adv Manag Res 14(3):256–269

Anand A, Aggrawal D, Agarwal M (2019) Market assessment with OR applications. CRC Press, Boca Raton

Attri R, Grover S (2018) Analysis of quality enabled factors in the product design stage of a production system life cycle: a relationship modelling approach. Int J Manage Sci EngManage 13(1):65–73

Botto S (2009) Tap water vs bottled water in a footprint integrated approach. Nat Prec. https://doi.org/10.1038/npre.2009.3407.1

Coria J, Villegas-palacio CLARA (2014) Regulatory dealing: technology adoption versus enforcement stringency of emission taxes. Contemp Econ Policy 32(2):451–473

Crosno JL, Cui AP (2014) A multilevel analysis of the adoption of sustainable technology. J Market Theory Pract 22(2):209–224

Fu Y, Kok RA, Dankbaar B, Ligthart PE, van Riel AC (2018) Factors affecting sustainable process technology adoption: a systematic literature review. J Clean Product 205:226–251

The electrolysis process|Enagic Kangen Water (n.d.) Enagic-Asia.com. https://enagic-asia.com/en/electrolysis-process

Gleick PH, Cooley HS (2009) Energy implications of bottled water. Environ Res Lett 4(1):014009
Gupta A, Sahu GP (2013) Factors influencing adoption of relationship marketing practices for overall growth of firm: an ISM-based model validation. Int J Electron Custom Relationsh Manag 7(1):21–44
Hamasaki T, Kashiwagi T, Aramaki S, Imada T, Komatsu T, Li Y, Teruya K, Katakur Y, Shirahata S (2005) Suppression of cell growth by platinum nanocolloids as scavengers against reactive oxygen species. In: Animal cell technology meets genomics: proceedings of the 18th ESACT meeting Granada, Spain, May 11–14, 2003. Springer. Netherlands, pp 249–251
Hayashi HW (1995) The chemisty of life. Explore 6:28–31
Henry M, Chambron J (2013) Physico-chemical, biological and therapeutic characteristics of electrolyzed reduced alkaline water (ERAW). Water 5(4):2094–2115
Hricova D, Stephan R, Zweifel C (2008) Electrolyzed water and its application in the food industry. J Food Prot 71(9):1934–1947
Ignacio RMC, Joo KB, Lee KJ (2012) Clinical effect and mechanism of alkaline reduced water. J Food Drug Anal 20(1):33
Ignatov I (2017) VITA intense-antioxidant and inhibition growth of tumor cells effects. Anti-aging influence. Negative oxidation-reduction potential (ORP) has important role in these effects. J Med Physiol Biophys 39:20–42
Ignatov I, Mosin O, Kirov P (2016) Mathematical model of Kangen water® Biophysical and biochemical effects of catholyte. Adv Phys Theor Appl 51(1):33–55
ISO 14001:2015 (2022) ISO 13 Dec 2022. https://www.iso.org/standard/60857.html
Jin D, Ryu SH, Kim HW, Yang EJ, Lim SJ, Ryang YS, Chung CH, Park SK, Lee KJ (2006) Anti-diabetic effect of alkaline-reduced water on OLETF rats. Biosci Biotechnol Biochem 70(1):31–37
Joshi Y, Rahman Z (2017) Investigating the determinants of consumers' sustainable purchase behavior. Sustain Product Consum 10:110–120
Jurison J (2000) Perceived value and technology adoption across four end user groups. J Org End User Comput 12(4):21–28
Kristensson P, Pedersen PE, Thorbjørnsen H (2020) New perspectives on consumer adoption and diffusion of innovations. J Bus Res 116:522–525
Kumar A, Bhaskar P, Nadeem SP, Tyagi M, Garza-Reyes JA (2020) Sustainability adoption through sustainable human resource management: a systematic literature review and conceptual framework. Int J Math Eng Manag Sci 5(6):1014
Kun-Kul R (2005) Permeability and dissolvability of cathodic electrolyzed water for electrophoretic gel and green tea components. J Korea Acad Ind Cooper Soc 6(1):87–93
LeBaron TW, Sharpe R, Ohno K (2022) Electrolyzed-reduced water: review I. molecular hydrogen is the exclusive agent responsible for the therapeutic effects. Int J Mol Sci 23(23):14750
Li CR, Lin CJ (2015) New product adoption and sales performance from the importer perspective. Indust Market Manage 44:98–106
Li W, Long R, Chen H, Geng J (2017) A review of factors influencing consumer intentions to adopt battery electric vehicles. Renew Sustain Energy Rev 78:318–328
Lin X, Cui S, Han Y, Geng Z, Zhong Y (2019) An improved ISM method based on GRA for hierarchical analyzing the influencing factors of food safety. Food Control 99:48–56
Lu H, Ayers E, Patel P, Mattoo TK (2023) Body water percentage from childhood to old age. Kidney Res Clin Pract 42(3):340
M (2023) The sustainable development agenda—United Nations Sustainable Development. United Nations Sustainable Development. https://www.un.org/sustainabledevelopment/development-agenda/
Majer JM, Henscher HA, Reuber P, Fischer-Kreer D, Fischer D (2022) The effects of visual sustainability labels on consumer perception and behavior: a systematic review of the empirical literature. Sustain Product Consum 33:1–14
Mathiyazhagan K, Govindan K, NoorulHaq A, Geng Y (2013) An ISM approach for the barrier analysis in implementing green supply chain management. J Clean Prod 47:283–297
Mohr R (2016) The healing power of Kangen water. Blurb, San Francisco
Murry JW Jr, Hammons JO (1995) Delphi: A versatile methodology for conducting qualitative research. Rev High Educ 18(4):423–436
Nandal V, Kumar R, Singh SK (2019) Barriers identification and analysis of solar power implementation in Indian thermal power plants: an Interpretative structural modeling approach. Renew Sustain Energy Rev 114:109330
Nguyen N, Johnson LW (2020) Consumer behavior and environmental sustainability. J Consum Behav 19(6):539–541
Okajima Y, Okajima M, Ikeda M, Wada Y, Shimokawa KI, Ishii F (2021) Antiviral effect of electrolyzed reduced water on SARS-CoV-2. Drug Discov Ther 15(5):268–272
Pandey S, Aggrawal D (2024) Sustainable practices and acceptance of water ionizers: a structural equation modeling approach. Manag Environ Qual: Int J 35(8):1912–1928
Patra SK, Raj T, Arora BB (2021) An analysis and modelling of selected barriers for a sustainable manufacturing system using ISM technique. Int J Oper Res 41(3):373–398
Peng RD (2015) R programming for data science. Leanpub, Victoria, pp 86–181
Pisoschi AM, Pop A, Iordache F, Stanca L, Predoi G, Serban AI (2021) Oxidative stress mitigation by antioxidants—an overview on their chemistry and influences on health status. Eur J Med Chem 209:112891
Pocock SJ, Shaper AG, Packham RF (1981) Studies of water quality and cardiovascular disease in the United Kingdom. Sci Total Environ 18:25–34
Ronen T, Kerret D (2020) Promoting sustainable well-being: Integrating positive psychology and environmental sustainability in education. Int J Environ Res Public Health 17(19):6968
Rubik B (2011) Studies and observations on the health effects of drinking electrolyzed-reduced alkaline water. Water Soc 153:321–327
Schreier M, Oberhauser S, Prügl R (2007) Lead users and the adoption and diffusion of new products: insights from two extreme sports communities. Mark Lett 18:15–30
Seal AD, Suh HG, Jansen LT, Summers LG, Kavouras SA (2019) Hydration and health. Anal Nutrit Res. https://doi.org/10.1016/b978-0-12-814556-2.00011-7
Shirahata S, Kabayama S, Nakano M, Miura T, Kusumoto K, Gotoh M, Hayashi H, Otsubo K, Morisawa S, Katakura Y (1997) Electrolyzed–reduced water scavenges active oxygen species and protects DNA from oxidative damage. Biochem Biophys Res Commun 234(1):269–274
Shirahata S, Hamasaki T, Teruya K (2012) Advanced research on the health benefit of reduced water. Trends Food Sci Technol 23(2):124–131
Shirahata S, Murakami E, Kusumoto KI, Yamashita M, Oda M, Temya K, Kabayama S, Otsubo K, Morisawa S, Hayashi H, Katakura Y (2002) Telomere shortening in cancer cells by electrolyzed-reduced water. In: Animal cell technology: challenges for the 21st century: proceedings of the joint international meeting of the Japanese Association for Animal Cell Technology (JAACT) and the European Society for Animal Cell Technology (ESACT) 1998, Kyoto, Japan. Springer, Dordrecht, pp 355–359
Silva JA (2023) Wastewater treatment and reuse for sustainable water resources management: a systematic literature review. Sustainability 15(14):10940
Sindhwani R, Saddikuti V, Vaidya OS (2023) Evaluation of enablers of supply chain resilience and responsibility in India during

large-scale disruptions: an ISM-ANP approach. Int J Oper Res 48(2):178–220
Singh AK, Samuel C (2018) modelling the strengthening factors for competitive position of apparel retailing in India. J Model Manag 13(4):884–907
Sugandini D, Purwoko P, Pambudi A, Resmi S, Reniati R, Muafi M, Adhyka Kusumawati R (2018) The role ofuncertainty, perceived ease of use, and perceived usefulness towards the technology adoption. Int J Civil Eng Technol 9(4):660–669
Sushil (2012) Interpreting the interpretive structural model. Glob J Flex Syst Manag 13:87–106
Tanner C, Wölfing Kast S (2003) Promoting sustainable consumption: determinants of green purchases by Swiss consumers. Psychol Mark 20(10):883–902
Trudel R (2019) Sustainable consumer behavior. Consum Psychol Rev 2(1):85–96
Tsai CF, Hsu YW, Chen WK, Ho YC, Lu FJ (2009) Enhanced induction of mitochondrial damage and apoptosis in human Leukemia HL-60 cells due to electrolyzed-reduced water and glutathione. Biosci Biotechnol Biochem 73(2):280–287
Ullah I, Narain R (2021) Analyzing the barriers to implementation of mass customization in Indian SMEs using integrated ISM-MICMAC and SEM. J Adv Manag Res 18(2):323–349
Usmani MS, Wang J, Waqas M, Iqbal M (2023) Identification and ranking of enablers to green technology adoption for manufacturing firms using an ISM-MICMAC approach. Environ Sci Pollut Res 30(17):51327–51343
Vacchi M, Siligardi C, Demaria F, Cedillo-González EI, González-Sánchez R, Settembre-Blundo D (2021) Technological sustainability or sustainable technology? A multidimensional vision of sustainability in manufacturing. Sustainability 13(17):9942
Warfield JN (1974) Developing interconnection matrices in structural modeling. IEEE Trans Syst Man Cybern 1:81–87
What percentage of the human body is water? (2020) May 27, 2010. www.medicalnewstoday.com. https://www.medicalnewstoday.com/articles/what-percentage-of-the-human-body-is-water#where-in-the-body
Ye J, Li Y, Hamasaki T, Nakamichi N, Komatsu T, Kashiwagi T, Teruya K, Nishikawa R, Kawahara T, Osada K, Toh K, Abe M, Tian H, Kabayama S, Otsubo K, Morisawa S, Katakura Y, Shirahata S (2008) Inhibitory effect of electrolyzed reduced water on tumor angiogenesis. Biol Pharmaceut Bull 31(1):19–26
Yoshida K, Achiwa N, Katayose M (2004) Application of electrolyzed water for food industry in Japan

**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

*International Journal of Communication Systems*

WILEY

RESEARCH ARTICLE

# Network Partitioning Problem and UAVs' Integration for Efficient Connectivity Restoration: A Systematic Review

Aditi Zear[1] | Virender Ranga[2] | Kriti Bhushan[3]

[1]School of Computer Science, UPES, Dehradun, Uttarakhand, India | [2]Department of Information Technology, Delhi Technological University, Delhi, India | [3]Department of Computer Engineering, NIT Kurukshetra, Kurukshetra, Haryana, India

**Correspondence:** Aditi Zear (aditizear93@gmail.com)

**Received:** 20 August 2023 | **Revised:** 28 August 2024 | **Accepted:** 10 December 2024

**Funding:** We are grateful to the Science and Engineering Research Board (SERB)—Department of Science and Technology (DST), Government of India, for sponsoring this research.

**Keywords:** connectivity restoration | FANET | network partitioning | partition detection | UAVs | UAV-enabled recovery

## ABSTRACT

Wireless sensor and actor networks (WSANs) are gaining substantial recognition because of their utility in inhospitable environments where humans have restricted accessibility. The extensive applications of WSANs in different domains require effectiveness, reliability, and some degree of robustness. These networks may experience frequent node failures due to their deployment in rough environments, such as energy depletion and onboard electronics failure of network nodes. These failures result in areas with no coverage, which may further deteriorate the standard of data accumulated. However, the network partitioning into disjoint fragments is the most severe repercussion that arises due to these failures. The network partitioning results in various negative impacts like obstruction in data exchange and restricted coordination among the nodes. Therefore, detecting partitioning in the network and restoring connectivity are important. This paper reviews the reported network partition detection and recovery techniques. The limitations of these recovery techniques are highlighted, along with the advantages of incorporating unmanned aerial vehicles (UAVs) in various ground wireless networks. UAVs are evolving to become a critical element of future wireless network technologies. This paper incorporates the analysis of UAV-assisted networks consisting of technical issues, challenges, and requirements related to the UAVs' employment in wireless networks. Thereafter, the reported UAV-assisted partition recovery techniques are discussed along with the application scenarios for incorporating UAVs in network partitioning detection and recovery problem. The paper also highlights the challenges for associating UAVs in the network recovery process.

## 1 | Introduction

WSANs are more advanced variants of WSNs, consisting of an overlay network of sensor and actuator or actor nodes. The actor nodes are more computationally powerful along with the capability to relocate (e.g., vehicles, robots, rovers). These actors gather data from the sensors and utilize it to take action accordingly. These systems are used increasingly because of their ability to perform functions without any real-time and direct human intervention. The advancement of mobile heterogeneous nodes has been spurred recently because of the developments in node hardware and the usability of WSANs in various applications. Such nodes include mobile sensors, actuators, vehicles, robots, and UAVs. A variety of networks emerged by combining these nodes like WSANs, wireless sensor and robot networks (WSRNs), and mobile sensor networks (MSNs). The focus of researchers for such networks is to deal with the issues associated with mobility and optimal node placement in order to ensure adequate network connectivity, coverage, and real-time data accumulation [1, 2]. The mobility of nodes in wireless networks

© 2024 John Wiley & Sons Ltd.

is considered challenging and advantageous. The challenges are difficulty in topology control and handling the mobility of network nodes. In spite of various difficulties, the advantages of mobility can be exploited to solve the issues related to data collection, fault tolerance, coverage, and real-time actuation. The tiny sensor nodes (SNs) are exposed to failures because of definite battery life, hardware malfunctions, and deployment in rough conditions such as landslides, glacier or volcano monitoring, battlefield surveillance, and fire detection. The failure of nodes deployed for monitoring the assigned area obstructs their communication and results in information loss [3–5]. Figure 1 shows WSAN partitioned due to node failures [6, 7].

### 1.1 | Motivation

Connectivity with the base station (BS) is eminent in WSANs; hence, the redundant SNs are also set up in the network to endure some node failures. However, even with these preventive measures, the concurrent failure of multiple nodes can cause network partitioning, which may extensively obstruct the entire network operation. The network's self-healing after partitioning is unattainable because the nodes are unaware of the scope of failure and information about the nodes in distinct partitions [8]. One mechanism to recover the network partitioning is to relocate existing active network nodes to re-establish network connectivity. The majority of node repositioning strategies reorganize the network structure intending to lower energy consumption due to movement. The solution is quick since any external involvement, such as extra node employment, is not required. In some cases, the nodes in the damaged network might not be mobile or sufficient for network recovery. Deploying new relay nodes (RNs) is also an alternative method to reconnect the partitioned segments. These nodes either move toward the assigned location or are deployed manually, and the information about partitions, the location of nodes present in these partitions, number of failed nodes, and their details are presumed [9–11].

Whether these techniques are distributed or centralized, they assume that the terrain in which the nodes move is uniform and obstacles are not present in the environment. Hence, all the movements are considered successful, which may not be possible. Moreover, the long-distance movement of the SNs is not favored because it straightaway increases energy consumption and requires complex hardware. However, the main requirement before starting the recovery procedure is complete information about the damaged region, which is crucial and challenging to obtain using conventional ground nodes (GNs) [9–11].

One mechanism to deal with this problem is assessing the shape and size of damaged area, and determining the actual failed nodes intending to compute new deployment locations for the RNs. Aerial devices are the ideal candidate that can aid the connectivity restoration process by gathering data about the affected region. UAVs have emerged from remote-controlled devices to small autonomous vehicles that can carry out intricate missions collectively. The continuous growth is experienced in the UAV market from $19.3 billion in 2019 to $45.8 billion market by 2025 due to the utilization of advanced technologies such as light detection and ranging (LiDAR) and hyperspectral sensors, object recognition using AI, multispectral targeting mechanisms, infrared sensors, environmental sensors, tensor/graphical processing unit (TPU/GPU). These sensors and chips introduce new level of intelligence and improve the performance of UAVs in different applications [12–15]. The promising benefits of UAVs raise the queries given below:

- Can UAVs be employed to deal with the drawbacks of the existing network recovery techniques?
- How can UAVs be used to perform the partition detection process?
- How can UAVs aid the connectivity restoration process after network partitioning?
- What are the implementation challenges and requirements for the UAVs' employment in detection and recovery processes?

In the case of network partitioning, quick sensing systems are required to get the complete state of the affected region. Therefore, according to the applications' requirements, these UAVs can collect data using various sensors and provide their gathered information to the central BS as shown in Figure 2. Afterward, different connectivity restoration algorithms can be implemented based on the received information to form robust postfailure network topology. Our work in [9, 10, 16] focuses on employing UAVs for detecting the network partitioning and aiding the recovery process. Interested readers can refer to these papers to understand UAVs' assisted network

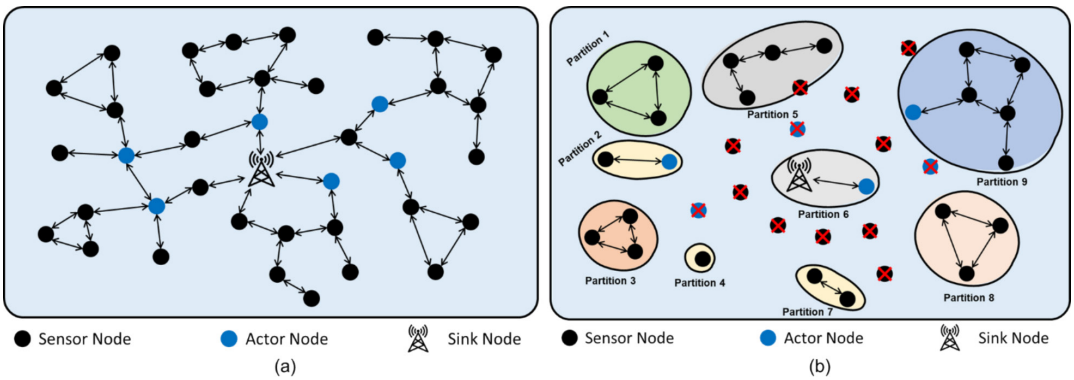



**FIGURE 1** | (a) Wireless sensor network, (b) partitioned network.



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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

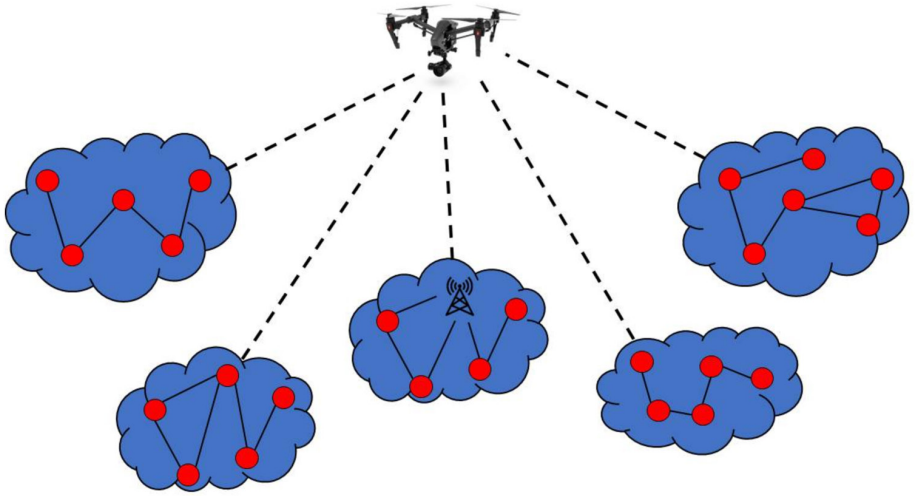


**FIGURE 2** | UAV-assisted network recovery.

recovery. Associating UAVs to build up connectivity in the partitioned network requires effective control, sensing, and communication algorithms to accommodate their dynamic mobility levels. The applications of the UAVs can be integrated with partition recovery algorithms in order to furnish excellent results. This paper summarizes the UAVs' applications in communication networks and proposes different methods to incorporate UAVs to deal with network partitioning problems. Table 1 shows the database of publications that are reviewed in this paper. As per our best knowledge, there is no extensive survey based on UAVs' assisted connectivity restoration from network partitioning.

## 1.2 | Contribution

This article is motivated by the earlier discussed limitations of network partition recovery techniques. This paper covers the important features and requirements for employing UAVs in the partition recovery process. The rising trend of employing UAVs as providers and users of communication facilities and as a relay to establish communication in different communication networks has been observed. The summary of various contributions of this survey is listed below:

- The popular methods employed for detecting network partitioning and recovery are discussed, along with their drawbacks.
- The limitations of the ground-based recovery solutions are discussed, along with the methods to incorporate unmanned aerial sensing (UAS) mechanisms to mitigate them.
- The article highlights the use cases of UAVs in different communication networks along with the requirements.
- The different methods to incorporate UAVs in partition detection and recovery are discussed.

Various challenges and requirements for employing UAVs to deal with the network partitioning problem are discussed.

**TABLE 1** | Publication database.

| Sr. no. | Literature source | Number of studies |
|---|---|---|
| 1 | Journal of Network and Computer Applications | 16 |
| 2 | Computer Networks | 12 |
| 3 | Ad Hoc Networks | 13 |
| 4 | IEEE Access | 9 |
| 5 | IEEE Internet of Things | 5 |
| 6 | Computer Communication | 8 |
| 7 | International Journal of Communication Systems | 2 |
| 8 | Wireless Networks | 3 |
| 9 | Transactions on Emerging Telecommunications Technologies | 2 |
| 10 | IEEE Communications Surveys and Tutorials | 2 |
| 11 | Wireless Personal Communication | 3 |
| 12 | IEEE Transaction on Vehicular Technology | 5 |
| 13 | IEEE Transactions on Computers | 3 |
| 14 | IEEE Transactions on Communication | 2 |
| 15 | IEEE Transactions on Wireless Communication | 2 |
| 16 | Journal of Parallel and Distributed Computing | 2 |
| 17 | Computers and Electrical Engineering | 2 |
| 18 | IEEE Wireless Communication Letters | 2 |
| 19 | IEEE Communications Letters | 3 |



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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

### 1.3 | Network Partitioning Problem in WSANs

WSANs have numerous advantages, including their ability to work in humans' inaccessible locations. The actor and SNs placed in such regions are vulnerable to adverse situations such as flooding, storm, fire, and wind. The failure of multiple network nodes can partition the network. Such partitioning prevents the association between the network nodes working together to share the data and perform collaborative operations. Hence, fault tolerance from these damages is a widely investigated problem. Detecting network partitioning is very challenging for the active network nodes since they can only connect with the nodes in their own partition. Hence, these nodes cannot even determine the actual failed nodes in the partitioned network. A successful network recovery operation requires detecting the failed nodes' count, total partitions in the damaged network, and nodes in these partitions [3–5]. Therefore, intending to handle such situations, more complex self-detection and recovery systems are required. The following sections and their subsections discuss the review of partition detection and recovery schemes, the utilization of UAVs in different ground networks, and the application scenarios of UAVs in mitigating the network partitioning problem [9, 11, 16]. Figure 3 illustrates the taxonomy of the network partitioning problem and UAVs' assistance for providing connectivity in ground networks. Overall, the paper is organized according to this diagram. Section 1.3 explains the network partitioning problem along with the review of detection and recovery techniques in Sections 2.1 and 2.2. Thereafter, the applications of UAVs and their utilization to deal with the network partitioning have been discussed in Section 3. Afterwards, the research challenges for employing UAVs to solve the network recovery problems are discussed in Section 5.

## 2 | Related Work

Various network partition recovery approaches are mainly classified according to the type of failure (single node or multiple nodes), coverage, connectivity, mobility of network nodes, and simultaneity of failures. Ranga et al. [17] presented a survey summarizing the partition recovery procedures. The existing recovery techniques are categorized according to the coverage, communication, coordination, and connectivity in WSANs. The survey infers that the efficient cooperation and controlled mobility of the actor nodes can maintain connectivity in the network. The study presented in [1] categorizes various topology management techniques for fault recovery from network partitioning in WSNs. The comparison analysis of these techniques has been provided according to the failed nodes' count and simultaneity of node failures. The recovery techniques are categorized as precautionary and reactive. The different features of the techniques belonging to these categories are discussed. The survey also points out some open issues and less investigated ideas to address the recovery problem. Pallavi and Prakash [18] demonstrated detection and recovery approaches in WSANs. The paper classified the recovery approaches as reactive, proactive, and hybrid. The survey concludes that the researchers focused more on connectivity and neglected other issues such as lifetime, path length, and coverage. Various limitations identified in the existing network partition recovery solutions [5, 17, 19–23] are as follows:

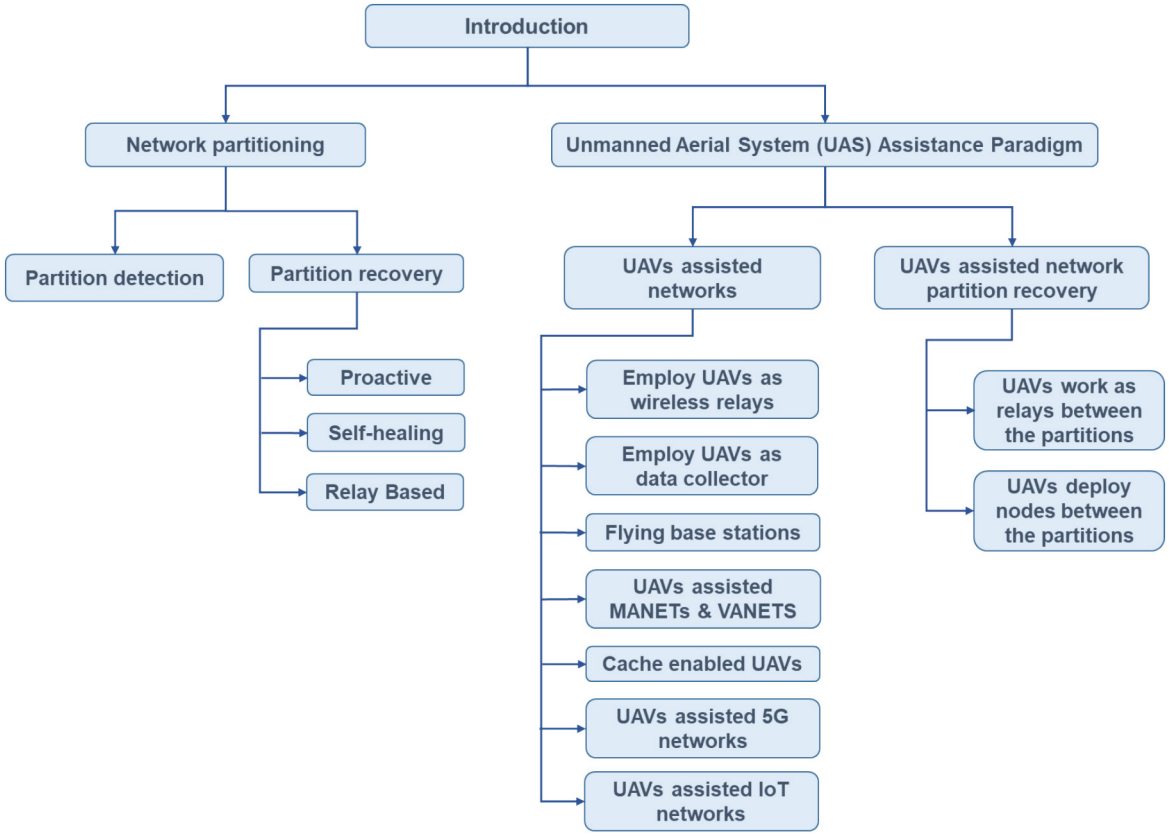



**FIGURE 3** | Taxonomy of network partitioning problem and UAVs' assistance paradigm.



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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 emphasis on detection and getting complete information regarding all the nodes in damaged network is less in the proposed recovery solutions.
- The previously reported techniques carry out the recovery process by assuming that the information about the damaged region has been obtained through satellites, drones, and mobile robots.
- Maximum recovery solutions assume that the network is obstacle-free and adequate extra nodes are available for recovery in all failure situations, which is impractical in real-life situations.
- In most reported techniques, the network partitions are connected without considering requirements of distinct network regions. However, the certain areas may have different connectivity demands, like the connections around the sink node should be more robust.
- Several recovery methods do not examine the size and shape of the partitions and represent them as a single node.
- Also, the maximum techniques do not confirm network recovery after executing the recovery steps.
- The far more complicated problem is decreasing the likelihood of future failures. The longevity of a WSAN following connection restoration has gained limited attention in the literature.

Recovery solutions can work better if accurate information about the failure scenario can be obtained in less time. This paper presents opportunities to incorporate UAVs to lessen the shortcomings of existing partition recovery approaches. UAVs are generally employed as communication providers to replace malfunctioning ground stations, increase coverage, fulfill the communication requirements of the users, and act as data collectors. Table 2 shows the list of notations referred in this paper, and Table 3 illustrates the review of different survey articles based on UAVs' assisted networks.

Alzahrani et al. [38] studied and summarized the UAVs' association with different applications, such as data gathering, routing, Internet of Things (IoT), cellular communications, and disaster management. The survey includes the description, categorization, and comparative assessment of the proposed UAV-assisted research algorithms. The critical challenges for each UAV-assisted application are described, and key references are also suggested for interested researchers who want to investigate UAV-assisted networks further. Hayat et al. [36] discussed the utilization of communication technologies such as ZigBee, Bluetooth, and LTE for aerial networks. The potential application domains of UAVs and the requirements of these domains, such as QoS, data transmission over the network, adaptability, connectivity, safety, security, scalability, and privacy, are also discussed. The qualitative or quantitative communication needs are identified to categorize different applications such as long range, short range, low capacity, high capacity, delay tolerant, and real time. The applicability of the current communication techniques for supporting aerial networking is investigated based on existing projects.

The revolution from voice-based to interactive multimedia–based communication has increased the connectivity and data rate

**TABLE 2** | List of notations.

| Notation | Description |
|---|---|
| 2CRNDC | 2-Connected relay node double cover |
| BIND | Boundary-aware optimized interconnection of disjoint segments |
| BS | Base station |
| C3R | Coverage conscious connectivity restoration |
| CDS | Connected dominating set |
| CISIL | Connecting isolated segments through intermittent links |
| CoRFL | Connectivity restoration mechanism using fuzzy logic |
| CRAFT | Connectivity restoration with assured fault tolerance |
| CRP | Cluster-based relay placement |
| DARA | Distributed actor recovery algorithm |
| DCR | Detection and connectivity restoration |
| DFS | Depth-first search |
| DT | Delaunay triangulation |
| ECR | Energy-centric optimized recovery |
| FH | Failure handler |
| FST | Full Steiner trees |
| IO-DT | Incremental optimization based on Delaunay triangulation |
| IoE | Internet of Everything |
| IoFT | Internet of Flying Things |
| IoT | Internet of Things |
| LDMR | Least distance movement recovery |
| LeDiR | Least-disruptive topology repair |
| LoS | Line of sight |
| MANETs | Mobile ad hoc networks |
| MDC | Mobile data carrier |
| MEC | Mobile edge computing |
| MILP | Mixed-integer nonlinear program |
| MSNs | Mobile sensor networks |
| NFV | Network function virtualization |
| NORAS | Node recovery through active spare designation |
| OACE | Obstacle-aware connectivity establishment |
| OBiC | Optimized biconnected |
| ORR | Optimized relay repositioning |
| PADRA | Partition detection and recovery algorithm |

(Continues)



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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)

| Notation | Description |
|---|---|
| PFR | Preventive failure recovery |
| QoS | Quality of service |
| RAM | Recovery algorithm for multiple node failures |
| RCR | Resource-constrained recovery |
| RFID | Radio frequency identification |
| RIM | Recovery through inward motion |
| RNC | Relay node count |
| RNFR | Rotating nodes based failure recovery |
| RTN | Round table negotiation |
| SBC | Single-board computer |
| SMT | Steiner minimum tree |
| TPU/GPU | Tensor/graphical processing unit |
| UAS | Unmanned aerial sensing |
| UAVs | Unmanned aerial vehicles |
| UCC | UAV control center |
| VANETs | Vehicular ad hoc networks |
| VCR | Volunteer-instigated connectivity restoration |
| WSN | Wireless sensor network |
| WSANs | Wireless sensor and actor networks |
| WSRNs | Wireless sensor and robot networks |

demands, which require new technologies to support these 5G and beyond wireless networks. Authors in [34] have considered UAVs for these advanced wireless networks because of low cost, mobility, line-of-sight (LOS) communication, and maneuverability features. The challenges of the UAV-associated networks are also highlighted, along with various benefits. Mainly, three use cases of wireless networks, such as flying relay, user equipment, and flying BS, are presented along with resource allocation and UAV placement problems. Dai et al. [30] surveyed mechanisms and frameworks developed for UAV-assisted networks. The architecture of these networks is overviewed according to four aspects (computing, sensing, communication, and application) and four domains (technology, framework, challenge, and solution). Moreover, the application and implementation of UAV-assisted networks along with the challenges and reported solutions are also discussed. Lastly, the authors have discussed the advanced techniques and research directions for UAV-associated wireless networks. Taimoor et al. [31] discussed the problems associated with resource management in aerial networks and ground networks (interference, inadequate bandwidth, and power). Communication, control, computing, and caching (4Cs) are the resources the network entities require for successful aerial networking. The resource management by formulating optimization problems based on the combination of these resources is discussed along with performance metrics, challenges, solution approaches, and future scope for resource management.

Alsuhli et al. [32] presented different roles of UAVs in communication processes. The role of the UAVs is categorized as communication provider, consumer of communication services, and relay. Zaidi et al. [35] addressed the constraints in deploying UAVs in IoT, such as signal interference and collision, limited resources, and intermittent accessibility. The paper has covered key concepts related to the Internet of Flying Things (IoFT), including description, classification, and comparison of various proposed works. Moreover, the application, challenges, and future scope of IoFT are also discussed in the paper. Rovira-Sugranes et al. [29] presented the application of AI for accommodating the requirements of dynamic aerial networks, such as efficient control, sensing, and communication. The paper has reviewed AI-enabled routing algorithms for UAV networks. Along with that, the role of AI is emphasized in UAV swarming, mobility, and networking models, and simulation tools are also discussed. Lastly, the challenges and future trends of AI-enabled UAV communication for different networking aspects are discussed. Liu et al. [33] presented the utilization of UAVs to deal with the expectations of the Internet of Everything (IoE), i.e., diversity, intelligence, and scalability, along with certain constraints such as security, coverage, and computing. Additionally, the opportunities and solutions to introduce the services of UAVs in IoE are discussed.

## 2.1 | Partition Detection

The network partitioning causes connectivity loss, due to which the network nodes cannot determine the span of the damage. Before proceeding with the recovery process, it is essential to verify network partitioning, and along with this confirmation, the recovery process requires information about total partitions in the damaged network and nodes in these partitions. Table 4 shows the summary of failure and partition detection approaches.

Baroudi et al. [39] attempted to detect partitioning by detecting some events in the network, such as the unexpected drop in messages, reduced reply messages while managing the two-hop neighbor's list, and single node incapable of communicating with its neighbors. Their proposed approach also reduces the message overhead by selecting a partition supervisor who lost the maximum neighbors. In [6], ideas are specified for sensing the failure scope by surveying the network using drones, robots, and the representative node from the segments. The approach presented in [7] also determines the node failures by evaluating the lost contact with their neighbors. Akkaya et al. [8] focus on detecting the failure of upstream neighbors, and the SNs confirm their failure when acknowledgement from them is not received. The algorithm attempts to apply the backoff mechanism of the network layer. In this mechanism, SNs send NEWPATH requests to their neighbors, and this querying process stops after attaining a certain number of hops by setting the k or time-to-live (TTL) value. If node does not receive a reply within a time-out value, then it assumes network partitioning. In [42], cut vertices in the network are identified initially, and the algorithm confirms network partitioning if a cut vertex fails.

Akkaya et al. [43] proposed distributed approach to determine the network's connected dominating set (CDS). Based on CDS,



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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** | Review of existing surveys on UAVs' assisted networks.

| Survey and year | Overview | Application domain | Applicable use cases | Key research challenges | Research gaps |
|---|---|---|---|---|---|
| [24] 2024 | UAV-assisted IoT applications | IoT communication | Resource management, connectivity, backhaul challenges, trajectory challenges, interoperability in UAV-assisted IoT communication | QoS challenges due to antenna transmission and reception | Existing simulation tools for UAV-assisted IoT communications can be discussed in detail |
| [25] 2024 | Multi-UAV networks | Disaster monitoring | Damage assessment, search and rescue, data collection, infrastructure inspection, resource delivery | Trajectory optimization, task scheduling, and data processing in UAV networks | Use cases in UAVs in different disaster scenarios are reviewed in detail |
| [26] 2023 | UAV-assisted data collection | IoT communication | UAV-based data collection from sensors, agricultural scenarios, emergency situations, smart cities | Channel modeling, energy limitation, interference, mobility, time-sensitive data, security, latency | Simulation tools could be investigated in detail. |
| [27] 2023 | Network management in UAV-enabled communication systems | Use cases of software defined network and network function virtualization | Cellular communication, satellite communication, security, placement, routing, UAV-assisted VANET and MANET | Wireless power transfer, 6G requirements, quantum communication, AI integration, UAV optical communication, MEC, blockchain-based security, catching | Simulation tools could be investigated in detail. |
| [28] 2023 | Network slicing | Role of UAV in 5G network slicing | UAV acts as a BS, data collector, MEC node, relay in different network slices | Security, energy, and handover among networks | UAVs' role in network slicing can be discussed in detail |
| [29] 2022 | AI-based routing, mobility, networking models, and simulation tools | UAV-assisted networks | Imaging, virtual reality, autonomous path control, resource allocation, sensing, security, positioning, channel estimation | Operational regulations, AI integration, connectivity, routing, energy efficiency, spectrum management, privacy and security | The applicability of AI-based routing is not discussed in detail |
| [30] 2022 | Description of UAV-assisted networks, implementation challenges, and advanced solutions | UAV-assisted networks | UAV-assisted aerial relaying, sensing and rescuing, mmWave communication, energy harvesting, cognitive UAV communication, mobile edge computing and caching, UAV-aided B5G networks, network function virtualization (NFV), software-defined UAV networks | Energy efficiency, space-air-ground networks' integration and quantum communication among these integrated networks, privacy, security, mmWave communication with artificial intelligence, and semantic communication in UAV networks | Applications of advanced UAV networks are not discussed |

(Continues)



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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** | (Continued)

| Survey and year | Overview | Application domain | Applicable use cases | Key research challenges | Research gaps |
|---|---|---|---|---|---|
| [31] 2022 | Resource management in UAV-assisted wireless networks | UAV-assisted networks, B5G, and 6G | Management of 4C's for the UAV-assisted and future B5G/6G networks | Dynamic and blockchain-enabled 4C's management, resource management for satellite-aerial-ground networks and UAV-assisted MEC, IoT devices | Analysis of 4Cs for different UAV-assisted networks |
| [32] 2022 | Role of UAVs as communication providers, relay, and consumer | UAV-assisted communication and applications | Smart farming, environment monitoring, aerial base stations, vehicular network recovery, LTE network offloading, healthcare, disaster management | Intercell interference in UAV networks, onboard energy, stability of mmWave communication using UAVs, security of Wi-Fi networks | UAVs' role as data collector is not discussed |
| [33] 2022 | Utilization of UAVs' deal with the challenges of IoE | Internet of Everything (IoE) | Ubiquitous connections, aerial intelligence, sensor powering, intelligent transportation system, rescue, logistics, aerial surveillance, unmanned military missions | Resource allocation, security, less complex AI algorithms, universal standard methods, coordinating computing facilities | Applications of UAV-assisted IoE are not discussed |
| [34] 2021 | Use cases and challenges for employing UAVs in 5G and beyond networks | 5G or beyond 5G (B5G) | UAVs' placement as flying relay, base station, and user equipment | Resource allocation, user association, channel modeling, energy, 3D UAV placement, privacy and security, backhaul connectivity, and trajectory planning | Examine more UAV use cases involving 5G technology |
| [35] 2021 | UAVs' deployment in IoT | Internet of Flying Things (IoFT) | Smart agriculture, environmental monitoring, disaster management, video surveillance, flying cloud, edge, fog, computing networks, flying cellular networks | Interference and collision, UAVs' deployment and selection, path planning, energy consumption, date rate, coverage, UAVs' management and control, privacy, and security | Analysis of IoFT challenges for efficient computing, storage and transmission of data collected by UAVs to Internet servers |

(Continues)



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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** | (Continued)

| Survey and year | Overview | Application domain | Applicable use cases | Key research challenges | Research gaps |
|---|---|---|---|---|---|
| [36] 2020 | UAVs' assistance in routing, IoT networks, data gathering, monitoring, cellular communication, disaster management, and secure communication | UAV-assisted networks | Intelligent surveillance, ground networks, UAV-WSN and UAV-IoT-based data gathering, traffic monitoring, prior-disaster monitoring, spectrum sharing, smart cities, agriculture, cybersecurity, and edge computing | Energy consumption, complex environments, optimal altitude for data collection, optimal UAVs' placement, UAVs' mobility, restricted flying times, intermittent connectivity, vulnerable communication channel | Investigate on the visible light communication concept and its usability in UAV networks. |
| [37] 2016 | The features and requirements of UAV-associated networks for the civil applications | Aerial networking | Connectivity, safety, adaptability, security, scalability, and privacy requirements for aerial networking in search and rescue, coverage, and delivery applications | Feasibility of existing IEEE 802.11, Wi-Fi technologies for large-scale UAV networks, improvement in existing media access protocols, security and privacy, MAC layer design, intermediate depots for delivery applications | To assess whether Wi-Fi is the feasible communication technology for UAV applications |

some dominator nodes are selected according to certain criteria. Afterwards, DFS is applied to these nodes to identify cut vertices in the network. Partitioning is confirmed when a cut vertex fails and the failure handler (FH) with that cut vertex initiates the recovery process. In [44], the partitions are detected using different methods, such as detecting abrupt communication drop, concurrent node failures, and the inability to reach certain remote nodes. Reachable nodes are computed in cluster-based relay placement (CRP) algorithm [22] and round table negotiation (RTN) [41] by sharing one-hop neighbor lists. These lists are exchanged between the network nodes in different iterations to compute the final connected components in the network. The UAVs collect information about the failed and active nodes in [10] and provide collected information to UAV control center (UCC). UCC rearranges the information provided by UAVs and applies DFS to find the network's connected components. If connected components are more than one, then UCC confirms network partitioning.

## 2.2 | Partition Recovery

The techniques for the network recovery from partitioning [18, 45–47], can be categorized into three groups: (1) proactive approaches, (2) self-healing, and (3) relay-based approaches (RN placement and data transfer using MDCs). Table 5 provides an illustration of these recovery schemes. The following subsections provide brief discussions and a literature review of the partition recovery approaches.

### 2.2.1 | Proactive Recovery Techniques

Connectivity with sink node is eminent in WSANs; hence, the redundant SNs are also set up in the network to endure some node failures. The proactive techniques focus on sustaining the network connectivity by pre-planning the network recovery from node failures to mitigate network partitioning [18, 43]. The reported techniques categorize proactive techniques into two groups. The techniques in the first group deploy redundant nodes to establish alternative paths among the pair of nodes. The alternate paths in the network must be separate such that a node failure could not break feasible routes. This process of creating disjoint paths refers to establishing k-connectivity. In k-connected network ($k \geq 2$), failure of $(k-1)$ nodes cannot fragment the network. The methodologies in the second group allocate additional nodes in network for the critical nodes (cut-vertex nodes). The failure of cut-vertex nodes causes partitioning, and these nodes are identified by employing various centralized or distributed methods [2, 18, 55]. Table 6 describes the proactive techniques for network recovery.

In [59], full fault-tolerant and partial fault-tolerant relay node placement (RNP) algorithms for heterogeneous WSNs are proposed. In heterogeneous networks, the nodes have different transmission radii. Full fault-tolerant algorithm establishes k-connectivity between all the sensor and relay node pairs. A partial fault-tolerant algorithm establishes k-connectivity between some selected pair of network nodes only. The variants of these two algorithms are also established to analyze different types of



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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** | Partition detection techniques.

| Reference | Failure type | Detection scheme | Type of detection | Advantages | Drawbacks |
|---|---|---|---|---|---|
| [21] 2024 | Single node | Sink identifies failure of cut-vertex node | Partition detection | Confirms network partitioning | Increased message overhead, UAV is required to identify actual failed cut vertex |
| [37] 2024 | Multinode | No response from neighbor nodes | Failure detection | — | Unable to detect network partitioning |
| [23] 2024 | Multinode | No response from neighbor nodes | Failure detection | — | Unable to detect network partitioning |
| [11] 2023 | Multinode | UAVs collect connected component sets from healthy network nodes and provide to UCC | Partition detection | Confirms network partitioning efficiently | Increases total messages transmitted count |
| [9] 2022 | Multinode | UAVs gather information about the network, and UCC confirms partitioning based on the data provided by UAVs using DFS | Partition detection | Confirms network partitioning efficiently | Area exploration process by UAVs is time consuming |
| [39] 2020 | Multinode | Observes sudden drop in reply messages, selects partition supervisor to announce failure | Failure detection | Reduces message overhead | Unable to detect network partitioning |
| [40] 2020 | Multinode | Computes reachable nodes by sharing one-hop neighbors to identify connected components in different iterations | Partition detection | Identifies network partitions if the network is partitioned | Cannot identify network partitioning accurately, Increased message overhead |
| [22] 2018 | Multinode | Computes reachable nodes by sharing one-hop neighbors to identify connected components in different iterations | Partition detection | Identifies network partitions if the network is partitioned | Cannot identify network partitioning accurately, increased message overhead |
| [5] 2017 | Multinode | Presents different ways to survey the network | Detection of failure and its scope | — | Methods given for detection are not implemented |
| [41] 2016 | Multinode | Evaluate lost contact with neighbors | Failure detection | — | Unable to detect network partitioning |
| [42] 2014 | Single node | Detects the failure of cut-vertex node | Partition detection | Confirms network partitioning accurately | Increased message overhead with two-hop neighbor's information |
| [8] 2013 | Multinode | Detects failure of upstream neighbors, failure detection process employs backoff mechanism of network layer | Failure detection | — | Unable to detect network partitioning |
| [43] 2010 | Single node | Determines CDS of network to identify cut vertices | Partition detection | Confirms network partition accurately | Increased message overhead with two-hop neighbor's information |
| [44] 2010 | Multinode | Detects communication drop, concurrent node failures, unreachability to remote nodes | Failure detection | — | Methods given for detection are not implemented |



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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** | Strategies for network connectivity restoration.

| References | Recovery approach | Description | Advantages | Limitations |
|---|---|---|---|---|
| [17, 19, 43, 48] | Redundant nodes' deployment | Provisioning redundant nodes as backups or for establishing k-connectivity | Improved network reliability | Significant deployment expenditure, huge communication overhead |
| [20, 22, 49, 50] | Relocation of active nodes | Network recovery by relocating healthy nodes in different partitions. | External resource provisioning is not required | Requires complex hardware, requires mobile active nodes |
| [3, 11, 21, 51, 52] | RN deployment | Deploy additional nodes for network restoration | Less overhead on network nodes, recovers network efficiently | Assume that enough RNs are available for deployment |
| [23, 53, 54] | Mobile data carrier (MDCs) for data transfer | Temporary recovery scheme which employs MDCs to provide intermittent links and to transfer data between different partitions | Works in resource-constrained sensor network | Increases the data latency |

fault tolerance by coupling them with one and two-way paths that exist among the nodes. Hao et al. [65] established polynomial time approximation method to place less RNs in such a way that two-node disjoint paths exist between all the pairs of RNs. The method is created by formulating 2-connected relay node double cover (2CRNDC) optimization problem. In case of RN failure, the SNs communicate with its designated backup RN. Bai et al. [63] find an optimal RN deployment pattern to accomplish k-connectivity ($k = 4$), and full network coverage. They proposed a "diamond" pattern, which coincides with triangle lattice pattern and a "square" pattern by changing some values. These patterns are examined systematically using mathematical theories.

Yun et al. [60] established set of new patterns to achieve (3-6)-connectivity. A universal hexagon–based pattern is also discovered in their approach. By varying this pattern, different types of connectivity can be achieved. Turjman et al. [58] proposed fault-tolerant optimized relay repositioning (ORR) to improve the network lifespan. ORR is a grid-based scheme that reduces the search space and locates the optimum RN positions that can retain more energy in different rounds. This node positioning problem is devised as mixed-integer linear program. Simulation results for ORR show that it sufficiently increases the network lifetime compared to other heuristic approaches. Imran et al. [57] proposed distributed partition detection and connectivity restoration (DCR) approach. DCR determines the critical nodes in network in a distributed manner and assigns them backup nodes. The backup nodes monitor their associated critical nodes. If critical node damages, its backup node performs network recovery. DCR is further extended to a recovery algorithm for multiple node failures (RAM) to handle multiple critical nodes' failures.

Vaidya and Younis [61] presented prefailure management strategy, node recovery through active spare designation (NORAS) to improve coverage and network connectivity. NORAS allots backups from one or two-hop neighbors of identified critical nodes. When critical node fails, the backup node replaces its position, or the movement of multiple nodes is initiated in a cascaded manner on the direct shortest path among the failed and back up node. Zamanifar et al. [62] focussed on reducing the message overhead and restoring actor–actor communication on the failure of critical actor node. The critical nodes in the network are identified using CDS of network. The candidate replacement nodes for the critical nodes are also selected in a hybrid way. When critical actor node damages, the neighbors of the damaged node communicate to initiate the recovery process. Akkaya et al. [43] presented partition detection and recovery algorithm (PADRA) to decrease the message overhead and restore connectivity. PADRA in advance determines the criticality of each node using CDS of the network in a distributed way. After determining the cut vertices, neighbor nodes are assigned to them to handle their failure. The neighbor node opts for dominatee node using greedy or dynamic programming algorithm to replace the damaged node. The movement load is shared using the cascaded relocations of the nodes present in the path of failed cut vertex and dominatee. PADRA is further extended to MPADRA to deal multiple node failures. Ranga et al. [42] presented distributed prioritized connectivity restoration algorithm (DPCRA) that proactively



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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** | Proactive recovery techniques.

| Reference | Objectives | Preventive methodology | Fault tolerance strategy | Parameters optimized | Simulator |
|---|---|---|---|---|---|
| [56] 2024 | Increase network lifetime | Cluster heads with more energy are selected | Sink selects the new cluster head from the available nodes | Terrain dimension, supported degree, total nodes in terrain, cluster size, transmission range | MATLAB |
| [42] 2014 | Balance energy consumption, enhance network life time. | Allots FHs to critical nodes. | Multiple FHs initiate the recovery process when a critical node fails. | Movement distance of nodes, number of nodes moved, message overhead, coverage reduction, energy consumption | NS2.34 |
| [57] 2012 | Fault tolerance | Backup node assignment for critical nodes | On failure of the critical node, backup node performs recovery | Travel distance, total nodes moved, exchanged messages count, percentage coverage reduction, average node degree | — |
| [58] 2011 | Fault tolerance and prolonged network lifetime | Select relay node positions that can save energy | Set up nodes in grid vertices and devise this problem as MILP | WSN lifetime | MATLAB |
| [59] 2010 | Fault tolerance, increased network connectivity in heterogeneous WSN | Establish full and partial k-connectivity ($k \geq 2$) | Deploy extra relay nodes to form k-vertex disjoint paths. | One-way and two-way full/ partial fault tolerance | QualNet |
| [60] 2010 | Full coverage | Establish k-connectivity ($k = 3, 4, 5, 6$) | New patterns are developed to establish (3-6)-connectivity, universal hexagonal pattern is discovered | Impacts of practical and communication sensing model | — |
| [43] 2010 | Reduce message overhead and restore connectivity | Selection of neighbor nodes for recovery | The neighbors of failed cut vertex selects dominatee node for replacement | Number of messages, movement distance of all nodes | — |
| [61] 2010 | Improve network connectivity and coverage | Backup node assignment for critical nodes | Failure of critical node triggers movement of spare node or cascaded movement of multiple nodes in the path of spare node | Total traveled distance, number of relocated nodes, reduction in area coverage | Visual C++ |
| [62] 2009 | Reduce message overhead and restore connectivity | Candidate replacement nodes for critical node | The neighbors of failed critical node initiate the recovery process | Number of transmitted messages, total movement by actors | — |
| [63] 2008 | Network coverage | Establish k-connectivity ($k = 4$) | "Diamond" pattern is presented to accomplish 4-connectivity | Number of nodes required to achieve 4-connectivity by different patterns | Tested mathematically |

(Continues)



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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** | (Continued)

| Reference | Objectives | Preventive methodology | Fault tolerance strategy | Parameters optimized | Simulator |
|---|---|---|---|---|---|
| [64] 2006 | Increase network efficiency, and fault tolerance | Building k-connected dominating set (CDS) | Establishing k disjoint paths in CDS of network | Backbone size, success ratio, tolerable failure ratio, message overhead | Custom simulator ds |
| [65] 2004 | Network scalability and prolonged lifetime | Establish k-connectivity ($k = 2$) | Deploy extra relay nodes to form 2-relay-node vertex disjoint paths. | Error ratios based on node density | — |

determines critical nodes and assign FHs. The FHs detect the failure and employ multiple backup nodes to repair the network. Instead of single FH, DPCRA selects multiple backup nodes for network recovery.

### 2.2.2 | Self-Healing Recovery Approaches

The majority of node repositioning strategies reorganize the network structure intending to lower energy consumption due to movement. The solution is quick since extra node placement is not required. These strategies vary in terms of the ways to select the nodes and compute new locations to relocate them. Euclidean distance between the nodes and their new destinations is the prime criteria adopted by these techniques to reduce the movement of the nodes. Whether these techniques are distributed or centralized, they assume that the terrain in which the nodes are moving is uniform and obstacles are not present in the environment. Hence, all the movements are considered successful, which may not be possible. Therefore, these movement algorithms should be revised to incorporate path planning mechanisms according to the terrain. Moreover, the long-distance movement of the SNs is not favored because it straightaway increases the energy consumption [8, 66–69]. Table 7 presents a review of self-healing approaches for network recovery.

Younis et al. [66] presented distributed "recovery through inward motion" (RIM) to reduce the travel distance of the network nodes during network reorganization and minimize message overhead. The network nodes move toward the failed node's location until they come into each other's transmission radius. The network reorganization continues recursively until the network becomes connected. RIM is tested through simulations and analyzed mathematically. Alfadhly et al. [67] proposed optimal node reorganization technique for one or more node failures to maintain the network coverage and minimize travel distance. The problem is framed as an integer linear program (ILP). The results show that the algorithm works better than distributed actor recovery algorithm (DARA) [77]. Chouikhi et al. [7] aimed to restore connectivity in multiple channel WSNs and proposed two centralized schemes: preventive failure recovery (PFR) and rotating nodes based failure recovery (RNFR). Both approaches employ the Steiner tree heuristics and graph coloring method. PFR is a proactive approach that works before the cut-vertex node fails and informs recovery decisions to the respective network nodes. RNFR works as a reactive approach and triggers recovery process after failure. The recovery process involves the cascaded reorganization of nodes toward the damaged part of the network. The evaluation results of these approaches are better than the existing RIM [66] approach.

The network reorganization using the neighbors of failed node sometimes breaks more connections, and the relocation process continues repeatedly. Imran et al. [69] proposed volunteer-instigated connectivity restoration (VCR) to address this problem. VCR employs the neighbors of failed node and checks how they can utilize their communication range. These neighbors voluntarily move toward the failed node. VCR avoids collision around the failed node by applying the diffusion force among the moving nodes according to transmission range. The results show that VCR performs efficiently with less message



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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 7** | Self-healing recovery approaches.

| Reference | Failure type | Network state | Recovery operation | Recovery technique | Parameters optimized | Simulator |
|---|---|---|---|---|---|---|
| [70] 2024 | Multiple node | 1-Hop | Distributed | Backup nodes move toward the location of failed critical nodes | Moving distance, message overhead, total relocated nodes | Coverage loss |
| [40] 2020 | Multiple node | 2-Hop | Distributed | Fuzzy logic employed to select the candidate nodes based on some parameters for repositioning. | Travel distance, energy, message overhead, time complexity. | MATLAB |
| [41] 2018 | Multiple node | 1-Hop | Distributed | Self-identification of clusters, negotiators meet to identify nodes for relocation and their new locations. | Reconnection time, travel distance, deployed node count | MATLAB v9 (R2015b) |
| [8] 2017 | Single node | 2-Hop | Centralized | Employs Steiner tree heuristics and graph coloring for network recovery through cascaded reorganization of nodes | Number of conflicts, control messages, recolored and relocated nodes, travel distance | Omnet++, MATLAB |
| [71] 2016 | Multiple node | 1-Hop | Centralized | Select boundary nodes that require least movement cost. | Success probability to establish biconnectivity, total nodes moved, travel distance | Customized C++ simulator |
| [45] 2016 | Multiple nodes | — | Distributed | Skeleton of the network is computed in advance to immediately start the recovery after failure | Number of RNs, Travel distance | CGAL computation library |
| [9] 2013 | Multiple node | 1-Hop and k-hops toward the sink | Distributed | Leader nodes exploit old routes to re-establish connection with the sink. | Recovery delay, false alarms, message overhead, travel distance. | Network simulator NS-2 |
| [72] 2013 | Multiple node | 1-Hop | Distributed | Relay nodes move toward the center | Travel distance, relay nodes' count | — |
| [5] 2012 | Multiple node | 1-Hop | Distributed | Game theory is employed to select the nodes for relocation. | Travel distance, coverage | — |
| [73] 2012 | Multiple node | 1-Hop | Centralized | Two heuristic methods select the nodes in partitions and identify their new locations. | Movement distance, recovery delay | — |
| [74] 2011 | Multiple node | — | Centralized | The recovery problem is mapped to MILP. Branch and bound algorithm solves MILP | Travel distance, delay | AMPL and Gurobi Optimizer 4.0 |
| [68] 2011 | Single node | 2-Hop | Distributed | Neighbor nodes move toward the failed node, and noncritical nodes move toward the location of relocated nodes. | Travel distance, coverage loss | MATLAB |

(Continues)



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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 7** | (Continued)

| Reference | Failure type | Network state | Recovery operation | Recovery technique | Parameters optimized | Simulator |
|---|---|---|---|---|---|---|
| [75] 2011 | Single node | 1-Hop | Centralized and distributed | Nearest block moves toward the failure location | Travel distance, number of relocated nodes and messages exchanged | Visual C++ |
| [66] 2010 | Single node | 1-Hop | Distributed | Neighbor nodes move until they reach in transmission radius of other moving nodes | Travel distance, number of relocated nodes and messages exchanged | — |
| [67] 2010 | Multiple node | — | Centralized | Network reorganization is formulated as integer linear program problem | Travel distance, coverage loss | GNU linear programming kit (GLPK), MATLAB |
| [69] 2010 | Single node | 1-Hop | Distributed | Neighbor nodes move toward the failed node without causing disconnections | Distance traveled, message overhead, total nodes moved, coverage change percentage | — |
| [76] 2010 | Single node | 1-Hop | Distributed | Failed node's neighbors provide intermittent connectivity by relocating according to their turns. | Travel distance, number of nodes relocated and messages exchanged, coverage reduction, network lifetime | |

overhead and travel distance. Alfadhly et al. [68] presented least distance movement recovery (LDMR). LDMR moves the failed node's neighbors toward its position. The positions of the relocated nodes are taken by selected non–cut-vertex actor nodes in the network based on their distance. The results show that LDMR works better than RIM [66]. Abbasi et al. [75] relied on the local information for planning network recovery and provided least-disruptive topology repair (LeDiR) algorithm. When the network nodes detect failure, they check their routing table and decide their role. The neighbor and the other nodes that belong to the smaller block move to ensure network recovery. LeDiR works well in dense networks with less reorganization and communication overhead. Akkaya et al. [8] focused on distributed partition recovery scheme that first detects the failure of upstream nodes and selects leader nodes from each partition to initiate recovery process. The SNs maintain information regarding the route to node. The leader SNs exploit this information and visit every node in their stored route until they find an active network node connected or on the path to the sink. Accordingly, nodes in different partitions move with the leader node in a cascaded manner. Baroudi et al. [39] focused on recovering the partitioned WSRNs. The proposed approach, "connectivity restoration mechanism using fuzzy logic" (CoRFL), is based on node repositioning and focuses on noteworthy parameters such as network lifetime, environmental conditions, and robot's battery level. This multiobjective optimization problem is solved using fuzzy logic, which selects the finest candidates for movement. CoRFL is further extended to CoRFL2 and CoRFLN, if the selected candidate is cut vertex. CoRFLN relocates nearest neighbors to the position of the recovery candidate, and CoRFL2 uses fuzzy logic to pick the replacement team based on the energy level. The proposed algorithm is evaluated using prototype experiments and simulations.

Shriwastav and Ghose [22] addressed the connectivity restoration problem for several disjoint clusters using RTN approach. The surviving clusters in the network execute an iterative self-detection procedure to determine active neighbors and other nodes in the clusters. The negotiators from each cluster then move to a predetermined location to engage in the RTN. The negotiators exchange information about their clusters and decide the connection paths to link the boundary nodes. The reconnection paths are decided based on the failed nodes' locations, and nodes are allocated for deploying on these paths. The negotiators then come back and start the reorganization process. RTN shows less recovery time as compared to some existing methods. Joshi and Younis [72] presented distributed approach for reconnecting disjoint segments (DarDs) to relink the fragmented network. DarDs populate the nodes from every segment toward the middle point of the network. Afterwards, it applies incremental optimization based on Delaunay triangulation (IO-DT) to determine the minimum node count needed to sustain connectivity and return extra deployed nodes to their cluster. The results demonstrate that DarDs lessen the travel overhead of nodes and connect the network with fewer RNs compared to some existing works.

Senturk et al. [4] focused on reducing the repositioning cost of the nodes in delay-tolerant WSNs while re-establishing network connectivity. The proposed approach can work in uncertain failure conditions with limited knowledge and employs game



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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

theory to select the nodes for relocation. In the proposed algorithm, each partition identifies the Nash equilibrium of other partitions to define the probability distributions of other partitions. The nodes in a partition move closer to the partition with the higher probability distribution. The algorithm performs almost similar to the centralized sub-optimal method. Senturk et al. [73] proposed a connectivity restoration mechanism using two heuristic approaches. The first heuristic method identifies the minimum relay node count (RNC) and the deployment locations required to re-establish connectivity. In the second heuristic method, nodes in different fragments that will not cause further partitioning are identified using the CDS of the network.

Sir et al. [74] aimed to reduce the travel distance of network nodes during recovery and modeled the connectivity restoration as a "mixed-integer nonlinear program" (MILP). The polygon approximation performs the linearization of nonlinear terms in the model. The branch and bound method solves the mixed-integer problem. The results obtained through simulations illustrate that this scheme outperforms the heuristic mechanisms. Joshi and Younis [45] proposed geometric skeleton–based reconnection (GSR) scheme to reconnect partitioned WSN in distributed mode. GSR models the network area containing disjoint segments as skeleton, and mobile relays in different segments move toward this skeleton to restore connectivity. GSR is a pre-failure strategy, and all the network nodes save the computed skeleton in order to start recovery immediately after failure. The relocated node count and travel distance of relocated nodes in the GSR approach are less as compared to DarDs [72] and DORMS [44].

### 2.2.3 | Relay-Based Approaches

In some cases, nodes in the damaged network may not be sufficient in numbers or mobile for network recovery. Deploying new RNs is also an alternative method to reconnect the partitioned segments. These nodes are deployed manually, and the information about the partitions, location of nodes present in these partitions, failed nodes and their locations is presumed. The techniques in this category are centralized and focus on reducing the deployed RNC. The centralized approaches are challenging to apply when assessing the scope of the damage is not feasible due to unreachable damaged regions. In such cases, the locations to deploy new RNs may not be determined accurately.

The primary goal of RNP strategies is to reduce the RNC and compute the deployment locations [4, 52, 80, 82, 83]. In case the RNs in network are inadequate for the recovery, the MDCs or data mules are utilized to traverse various network partitions in order to transmit information. The MDCs are assumed to possess sufficient storage capacity due to the advanced computing and communication capabilities of modern devices. Nevertheless, this approach suffers from inefficiency due to increased overhead and data latency. Table 8 presents the summary of reported recovery methods based on RN deployment.

Ranga et al. [52] aimed to reconnect the fragmented network. The proposed algorithm picks up representative nodes from each fragment and computes the convex hull of these selected nodes. Afterwards, the algorithm finds zero gradient points within the convex hull and obtains positions for the new RNs, which are deployed toward the calculated point. Senel and Younis [80] focused on providing connected topology and devised approaches that compute discrete Fermat point (DFP) for all the possible sets of triangles and then select the best triangles among them that requires minimal extra RNs for connectivity. Lee and Younis [83] proposed an optimized RNP algorithm using a MST on convex hull (ORC) to determine Steiner points (SPs) for positioning extra RNs to connect partitioned fragments. ORC identifies the SPs in different rounds and deploys the extra nodes inwards from the boundary of convex hull. The results demonstrate the effectiveness of ORC in generating efficiently connected topology with fewer RNs. Senel et al. [3] strive to provide robust network topology with extended coverage during partition recovery. The proposed scheme identifies the boundary partitions and computes these segments' center of mass (COM). Primarily, the 1C-SpiderWeb approach is applied, and RNs are deployed in a ring formation to connect the segments in the network's boundary. After that, 2C-SpiderWeb is applied to create two disjoint vertex connections between all the segments by joining the network partitions to the ring through alternative paths. The experimental results of 1C-SpiderWeb and 2C-SpiderWeb are compared with baseline techniques. The results demonstrate that network topologies created from both algorithms provide more coverage and average node degree with reduced path length.

Abbas and Younis [49] solved the problem of federating disjoint fragments in network using fixed number of mobile and stationary nodes. The MDCs travel around the network partitions and employ RNs to decrease the tour length of MDCs. The convex hull of terminal is computed, and RNs are added inwards to act as access points. The algorithms then attempt to find the shortest tours for MDCs to cover all the terminals through relays. The results illustrate that the tour length of MDCs is less than other recent schemes. Lalouani et al. [5] presented boundary-aware optimized interconnection of disjoint segments (BIND) that generate full Steiner trees (FSTs) of some boundary nodes and concatenates these FSTs to build a SMT. The segments are modeled as hypergraphs to find the combinations of terminal nodes. BIND employs Geosteiner methods to resolve the exponential growth of FSTs. The results demonstrate that BIND is advantageous compared to other competing schemes.

Lalouani et al. [47] presented a mechanism to recover the network using mobile nodes' intermittent links when sufficient RNs are not available in network. The connecting isolated segments through intermittent links (CISIL) algorithm employs DT to identify the possible tours among the partitions. The optimal set of tours is selected from the available network tours. CISIL formulates this selection process to k-edge MST problem in a hypergraph of the computed tours. Kumar and Amgoth [40] presented a CRP algorithm that first finds two nearest partitions based on their centroids. After that, convex hull of the partitions is computed to determine the boundary nodes to connect the selected centroids. This process repeats iteratively till the network becomes connected. Experimental results demonstrate that CRP performs better than ORC [83] and CIST [84]. Lee et al. [6] aim to produce fault-tolerant topology and proposed CRAFT algorithm. CRAFT creates an inner ring by computing the partitions' convex hulls, and SPs iteratively. The ring forms



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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 8** | Relay node deployment approaches.

| Reference | Failure type | Network state | Recovery operation | Recovery technique | Parameters optimized | Simulator |
|---|---|---|---|---|---|---|
| [11] 2024 | Multiple nodes | — | RN deployment | UAVs place the relay nodes in computed regions | Number of relay nodes (RNs), convergence time, delay, packet delivery ratio, travel distance | MATLAB, NS2.35 |
| [5] 2024 | Single node | — | MDC | UAV moves to the location of the identified failed node to relay messages among the partitions | Detection time, travel distance, messages transmitted, recovery time | Omnetpp 5.4.1 |
| [24] 2024 | Multiple nodes | 1-Hop | MDC | MD carriers find obstacle-free path to collect data from segments | Travel distance, travel time, deviation in tour length, total sojourn locations, latency, energy, end-to-end delay | NS2.34 |
| [12] 2023 | Multiple nodes | 1-Hop | RN deployment | UAVs identify partitions and reconnected them in biconnected topology | Travel distance, relay node count, detection time, recovery time, message transmitted | Omnet++ |
| [46] 2021 | Multiple nodes | 1-Hop | MDCs' tour network partitions | Obstacle-free paths for MDCs are computed for data transportation. | Average latency, path length, energy consumption | — |
| [22] 2020 | Multiple nodes | 1-Hop | RN deployment | Nearest partitions based on the centroids are computed in an iterative manner | Number of RNs | Python |
| [78] 2019 | Multiple nodes | 1-Hop | RN deployment | Quadrilateral Steiner trees are used to compute obstacle-free RN positions | Number of RNs, node degree | — |
| [47] 2018 | Multiple nodes | — | MDCs' tour network partitions | k-Edge MST and DT are used to compute the MDCs' tours | Tour length | Python. Geosteiner library |
| [79] 2018 | Multiple node | — | RN deployment | Whale optimizer iteratively computes the locations of new relay nodes. | Number of RNs, execution time | Python |
| [6] 2017 | Multiple nodes | — | RN deployment | FSTs are formed by considering segments as hypergraphs to compute SMT. | Number of RNs, path length, hop count | Python |
| [80] 2016 | Multiple nodes | — | RN deployment | Select best triangles for node deployment according to three heuristic methods that involve DFP computation. | Relay node degree, node degree | Java |

(Continues)



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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 8** | (Continued)

| Reference | Failure type | Network state | Recovery operation | Recovery technique | Parameters optimized | Simulator |
|---|---|---|---|---|---|---|
| [81] 2016 | Multiple nodes | 1-Hop | RN deployment | Concave hull of the partitions is computed to biconnect the network partitions | Number of RNs, node degree, hop count | C |
| [1] 2016 | Multiple nodes | — | MDCs, RN deployment | MDCs and RNs are selected and unutilized RNs function as MDCs | Standard deviation, tour length | — |
| [7] 2015 | Multiple nodes | 1-Hop | RN deployment | Inner polygon is created, among the partitions to connect them using two disjoint paths. | Number of RNs, hop count, node degree | C |
| [52] 2015 | Multiple nodes | — | Relay node (RN) deployment | Deploy relay nodes toward the zero gradient position inside the convex hulls of partitions. | Number of RNs, average path length, recovery time | NS2.34 |
| [82] 2014 | Multiple nodes | 1-Hop | RN deployment | Repelling magnetic forces and Nash equilibrium are used by FBA and GTA to place relay nodes | Number of RNs, travel distance | — |
| [49] 2013 | Multiple nodes | — | MDCs, RN deployment | MDCs carry data between different segments, and RNs act as access points to reduce their tour length | Tour length | C++ |
| [83] 2012 | Multiple nodes | 1-Hop | RN | SPs are computed in different rounds on the convex hull of the partitions | Number of RNs, hop count, total searched cells | C language |
| [4] 2011 | Multiple nodes | — | RN deployment | 1C-SpiderWeb connects the boundary partitions in a ring formation, and 2C-SpiderWeb connects the partitions with ring through alternate paths | Number of RNs, path length, recovery time, average node degree | — |



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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

toward the center of the network such that all the partitions lie outside or on the computed ring. The partitions outside the ring are connected using two disjoint paths. The results illustrate that CRAFT performs better than the 2C-SpiderWeb [4] and 2C-MSS [85]. Kumar et al. [46] engaged mobile collectors for data ferrying and proposed OACE. OACE employs DT to distribute the area and compute meeting points (MPs) to deploy mobile nodes. Afterwards, the network graph consisting of obstacles is constructed to determine the obstacle-free trajectory of mobile nodes.

Lee et al. [81] aimed to create a Hamiltonian cycle of network partitions to guarantee two-vertex disjoint connections among the partitions and increase efficiency in average node degree. The "optimized biconnected" (OBiC) algorithm builds a concave hull or cave of the network partitions toward the center of the network. Afterwards, OBiC biconnects the network segments by populating RNs along with the computed cave, which is eventually similar to the Hamiltonian cycle. OBiC performs better than CRAFT [7], 2C-SpiderWeb [4], and 2C-MSS [85] in the simulation results. Chen et al. [78] proposed a connectivity restoration scheme to connect fragmented network partitions using RNs. This scheme also aims to avoid obstacles while placing RNs. If the computed position of the RN contains an obstacle, then it is placed in a nearby location calculated according to the border of the obstacle. The RNs' positions are computed by selecting quadrilaterals of the network partitions and calculating their SPs. After that, these SPs are connected using the minimum spanning tree (MST) or triangle Steiner tree. Simulation results illustrate that the reconnected network topology provides high coverage and node degree with fewer extra RNs. Verma and Ranga [79] aimed to repair the partitioned network using a nature-inspired whale optimizer scheme. The RNs are initially placed in the disjoint partitions. Whales find optimal locations while attacking their prey. This nature of whales is mimicked to find the new RNs' locations for connecting the partitions.

## 3 | Unmanned Aerial System (UAS) Assistance Paradigm

The developments in electronics and sensor devices have expanded the employment of UAVs in different applications. Hence, the UAS assistance paradigm is becoming a prominent research area to aid various applications, terrestrial networks, communication entities, and users. UAVs embedded with low price and efficient transceivers, such as IEEE 802.11 and cellular networks, can be engaged in enhancing connectivity in wireless networks. The easy and swift deployment of UAVs can help in establishing low-cost UAV–based networks to provide wireless connectivity. UAVs can collect, disseminate, and process data in wireless ground and underwater networks [14, 31, 86].

UAS comprises three main components: ground station, UAV, and communication links. The GCS controls the system's operations and facilitates communication between UAVs and human operators. UAVs perform assigned operations in their deployment locations and communicate with ground networks using particular communication links. The flight controller signifies the central processing unit of UAV. The UAVs are equipped with communication interface to accept commands from GCS and exchange data. Figure 4 shows important UAV components, and a brief discussion about the components of UAVs is given below:

- Airframe: The UAV's airframe carries different components of the UAV. Stability, lightweight and space characteristics differentiate it.

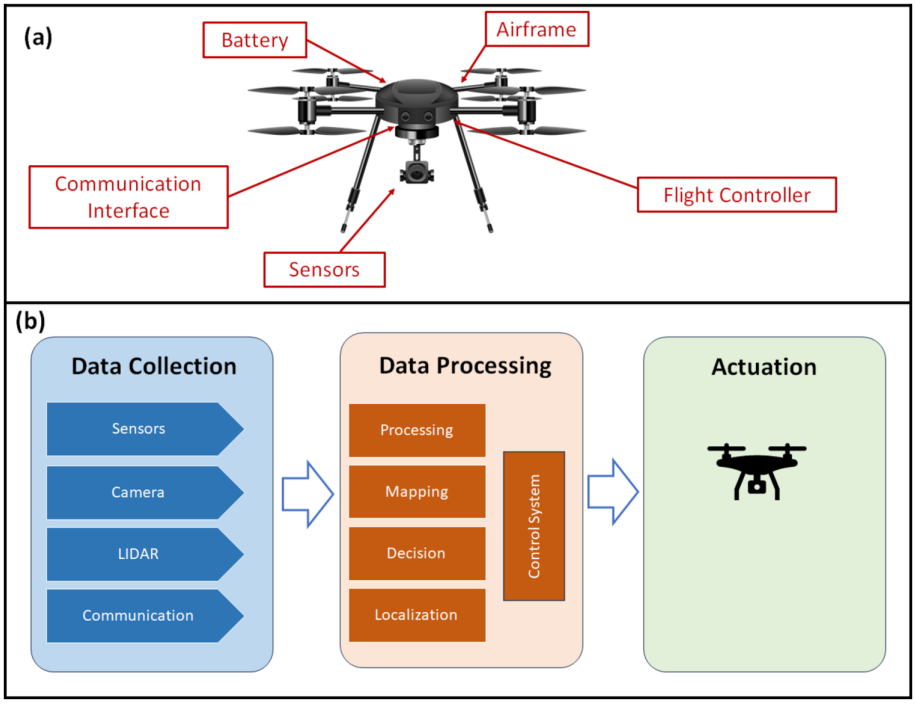



**FIGURE 4** | UAV components.



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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 9** | UAV-assisted wireless networks.

| Reference/ year | Application | Target attribute | Advantages | Limitations |
|---|---|---|---|---|
| [89] 2024 | Maximize energy efficiency in UAV networks | Energy harvesting, energy efficiency | Caters energy-effective resource allocation and connectivity problem | More energy-efficient systems such as integrated sensing UAV networks with nonlinear energy harvesting can be included |
| [90] 2024 | Optimize allocation of network resources to improve QoS | Packet delivery ratio, fairness, throughput, time delay, energy utilization, computation time | Manages the dynamic and complex ultradense UAV–assisted network conditions efficiently | Optimize resource allocation without increasing the computational complexity |
| [91] 2024 | UAV-assisted vehicular network to locate trapped vehicle | Search time, root mean square error, convergence rate | Ensures less search time and number of UAVs | Difficulty in approximating actual SAR environment with unpredictable and complex constraints |
| [92] 2024 | Ensures fairness service among multiple users | Fairness, throughput, energy consumption | The reward function to encourage fairness and energy efficiency is consistently defined | The complexity of the algorithm and latency in message transmission is not mentioned. Deep reinforcement learning method that takes dynamic observations to act may increase the complexity of algorithm |
| [93] 2023 | UAVs' deployment in wireless networks using DNN model | Coverage probability and average throughput | Proposed mechanisms can be applied to any network | Performance depends on accuracy of responses especially noisy responses. Responses to train DNN increase as the network scales up. Difficult to obtain real world implementation |
| [89] 2023 | Maximizing energy efficiency in UAV networks | Harvested power, energy efficiency | NOMA technology is exploited to improve access of IoT users, and UAVs use harvested energy to transmit uplink data to base station. | MEC technology can be integrated with nonlinear EH–based UAV networks. The work can be extended to more energy-efficient scenarios |
| [94] 2023 | UAV trajectory planning or data collection | Packet delivery rate, number of sensor nodes covered | Fuzzy path is constructed based on software defined sensor network. Improve UAV's energy efficiency | Algorithm can be further validated for different use cases, network scales, and operation domains. |
| [95] 2023 | UAV-enabled vehicular communication system | System average bit rate, U2V and V2V-vehicles scheduling | Provide communication services for both V2V and U2V vehicles | Need to achieve seamless mutual control and integration between the UAV and IRS |
| [96] 2023 | Path planning for UAV-assisted networks | Energy consumption, travel time | UAV's path planning to provide date at required rate | Performance should be evaluated for multiuser and multi-UAV scenarios |

(Continues)



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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 9** | (Continued)

| Reference/ year | Application | Target attribute | Advantages | Limitations |
|---|---|---|---|---|
| [97] 2023 | Resource allocation and UAVs' positioning with the goal of reducing the number of deployed UAVs | Average deployed UAVs | Minimize the UAVs required to provide real-time communication in wireless networks | Comparison with similar approaches is required |
| [98] 2023 | UAV-assisted sparse sensing in WSNs | Relative error, energy consumption | UAV collects data from sparse sensor node with less error rate to minimize energy consumption | Comparison with similar approaches is required |
| [99] 2023 | Data collection from UAV-enabled IoT networks | Transmission energy, number of clusters, computational complexity | Formation of node clusters, cluster heads, and UAV's routing paths for data aggregation in large-scale IoT networks | Need to include various constraints such as UAV battery power and completion time |

- Flight controller: This component monitors UAV's navigation and stability. The flight controller also generates the control signals to provide manual control for the users to manage different UAV states.
- Sensors: UAVs are equipped with various sensors, like humidity, temperature, gas, and pressure, to gather information about the environment. UAVs can either partially process or transmit their collected data to the GCS for processing and analysis.
- Global positioning system (GPS): GPS is an essential component necessary for UAV's autonomous flight and navigation. This module computes the location, direction, latitude, longitude, compass information, elevation, and speed of the UAVs at particular time intervals.
- Single-board computer (SBC): It processes the data collected by the radio frequency identification (RFID) system and transmits it to the GCS through communication links.
- Communication interface: The communication devices, such as an omnidirectional antenna, are equipped in UAVs to facilitate wireless communication with ground devices and other UAVs.
- Battery: The UAVs supply power to their components through the embedded battery. The UAVs have restricted battery power and require effective energy utilization techniques.

The capability of UAVs to carry transceivers for wireless communication depends on their design, size, and payload. For example, balloons can lift more than 9–10-kg LTE equipment. However, the mini UAVs are restricted to carrying very limited weights. The employment of UAVs in diverse applications is prevalent because of advantageous characteristics, such as flexibility, adaptive altitude, agility, and mobility. The different use cases of UAVs can considerably improve the potency of wireless networks according to the overall quality of service (QoS), coverage, throughput, and capacity. The following subsections discuss the application of UAVs in different wireless networks [87, 88]. Table 9 summarizes some of the UAV-assisted wireless networks.

### 3.1 | UAV Deployment as Wireless Relays

UAVs can either work as a relay between the detached ground devices or create a flying ad hoc network (FANET) as a reliable relay network to facilitate connectivity in obstacle-prone areas or damaged ground networks as shown in Figure 5. These relay networks can provide connectivity to disconnected network regions and act as backbone networks for terrestrial communication infrastructure. UAVs deployed as RNs can assemble data from ground devices and transmit it to the BS. UAV-assisted relay networks have numerous advantages compared to terrestrial relay networks. The cost-effective, easy, and swift deployment of UAVs entitles them to be appropriate for temporary and emergency circumstances. The high mobility feature of UAVs enables them to adjust their positions dynamically according to requirement, which aids in improving the network's performance [36].

De Freitas et al. [100] investigated the benefits of employing UAVs as relays to guarantee the connectivity of ground SNs. UAVs are



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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

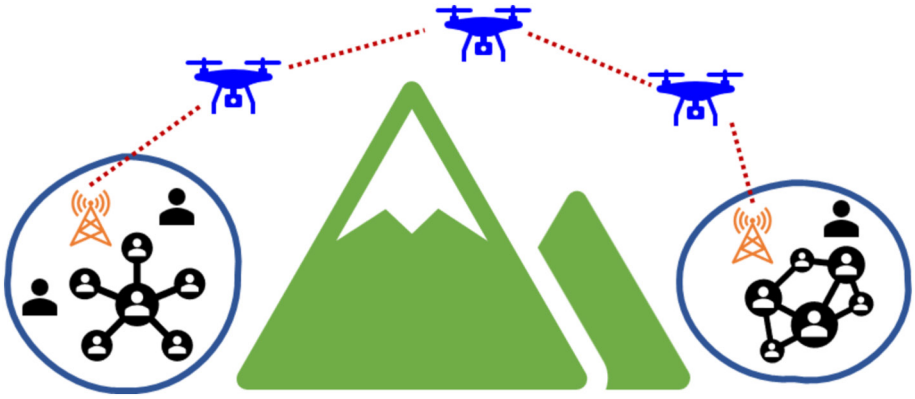


**FIGURE 5** | UAVs deployed as wireless relays.

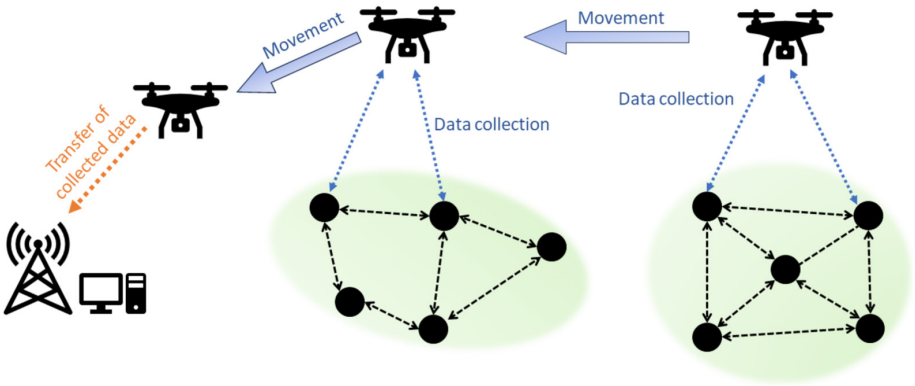



**FIGURE 6** | Data collection using UAVs.

deployed optimally to assure message delivery between the SNs. Chen et al. [101] estimated the altitude of static and mobile UAV working as a flying relay to maximize reliability. The reliability is determined using metrics such as bit error rate (BER), outage probability, and power loss. Esrafilian et al. [102] focussed on UAV placement between fixed BS and other devices according to the terrain conditions to maximize throughput. The proposed work considers initial information about the users' locations and path loss to model them in a 3D map. UAVs further collect information about users, and based on this gathered data, the optimum position of a UAV is determined iteratively. Chen et al. [103] also considered placing a UAV in a 3D area to ensure connectivity between the user and BS in case of obstacles in the terrain. The work considers real terrain formation and focuses on the overall quality of experience of the target user. Gaofeng et al. [104] implemented clustering methods for placing UAVs. The algorithm uses UAVs' count as the performance metric and deploys multiple UAVs as RNs to improve network's performance. Orfanus et al. [105] proposed an approach to deploy multiple UAVs as relays to deliver services to the ground sensors. The approach focuses on maintaining connectivity between UAVs and maximizing the UAVs' coverage for ground SNs.

UAV-based relay networks effectively improve the throughput and reliability of associated ground network. However, UAVs require attentive path planning techniques to identify the channel variations when the UAVs move. Along with trajectory, the computation of optimal heading angles for UAVs is also required to maintain connection between UAVs and GNs. The communication protocols for UAVs should ensure robust and reliable data transfer that can adapt to continuous changes in UAVs' positions. Other requirements for minimizing the energy dissipation and maximizing the throughput of UAV-based relay networks are proper flight formation of UAVs, cooperation, and coordination with other network nodes.

### 3.2 | Data Collection by UAVs in Wireless Networks

The SNs in various applications maintain connectivity using wireless standards such as Z-wave, ZigBee, and IEEE 802.15.4. These SNs have limited power, storage, and processing capability. These issues in SNs affect their functionality to achieve application-related goals. The UAVs equipped with wireless devices can support WSNs to reduce their challenges. The UAVs can increase the throughput and lifetime of WSNs by intermittently establishing connections with the SNs to collect their data and deliver it to the corresponding BS as shown in Figure 6.

Polo et al. [106] utilized UAVs to collect information from SNs deployed in agriculture fields. The UAVs also collect additional information or detect changes related to crops, such as diseases, pests, and variations in soil moisture. Ma et al. [107] investigated data collection by UAV in mobile WSNs. The mobile sensors are move in their fixed paths with diverse constant velocities, and UAV at different altitudes collects data from them. Four UAV-enabled data collection algorithms are presented by considering the duration



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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

time among the UAV and mobile sensors and data rate. Jawhar et al. [108] presented models and strategies to employ UAVs for information collection in WSNS for minimizing the energy consumption and frequency interference and prolong the network lifespan. Several UAV-assisted frameworks for data collection, such as real-time and store-and-forward data transfer, are used for different types of devices and nodes. The technique presented in [109] employs UAVs to assemble data from linear sensor networks (LSNs) monitoring linear infrastructures. SNs transmit data to the nearest RNs, and a UAV collects data from the RNs by moving back and forth to deliver collected data to BSs on both sides of the LSN. Mohamed et al. [110] investigated the use of UAVs to sustain connectivity in faulty LSNs and increase the network's reliability. UAVs move toward particular locations to work as a relay in faulty LSNs. The allocation models for UAVs are developed and compared to increase their coverage.

Employing UAVs for data collection in WSNs comes with various challenges. The challenges include identifying the optimal speed of UAVs for reliable data transfer, resolving conflicts when multiple SNs transmit data to the same UAV, and developing autonomous data collection and processing methods for UAVs. The UAV-assisted data collection methods also require optimal deployment locations for UAVs and path planning algorithms to reduce energy consumption and data transmission delay.

### 3.3 | Flying BSs

Unexpected natural disasters such as hurricanes, floods, snowstorms, and tornados can damage or totally destroy the existing ground communication networks. Consequently, an immediate communication system is required to carry out the search, public safety, and rescue operations to save lives. UAVs can dynamically change their locations according to emergencies. UAVs embedded with communication devices can provide such networking infrastructure because of their low cost, mobility, rapid configuration, and swift deployment characteristics. UAVs can also be used as temporary BSs to provide coverage in minimum time and monitor large gatherings such as sports events or political gatherings as shown in Figure 7.

Lyu et al. [111] provided an efficient approach for wireless communication using mobile base stations (MBSs). The approach focuses on providing coverage to ground terminals with minimum MBSs. The MBSs are placed starting from the boundary toward the center of the network area in a spiral path. Soorki et al. [112] proposed a cluster-based framework for machine-to-machine (M2M) connectivity. UAVs are employed as BSs to accumulate data from the cluster heads of M2M devices' clusters. The resource allocation and scheduling procedure are also proposed to minimize the CHs' power requirement. Mozaffari et al. [113] investigated deployment and trajectory planning of UAVs employed as aerial BSs. The approach for optimum mobility of the UAVs with the least energy consumption is also proposed to ensure uplink communication for IoT devices. The transmission power of the IoT network is reduced by clustering the devices and assigning a cluster to each UAV. The deployment of static or mobile UAV as aerial BS to provide connectivity in a geographical area is analyzed in [114]. The scenarios for both static and mobile UAVs are considered. The coverage probability for static UAV is computed by considering total device-to-device (D2D) users and UAV's altitude. The minimum count of stop points required to cover the entire area is computed for the scenario of a mobile UAV.

Employing UAVs as aerial BSs improves the overall network's performance such as load balancing, spectral efficiency, coverage, and user requirements. The issues that need to be addressed

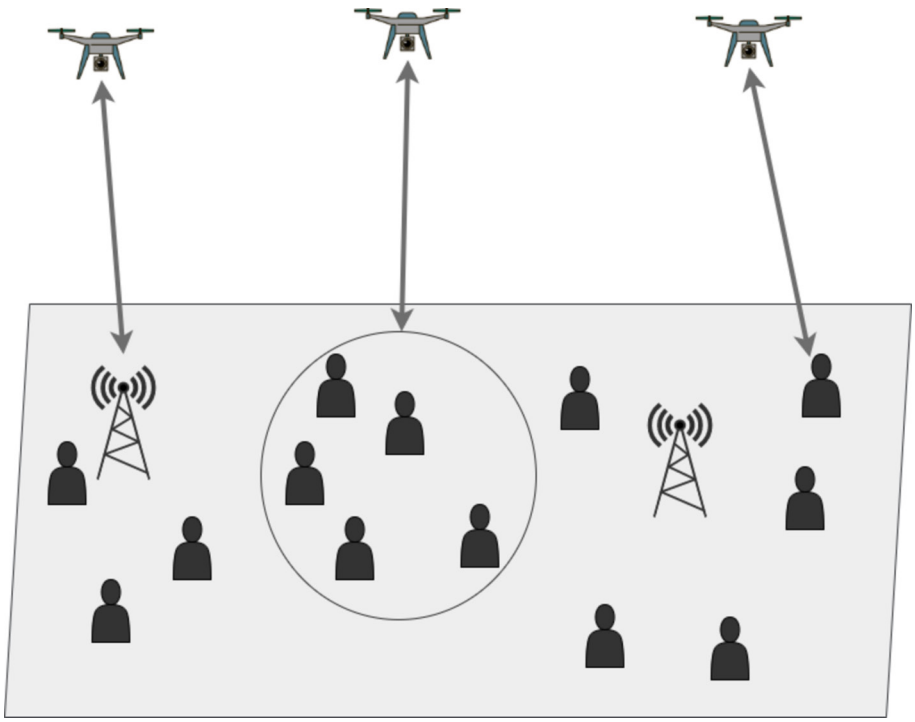


**FIGURE 7** | Flying base stations.



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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 deploying flying BSs are finding locations for UAVs and mobility in a 3D environment, interference and collision avoidance, network security, path planning, and reducing power consumption.

### 3.4 | UAV-Assisted Mobile Ad Hoc Networks (MANETs) and Vehicular Ad Hoc Networks (VANETs)

MANETs consist of backbone nodes to manage motion coordination and establish connectivity with disconnected networks. UAVs can support ground MANETs to handle various challenges such as load balancing, linkage, and reliability. UAVs can be deployed for network coordination and to ensure multihopping, scalability, and energy conservation while providing connectivity to the MANET nodes as shown in Figure 8. UAVs can also reduce the delay, routing overhead and maximize the throughput.

Lysenko et al. [115] considered the unpredictable and rapid movement of mobile users and proposed quasioptimal placement control (QPC) method to increase the throughput of MANETS. This method's application increases the network's throughput by 15%–20% compared to existing methods. Huo et al. [116] proposed an optimization algorithm to formulate the mobility of UAVs. The UAV works as RN for the GNs in the MANET. The proposed algorithm considers ground users' data requirements and focuses on data transmission rate, transmission delay, and network connectivity. Gao et al. [117] explored the application of UAVs in damaged MANET communication infrastructure and proposed a priority-based dynamic packet scheduling method to manage the QoS requirements. The method incorporates issues related to movement nodes, unstable topology, and channel quality while assigning priority to packets. The packet transmission delay is estimated for different channels such as device-to-UAV, D2D, and UAV-to-UAV.

VANETs are specific class of MANETs that supports vehicles' communication with other vehicles and roadside units. The objectives of VANETs are to provide safety, internet services and manage traffic. The massive increase in vehicles and data exchange in urban areas affects the throughput of VANETs. The significant traffic, different mobility patterns of vehicles, and obstructions in urban areas cause difficulty in finding the shortest routes. UAVs can assist VANETs as relays or FANETs to notify ground users about the current traffic conditions as shown in Figure 8. They can report accidents and also help in determining the shortest routes for ground vehicles [118].

Oubbati et al. [119] presented a UAV-assisted routing protocol for VANETs. The UAVs accumulate information and notify the ground users about the connectivity status among the vehicles and traffic density. UAVs exploit this information to find their placement locations when the ground vehicles are disconnected. Seliem et al. [120] presented routing protocol to strengthen VANET communication and focused on reducing the packet delivery delay between vehicles and UAVs. The closed-form representation is used to find the highest distance among two drones and estimate the probability distribution of packet delays between vehicles and drones. Azzoug and Boukra [121] reported the application of UAVs in vehicular delay-tolerant networking for sparse areas and proposed a bioinspired routing method to improve QoS in VANETs. The proposed approach selects the best store-carry-forward nodes for both uplink and downlink connections between UAVs and vehicles.

The main challenge for employing UAVs in MANETs is determining mobility models that support continuous communication and information flow. The mobility model and deployment positions for UAVs should consider the dynamic network topology that permits network scalability and reduces packet delay. The efficient routing protocols must be devised to accommodate UAVs with MANET backbone and help the mobile GNs to access the UAVs. UAVs will be critical enablers in intelligent transport systems (ITS), and combining UAVs with VANETs involves several challenges. These challenges include designing data-gathering algorithms, mobility strategies, path planning techniques, routing protocols supported by VANETs and UAVs, deployment algorithms for UAVs, and mobility control methods.

### 3.5 | Cache-Enabled UAVs

Caching in static BSs helps in reducing the transmission delay and throughput. This technique requires data to be stored in

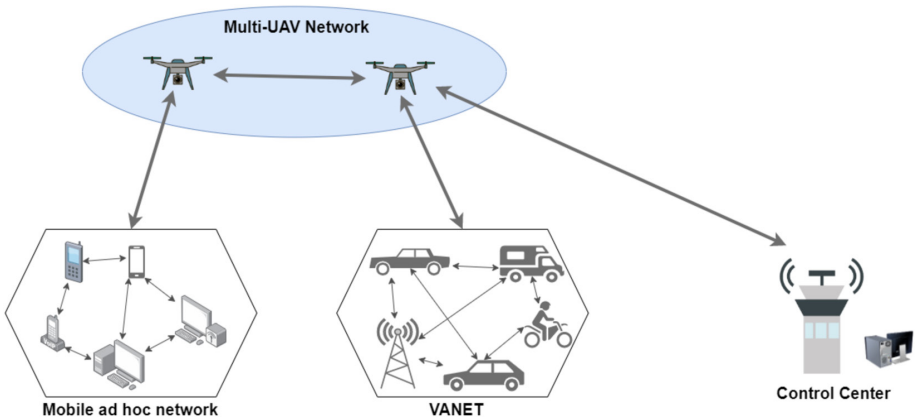



**FIGURE 8** | UAV-assisted MANET and VANET.



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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

multiple static (BSs) to cooperate with the mobile users' requirements, which is ineffective when the mobile users and their data grow. The dynamic movements of mobile nodes and frequent handovers reduce the effectiveness of caching in static BSs. Hence, in such cases, flexible and mobile BSs are required to fulfill the needs of moving users. The UAVs can be flexibly and swiftly deployed as BSs to cache contents and serve mobile users by tracking them. Caching in UAVS also helps traffic offloading in the ground networks and reduces the catching complexity compared to static BSs.

Xu et al. [122] proposed proactive caching method for UAV-assisted communication in wireless networks. The authors focused on content-centric systems, and UAVs serve some GNs requesting the same content at different times. UAV sends files to the GNs in file caching and retrieval phase. GNs with file requests extract the files from either their cache or the neighbor's cache. Chen et al. [123] examined cache-enabled UAVs' placement to enhance the QoE of wireless devices in the network. The contemplated model considers user information, such as mobility patterns, visited locations, information requests, and type of device. Based on this information, the locations for UAVs and contents to be cached at UAVs are predicted. The location prediction problem is devised as an optimization problem and solved using the machine learning framework. Fang et al. [124] presented a game-theoretic method to arrange the cache contents for UAVs and compute the UAVs' deployment locations for serving wireless networks. The authors focused on balancing the communication overhead and service probability for users.

Employing cache-enabled aerial vehicles minimizes congestion and energy consumption and enhances QoE of users. Realizing caching using UAVs requires progressive frameworks to store the popular contents according to the user's demand and mobility patterns, which can be immediately delivered according to the user's request. These frameworks should guarantee the security of the content and UAVs, avoid interference, and restrict eavesdropping. The optimization methods are also required to determine the UAVs' positions, plan trajectories for UAVs, and reduce energy consumption.

### 3.6 | UAV-Assisted 5G Networks

The enormous data transfer and video streaming increase has resulted in immense data traffic. With the increased application of IoT devices and innumerable devices being connected to the internet, the demand for high-speed data will grow exponentially. The 5G network is expected to fulfill the data demands of increasing network devices (maximum IoT devices) by providing connectivity, high data rates, lower latency, higher capacity, and better QoE with less cost and energy requirements. Communication links with low latency are required in various critical applications such as healthcare, ITS, financial, and industrial internet. 5G consists of heterogeneous networks, and UAV-backed heterogeneous networks are a remarkably efficient connectivity solution to handle emergency situations like infrastructure damage due to disasters, equipment failures, and systems flooded with sudden high user requests. UAVs can fly in hard-to-reach areas to provide connectivity. UAVs can help in

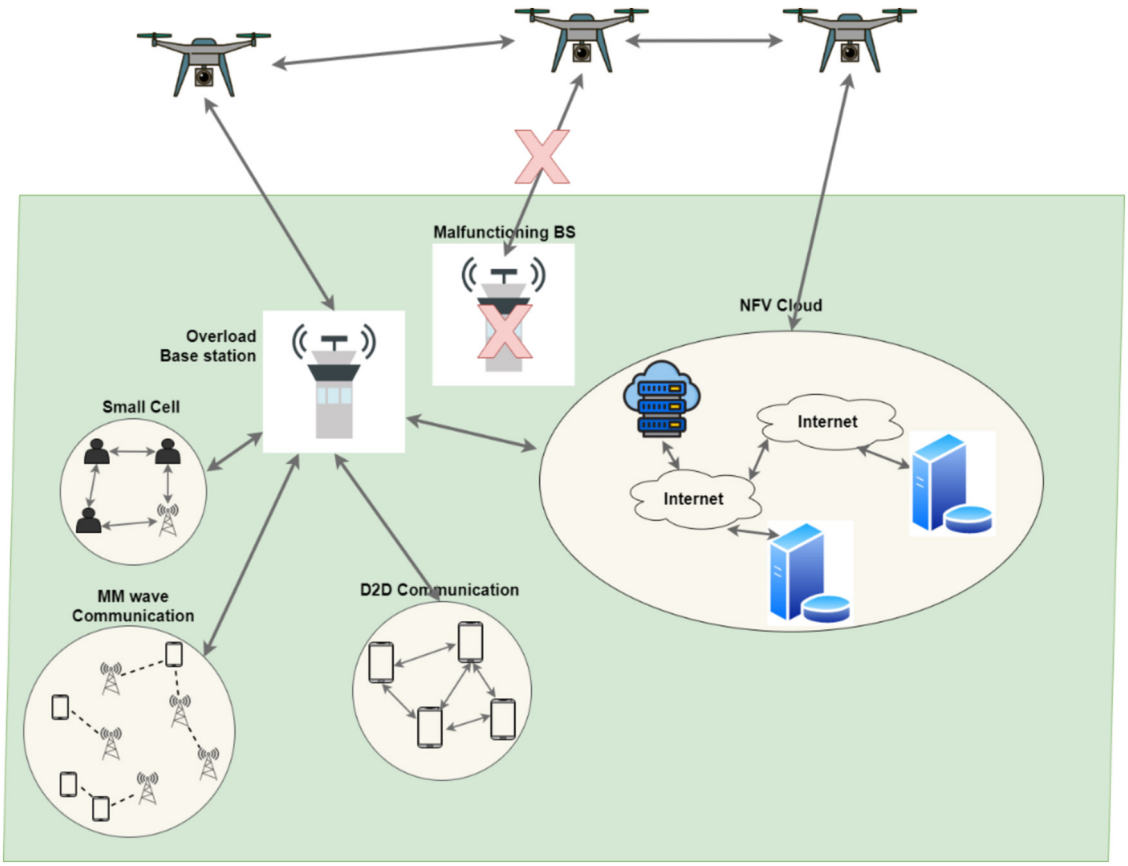



**FIGURE 9** | UAV-assisted 5G networks.



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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

applying the millimeter wave (mmWave) band in cellular networks, which can fulfill the requirements of the 5G network. UAVs can also help to facilitate D2D communication as shown in Figure 9.

Wang et al. [125] investigated resource allocation issues in UAV-assisted network. UAV works as a radio frequency energy source for (D2D) pairs using the same spectrum for data transmission without interference. MILP is used to frame the resource allocation problem. Zhang et al. [126] presented a solution for trajectory design and subcarrier allocation problems to increase the uplink sum rate. The communication between a cellular network and UAVs is considered using different modes. UAVs provide collected data to BS, and the proposed algorithm considers the delay in accumulating sensing data. Sharma et al. [127] addressed the issue of deploying more new stations to solve coverage problem in 5G networks. UAVs are incorporated into the existing networking system to solve this issue and placed in the area of demand to increase capacity and coverage. The overall UAV placement problem is devised as a joint UAV allocation, decision problem.

Incorporating UAVs in dense and heterogeneous networks is challenging in terms of placement, energy, and security issues. The UAV-enabled networks' security is extremely important, and various security issues such as eavesdropping, interference, and data security must be handled with priority. Advanced routing protocols are also required to enhance network efficacy.

### 3.7 | UAV-Enabled IoT Communications

IoT is an important aspect of 5G networks that combines various devices, sensors, and actuators through technologies such as embedded and cloud computing to form a network of heterogeneous devices. The IoT devices require reliable communication, low latency, and guaranteed data delivery. The key issues with IoT devices are definite battery power, incapable of transmitting at longer distances, and deployment in poor wireless infrastructure areas. As a promising technology, UAVs can solve these issues by enhancing connectivity through the LoS connections, enabling mobile edge computing (MEC), and their dynamic change in positions can provide connectivity to disconnected IoT devices as shown in Figure 10.

Motlagh et al. [128] employed UAVs embedded with IoT devices for providing services to ground devices from high altitudes. The UAVs are employed for crowd surveillance using a face recognition system. The case study compares the processing of video data offloading at a MEC node as compared to video data present at a UAV node. Zhan et al. [129] employed UAVs equipped with edge servers to enable computation support to the IoT devices in the UAV-enabled MEC framework. The proposed approach considers various factors such as channel interference, optimized trajectory, computation offloading, and resource distribution for expanding the services for IoT devices. Li et al. [130] presented an energy-harvesting and content-catching approach for UAV communications in IoT networks. The UAV works as RN to fulfill on-demand requests from users. The throughput in terms of user service is optimized by combining time scheduling, power allocation, and energy-harvesting schemes.

Maximum IoT devices consist of limited battery power, and employing UAVs to provide communication and energy-related services to IoT devices is a challenging problem. In addition, security procedures are required to avoid any damage to UAVs and IoT devices.

## 4 | UAV-Assisted Network Partition Detection and Recovery

Involving UAVs to solve network partitioning-related problems is a new research area for future researchers with a broad scope for further improvements and optimizations. Table 10 illustrates the review of reported UAVs' assisted network recovery techniques.

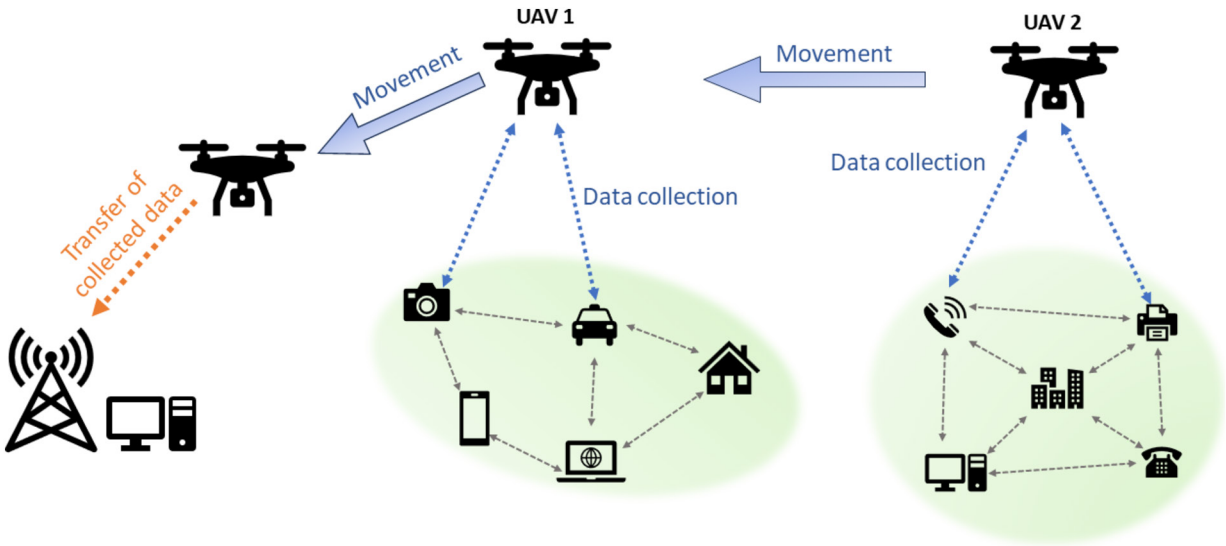



**FIGURE 10** | UAV-enabled IoT communications.



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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 10** | Analysis of network recovery methods using UAV/UAVs.

| Reference | Failure type | Overview | Parameters optimized | Limitations |
|---|---|---|---|---|
| [21] 2024 | Single node | UAVs move to the location of failed node to act as relay between partitions | Detection time | Increased messages transmitted, recovery time, and UAVs' travel distance |
| [37] 2024 | Multinode | UAVs deploy nodes at allocated areas | End-to-end delay, convergence time, number of relay nodes | Partition detection scheme is not mentioned |
| [11] 2023 | Multinode | UAVs collect the information about connected component sets in network and deploy relay nodes for recovery | Detection time, UAVs' travel distance, number of relays | Increases total messages transmitted count |
| [9] 2022 | Multinode | UAVs aid the partition detection process and recover network by deploying new relay nodes | Detection time, total messages transmitted, UAVs' travel distance, recovery time | UAVs' travel distance and detection time are large. |
| [11] 2022 | Multinode | UAVs help in detecting network partitioning and deploy relay nodes to create fault-tolerant topology | Detection and recovery time, total messages transmitted and relay nodes deployed, UAVs' travel distance | Increased message overhead and total relay nodes |
| [10] 2021 | Single node | UAV is deployed to restore network partitioned due to single-node failure. | Detection time, total messages transmitted, UAVs' travel distance, recovery time | Multiple node failures are not considered |
| [131] 2018 | — | UAVs are deployed as relay nodes to restore connectivity between isolated partitions | Network traverse time, coverage, travel distance | Network partitioning is not identified |
| [132] 2014 | — | UAVs work as data mules to carry messages among the isolated partitions | Number of packets delivered and latency time | Network portioning is not identified, increased latency time |
| [100] 2010 | — | UAV network is deployed as backbone to maintain connectivity in disjoint partitions | Average distance traveled by UAVs, connection rate, total disconnected nodes, average neighbors' count | Backbone connection is only maintained with the base station |



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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

Heimfarth and De Araujo [132] presented a method to employ UAVs for message ferrying to connect disjoint network segments. The method is suitable for delay-tolerant networks. The disjoint network partitions elect their cluster heads, responsible for coordinating with UAVs. The cluster heads store the packets received from the nodes in their partition in the local buffer and share them when the UAV arrives. The performance results of the algorithm show slight packet loss and significant latency. In [100], the idea of creating a controlled UAV network as the backbone to connect the isolated network fragments is presented. The UAV network is basically designed to connect the GNs with the BS. The paper also focusses on maintaining the network connectivity among the UAVs in their sub-network. The simulation results are analyzed according to connection rate, disconnected nodes' count, the average neighbors' count, and the average distance between the network nodes. Similarly, Park et al. [131] employed UAVs to connect the isolated network fragments using aerial links. The path planning algorithms for UAVs are established, and optimal locations for the UAVs are computed for effective data delivery. The simulation results are evaluated according to different parameters such as network traverse time, coverage time, UAVs' average travel distance, and duplicate coverage.

Our proposed work in [9–11, 16] has illustrated different ways to employ UAVs for detection and network recovery. In [10], the network fragmented because of single cut-vertex node damage is recovered using UAV. The network nodes identify the failure of their neighbor and report this to sink node. The sink node identifies if failed node is cut vertex and sends UAV to connect the partitions in case of cut-vertex node failure. The UAVs explore the entire damaged network and provide this information to their UCC. Based on the information furnished by UAVs, the network partitioning is determined along with the total partitions formed and nodes in such partitions. The UCC identifies the network partitioning with the help of the depth-first search (DFS) algorithm. Afterwards, the sink node computes the locations for the deployment of extra nodes using a network recovery algorithm. Lastly, the UAVs facilitate the recovery by deploying new nodes at the provided locations. The recovery algorithms in [9, 16] work to reduce the additional RN count, which can also include more cut vertices and make the network more susceptible to further partitioning. The recovery mechanism in [12] aims to create fault-tolerant network topology by establishing two-vertex disjoint connections between partitions. A closed curve algorithm connects them in a ring topology. These proposed algorithms are tested according to different parameters such as recovery and detection time, total messages transmitted, required RNs, and UAVs' travel distance. The algorithms are also compared with similar recovery techniques, and their performance is better in terms of recovery time, distance traveled, and required RNs.

The network partition recovery techniques mentioned previously can work better only if accurate network information state after the damage is available. The UAVs can interact with GNs strengthening and recovering the connectivity in damaged wireless networks [11, 55, 133, 134]. The different methods for utilizing UAVs for the network recovery are explained below:

- The UAVs can explore the damaged network region to accumulate complete information of failed and active nodes in network. Based on this accumulated information, network partitioning can be confirmed. The main requirements for UAV-enabled information accumulation require optimum area exploration techniques, obstacle avoidance and path planning mechanisms, and communication and coordination techniques for GNs and UAVs.
- The recovery operations can be conducted accordingly depending on the information received from UAVs about the network state. The UAVs can be deployed to relay the information between the network partitions as shown in Figure 11a. The UAVs' deployment process requires computing optimum altitude, locations, and angle for UAVs to hover, proper communication architecture, data storage capabilities based on the applications' requirements, and proper communication between air and ground network nodes.
- The UAVs can also deploy extra nodes between the partitions as shown in Figure 11b. Employing UAVs to deploy extra nodes in the network requires efficient payload management methods, computation of optimum heading angles, obstacle avoidance and path planning methods, and mechanisms to deploy or deliver nodes at the provided locations.
- Suppose the extra node deployment is not feasible in case of disaster or obstacle-prone areas, and real-time information is required. In that case, the ad hoc network using UAVs, i.e., FANETs, can be employed in such a scenario. The UAVs can create FANET connecting different network partitions in order to get the current network state and establish connectivity as shown in Figure 11c. This method requires computing 3D-placement positions for the UAVs, heading angles and UAVs' altitude, air-to-air and ground-to-air communication, and data storage facilities.
- The self-healing recovery methods can also be combined with UAV-enabled recovery solutions. Based on the information provided by the UAVs, the network partitions can be determined accurately. Most of the times, the network partitions consist of one or two nodes, and these nodes can be moved between the other network partition to establish connectivity. This process can reduce the deployment cost required to place extra RNs. However, this method requires complex hardware for the network to support their movement for connecting network partitions.

The successful utilization of UAVs in detection and recovery necessitates the development of methodologies for deploying RNs, devising optimal path planning strategies, implementing area exploration approaches tailored for UAVs, and integrating effective obstacle avoidance mechanisms. The distance traveled by UAVs for assessing the damaged network and the node deployment procedure should be optimized because of UAVs' limited onboard energy. The UAVs should work in their own application areas to avoid interference and collisions with other UAVs. In order to avoid conflicts, clustering approaches can be employed to assign nodes to the UAVs in different phases.

## 5 | Lessons Learned

Several limitations of existing network partition recovery techniques can be addressed by employing UAV-enabled solutions. Researchers have analyzed and proposed different



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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

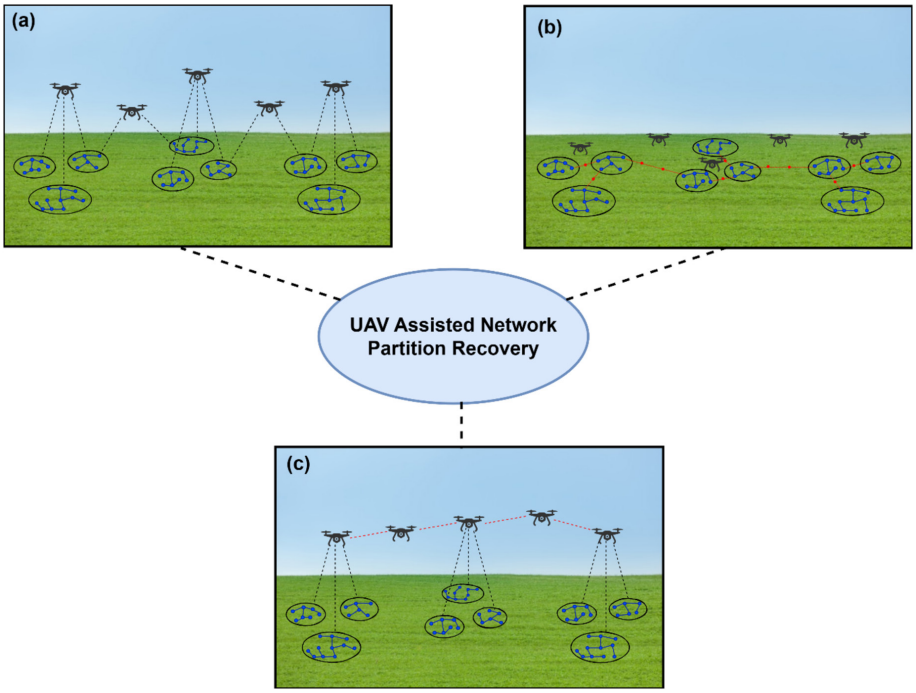



**FIGURE 11** | UAV-assisted network partition recovery.

communication technologies for use with UAVs, considering factors like range and data transfer. These assessments have led to the advancements in existing technologies and development of new techniques especially designed for UAV communications. Compared to ground-based networks with expensive infrastructure, UAV networks are significantly cheaper to deploy for temporary network setups. Unlike terrestrial networks that may face signal interferences, UAVs can often establish clear LoS connections where reliable data connections are required, leading to stronger and more reliable communication links. This is particularly beneficial in situations where reliable data transmission is required. They can move to different locations, making them ideal for remote areas for providing temporary network coverage. UAV networks can adjust on-demand network changes based on unforeseen circumstances. This is notably valuable in scenarios with unpredictable traffic patterns or abrupt changes in network requirements. Various methods to utilize UAVs are being explored to further enhance the deployment and utilization of UAVs with different terrestrial wireless networks. These noteworthy characteristics of UAV and UAV networks can be exploited to mitigate the limitations of existing network recovery technologies. UAVs can assist in identifying network partitions and aid the recovery process in different ways.

However, exploring different methods to utilize UAVs also requires addressing the unique requirements and challenges of deploying UAVs and the specific communication technologies involved for creating adaptable, reliable, and efficient UAV networks.

## 5.1 | Implementation Challenges

UAVs are evolving to become the critical element of future wireless network technologies. Several advantageous characteristics of UAVs promote research and growth in this domain. Table 11 shows some challenges that must be mitigated to successfully implement the UAV-based network partition recovery system, which are explained below:

- The optimal deployment of UAVs in a 3D environment is a critical concern that needs specific consideration while deploying UAVs to maximize capacity and coverage in IoT networks, smart cities, caching, and public safety. The UAVs' placement is a crucial parameter as it can affect the system's QoS, power, and coverage.
- The data collection using UAVs comprises important technologies such as routing, path planning, access to numerous SNs scattered in the deployment region and clustering of these nodes, and resource management. Various data-gathering techniques using UAVs attempt to reduce flight time and UAVs' travel distance, maximize the overall collected data, and reduce the energy utilization of UAVs or sensors. The effective optimization and utilization of three essential technologies, such as clustering methods for the SNs, UAVs' path planning, and data collection modes of UAVs, can ensure the timeliness of gathered data and reduce the energy usage of SNs and UAVs.
- The UAVs' localization requires various other capabilities, such as perception, motion control and planning, along with mapping and obstacle avoidance. These processes are supported by various algorithms, sensors, and actuators, which are also considered fault sources. In autonomous systems, these faults should be observed via sensors or integrity monitoring techniques.
- The channel modeling for UAV-based networks is essential mainly when UAVs are employed for the applications like coverage improvement, IoT communications, and cellular



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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 11** | Implementation challenges for UAVs' assisted network recovery.

| Reference | Implementation challenges | Techniques | Parameters |
|---|---|---|---|
| [31, 135–138] | Resource management | Management of spectrum, trajectory, path planning, data offloading, cost, frequency and energy | Per node power, latency, cost, interference, UAVs' count, flight time, signal-to-noise ratio, path loss, outage probability, security, network density, utility, revenue, traffic control. |
| [139–144] | Path planning | Grid-based, sampling-based, AI-based, machine learning-based, mathematical models, bioinspired, multiobjective optimization, graph search methods | Completeness, path length, cost-efficient, optimality, collision avoidance, time efficiency, robustness, energy efficient. |
| [145–147] | Mobility | Random models, learning-based models, geographical dependency models, temporal dependency, spatial dependency and hybrid models | Throughput, overhead, packet delivery ratio end-to-end delay. |
| [14, 148–151] | Payload management | Load attached rigidly, load suspended using cables. | UAVs' sustainable payload, delivery distance, flight time, UAVs' speed, path planning, waypoint navigation, safety monitoring. |
| [86, 152–154] | Data collection | Flying mode, hovering mode, hybrid mode | Flight time, energy consumption of UAVs and sensors, age of the accumulated information, computing and collection efficiency, transmit power of sensors, UAVs' trajectory |
| [139, 155–159] | Routing | Proactive, reactive, hybrid, topology-based, position-based, learning-based, bioinspired. | Connectivity, delay, throughput, successful route setup, data delivery rate, energy consumption, end-to-end delay, average hop count |
| [160–163] | Optimal UAVs' placement | Horizontal and vertical placement in 2D and 3D | Quality of experience (QoE), lifetime, network throughput, spectrum efficiency, and fairness UAVs' count and outage probability |
| [164–168] | Localization | GPS, GNSS, VLC, IMU, LiDAR, SLAM | Accuracy, robustness, error dynamics, reliability, representational richness, computational speed, update rate and power consumption. |
| [102, 169, 170] | Channel modeling | Empirical, deterministic, stochastic, geometry-based stochastic | Antenna design, propagation environment, UAVs' size, payload capacity and flight altitude, position, shape and size of obstacles, accuracy, computational complexity and fading statistics estimation. |
| [12, 15, 171, 172] | Spectrum management | Spectrum sharing (interweave, overlay, and underlay using cooperative, machine, and deep learning), access control, auctioning, spectrum monitoring (cloud-based, fog-based, NFV-based, SDN-based, UAV-based, crowdsourcing). | Spectrum utilization, efficiency of network operators based on channel coverage and capacity, reduced operational cost, less infrastructure, better revenue, traffic offloading using relay nodes, security, interference control |
| [13, 173–175] | Security | Detect jamming by controlling radio signals, detecting malicious UAVs, encrypting UAVs' data, embedding watermarks with UAVs' data, dropping repeated data from same source, using unique timestamps during the authentication processes, employing blockchain-enabled gateways | Secrecy rate, secrecy outage probability, secrecy capacity, ergodic secrecy capacity, intercept probability, secrecy energy efficiency, achievable secrecy diversity order |



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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

networks. The channel characteristics of flying networks rely on elevation angle, altitude, location, and type of UAV. The UAV-based communication channels possess unique features as compared to commonly utilized satellite, cellular, and Wi-Fi systems. The cochannel hindrance is a critical concern in the frequently used channels (WI-Fi, satellite, cellular). Therefore, the improved and cost-effective UAV-based communication requires the accurate investigation of the UAVs' channel characteristics.

- The UAVs carry the load by attaching it rigidly or using cables to suspend the load. The rigidly connected load can result in agile disruption, sluggish for fast response, and increased inertia. The cable-load suspended system increases the rate of under-actuation, which increases the difficulty in control and planning. Due to multiple applications and challenges, load transportation through UAVs is a growing research domain. Various parameters that need to be monitored for designing an efficient UAV-based payload system for deliveries at multiple target locations are UAVs' sustainable payload, delivery distance, flight time, UAVs' speed, path planning, waypoint navigation, and safety monitoring.
- UAVs' mobility is an important factor that affects communication. The large mobility of UAV nodes in a network can cause link outages and frequent topology changes. Hence, different approaches are required to deal with the connectivity issues in UAVs' assisted networks. The primary approach is topology control, which optimizes the control and networking effects in order to lessen the connectivity challenges.
- The reliable and efficient communication among the UAVs and GNs as well as among the UAVs is necessary and challenging due to the unstable links, power constraints, frequent addition and removal of UAVs due to mobility, load balancing, and UAVs' localization. Thus, the routing protocols devised for the UAVs should adapt to intermittent links, UAVs' mobility, topology changes, and security requirements to transmit critical messages.
- The flight path planning of UAVs is challenging because of various parameters and constraints such as energy consumption, UAVs' dynamics, flight constraints, channel variations due to dynamic movements, and finding numerous UAVs' locations. Moreover, determining the optimal trajectory for UAVs is difficult because significant attention is required to plan the movements of UAVs along with maintaining QoS standards in wireless communication. The appropriate path planning is required for the UAVs for the successful accomplishment of their mission.
- The UAVs' employment in diverse applications has raised cyberattacks substantially. Mainly, the LoS radio links containing valued information like network and routing details, navigation information, and CNPC signals, are more vulnerable to intrusions. These intrusions affect the QoS, secrecy, and reliability of UAV-associated communication.

### 5.2 | Takeaways From the Presented Survey

While UAVs offer significant benefits, researchers are further working to enhance their reliability in less cost. This involves developing communication protocols to withstand disruptions and cyberattacks, implementing backup systems, and exploring energy-efficient technologies to extend flight times. Additionally, researchers are striving to standardize communication technologies across applications and manufacturers, and encouraging the use of open-source solutions to drive down costs associated with proprietary technologies. These ongoing efforts aim to address limitations and maximize the potential of UAV communication technologies in various sectors. While UAVs offer significant benefits, researchers are further working to enhance their reliability in less cost. This involves developing communication protocols to withstand disruptions and cyberattacks, implementing backup systems, and exploring energy-efficient technologies to extend flight times. Additional aims to improve UAV-based communication include standardizing communication technologies across applications and encouraging the utilization of open-source solutions to reduce the associated costs to maximize the potential of UAV-enabled communication technologies.

Future research in UAV-assisted network recovery methods should focus on maximizing practical deployment, and scalability. This includes exploring wireless energy–harvesting techniques to address power limitations and establish commercially applicable network setups designed for UAVs. Researchers should examine efficient network backhauling techniques for the fast UAVs' mobility and deployment of UAV networks. In order to develop efficient UAV–assisted partition recovery solutions, the channel interference and errors in packet transmissions should be considered. The UAVs also require efficient data collection and the caching strategies while collecting data about the active and the failed nodes in the network. Moreover, employing UAVs to set up the SNs in the computed locations require payload management techniques and strategies to deliver the nodes. In order to recover the network by establishing FANET, the optimal deployment positions for the UAVs need to be computed. UAV-based communication systems are simulated in various languages and simulation tools like python, OMNeT++, and Matlab. However, various challenges associated with these simulators include environmental factors, security, infrastructure costs, and communication issues. Therefore, independent UAV simulation tools should be developed to consider hardware features and to provide realistic simulation environment.

## 6 | Conclusion

This paper concentrates on integrating UAVs for network partitioning detection and recovery to reduce the limitation of existing recovery solutions. The network partitioning problem and the partition recovery techniques are described in the beginning. The existing network recovery techniques are systematically reviewed, considering their advantages and limitations. Furthermore, the features of UAS and the motivation to employ UAVs to solve the network partitioning problem are highlighted. The paper is mainly structured based on three aspects. Firstly, the algorithms based on existing recovery techniques are surveyed separately. Afterwards, the accelerating growth of UAVs for various networking applications and UAS architecture has been discussed. The applications of UAVs like wireless relays, data



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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

collection, flying BSs, and UAV-assisted IoT, MANET, VANET, and 5G networks have been reviewed. Lastly, the paper further directs that the increasing use of UAVs can also help in dealing with network partitioning problems. Some research questions are raised in Section 1 of the paper, and the answers to those questions, which are discussed throughout the paper, are given below:

- RQ1: Can UAVs be employed to deal with the drawbacks of the existing network recovery techniques?
  - *RA1:* In the case of network partitioning, quick sensing systems are required to get the complete state of the affected region. Therefore, according to the applications' requirements, the UAVs can collect data even from the obstacle-prone areas using various sensors and provide their gathered information to the central BS. Afterward, different connectivity restoration algorithms can be implemented based on the received information to form robust postfailure network topology.
- RQ2: How can UAVs be used to perform the partition detection process?
  - *RA2:* One mechanism to deal with this problem is assessing the shape and size of damaged area, and determining the actual failed nodes intending to compute new deployment locations for the RNs. UAVs can get the information about the network state such as active and failed nodes. Based on this information, the network partitions can be identified.
- RQ3: How can UAVs aid the connectivity restoration process after network partitioning?
  - *RA3:* UAVs can act as relays to connect partitions, UAVs can deploy extra nodes between the partitions, and UAVs can create a backbone network (FANET) to connect partitions.
- RQ4: What are the implementation challenges and requirements for the UAVs' employment in detection and recovery processes?
  - *RA4:* The challenges that must be mitigated to successfully implement the UAV-based network partition recovery system are resource management, path planning, mobility, payload management, data collection, routing, optimal UAVs' placement, localization, channel modeling, spectrum management, and security.

In future, we will focus our research on analyzing and implementing the optimization methods required to determine the UAVs' positions, plan trajectories and obstacle avoidance mechanisms for UAVs, reduce energy consumption, and coordinate between the ground and UAV nodes.


### Acknowledgments

We are grateful to the Science and Engineering Research Board (SERB)—Department of Science and Technology (DST), Government of India, for sponsoring this research.


### Data Availability Statement

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

### References


1. M. Younis, I. F. Senturk, K. Akkaya, S. Lee, and F. Senel, "Topology Management Techniques for Tolerating Node Failures in Wireless Sensor Networks: A Survey," *Computer Networks* 58 (2014): 254–283, https://doi.org/10.1016/j.comnet.2013.08.021.

2. S. Verma, S. Bhatia, S. Zeadally, and S. Kaur, "Fuzzy-Based Techniques for Clustering in Wireless Sensor Networks (WSNs): Recent Advances, Challenges, and Future Directions," *International Journal of Communication Systems* 36 (2023): e5583, https://doi.org/10.1002/dac.5583.

3. F. Senel, M. F. Younis, and K. Akkaya, "Bio-Inspired Relay Node Placement Heuristics for Repairing Damaged Wireless Sensor Networks," *IEEE Transactions on Vehicular Technology* 60 (2011): 1835–1848.

4. I. Senturk, S. Yilmaz, and K. Akkaya, "Connectivity Restoration in Delay-Tolerant Sensor Networks Using Game Theory," *International Journal of Ad Hoc and Ubiquitous Computing* 11 (2012): 109–124, https://doi.org/10.1504/IJAHUC.2012.050268.

5. W. Lalouani, M. Younis, and N. Badache, "Optimized Repair of a Partitioned Network Topology," *Computer Networks* 128 (2017): 63–77, https://doi.org/10.1016/j.comnet.2017.02.003.

6. S. Lee, M. Younis, and M. Lee, "Connectivity Restoration in a Partitioned Wireless Sensor Network With Assured Fault Tolerance," *Ad Hoc Networks* 24 (2015): 1–19, https://doi.org/10.1016/j.adhoc.2014.07.012.

7. S. Chouikhi, I. El Korbi, Y. Ghamri-Doudane, and L. Azouz Saidane, "Centralized Connectivity Restoration in Multichannel Wireless Sensor Networks," *Journal of Network and Computer Applications* 83 (2017): 111–123, https://doi.org/10.1016/j.jnca.2017.01.040.

8. K. Akkaya, I. F. Senturk, and S. Vemulapalli, "Handling Large-Scale Node Failures in Mobile Sensor/Robot Networks," *Journal of Network and Computer Applications* 36 (2013): 195–210.

9. A. Zear and V. Ranga, "UAVs Assisted Network Partition Detection and Connectivity Restoration in Wireless Sensor and Actor Networks," *Ad Hoc Networks* 130 (2022): 102823, https://doi.org/10.1016/j.adhoc.2022.102823.

10. A. Zear and V. Ranga, "Distributed Partition Detection and Recovery Using UAV in Wireless Sensor and Actor Networks," *Kuwait Journal of Science* 48 (2021): 1–16, https://doi.org/10.48129/kjs.v48i4.10819.

11. A. Zear, V. Ranga, and K. Bhushan, "Coordinated Network Partition Detection and Bi-Connected Inter-Partition Topology Creation in Damaged Sensor Networks Using Multiple UAVs," *Computer Communications* 203 (2023): 15–29, https://doi.org/10.1016/j.comcom.2023.02.018.

12. H. Kang, X. Chang, J. Mišić, V. B. Mišić, J. Fan, and Y. Liu, "Cooperative UAV Resource Allocation and Task Offloading in Hierarchical Aerial Computing Systems: A MAPPO Based Approach," *IEEE Internet of Things Journal* 10 (2023): 10497–10509.

13. S. A. H. Mohsan, N. Q. H. Othman, Y. Li, M. H. Alsharif, and M. A. Khan, "Unmanned Aerial Vehicles (UAVs): Practical Aspects, Applications, Open Challenges, Security Issues, and Future Trends," *Intelligent Service Robotics* 16 (2023): 109–137.

14. B. Shahzaad and A. Bouguettaya, "Service-Oriented Architecture for Drone-Based Multi-Package Delivery," *Proceedings - IEEE International Conference on Web Services, ICWS* 2022 (2022): 103–108, https://doi.org/10.1109/ICWS55610.2022.00030.

15. N. Liao, P. He, Y. Du, Y. Zhang, Y. Chen, and T. Liang, "Joint Mission Planning and Spectrum Resources Optimization for Multi-UAV Reconnaissance," *IET Communications* 17 (2023): 324–335.

16. A. Zear, V. Ranga, and K. Bhushan, "Partition Detection and Recovery by UAVs in Damaged WSANs Using N-Angle Clustering,"





10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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 Communication Systems* 35 (2022): 1–22, https://doi.org/10.1002/dac.5190.

17. V. Ranga, M. Dave, and A. K. Verma, "Network Partitioning Recovery Mechanisms in WSANs: A Survey," *Wireless Personal Communications* 72 (2013): 857–917, https://doi.org/10.1007/s11277-013-1046-7.

18. R. Pallavi and G. C. B. Prakash, "A Review on Network Partitioning in Wireless Sensor and Actor Networks," in *2015 International Conference on Applied and Theoretical Computing and Communication Technology (ICATccT)* (Davangere, India: IEEE, 2015): 771–782, https://doi.org/10.1109/ICATCCT.2015.7456987.

19. K. Akkaya, A. Thimmapuram, F. Senel, et al., "Distributed Recovery of Actor Failures in Wireless Sensor and Actor Networks," in *IEEE Wireless Communications and Networking Conference, WCNC* (Las Vegas, NV: IEEE, 2008): 2480–2485, https://doi.org/10.1109/wcnc.2008.436.

20. S. Chouikhi, I. El Korbi, Y. Ghamri-Doudane, and L. Azouz Saidane, "Recovery From Simultaneous Failures in a Large Scale Wireless Sensor Network," *Ad Hoc Networks* 67 (2017): 68–76, https://doi.org/10.1016/j.adhoc.2017.10.008.

21. A. Zear, V. Ranga, and K. K. Gola, "Network Partition Detection and Recovery With the Integration of Unmanned Aerial Vehicle," *Concurrency and Computation: Practice and Experience* 36 (2024): e8048, https://doi.org/10.1002/cpe.8048.

22. S. Shriwastav and D. Ghose, "Round-Table Negotiation for Fast Restoration of Connectivity in Partitioned Wireless Sensor Networks," *Ad Hoc Networks* 77 (2018): 11–27, https://doi.org/10.1016/j.adhoc.2018.04.008.

23. G. Rajeswari, M. K. Sandhya, and K. Murugan, "Obstacle-Aware Connectivity Restoration for the Partitioned Wireless Sensor Networks Using Mobile Data Carriers," *Wireless Networks* 29 (2023): 1703–1720, https://doi.org/10.1007/s11276-022-03221-4.

24. M. Adil, H. Song, M. A. Jan, et al., "UAV-Assisted IoT Applications, QoS Requirements and Challenges With Future Research Directions," *ACM Computing Surveys* 56 (2024): 1–35, https://doi.org/10.1145/3657287.

25. I. Chandran and K. Vipin, "Multi-UAV Networks for Disaster Monitoring: Challenges and Opportunities From a Network Perspective," *Drone Systems and Applications* 12 (2024): 1–28, https://doi.org/10.1139/dsa-2023-0079.

26. K. Messaoudi, O. S. Oubbati, A. Rachedi, A. Lakas, T. Bendouma, and N. Chaib, "A Survey of UAV-Based Data Collection: Challenges, Solutions and Future Perspectives," *Journal of Network and Computer Applications* 216 (2023): 103670, https://doi.org/10.1016/j.jnca.2023.103670.

27. O. Sami Oubbati, M. Atiquzzaman, T. Ahamed Ahanger, and A. Ibrahim, "Softwarization of UAV Networks: A Survey of Applications and Future Trends, IEEE," *Access* 8 (2020): 98073–98125, https://doi.org/10.1109/ACCESS.2020.2994494.

28. T. Bouzid, N. Chaib, M. L. Bensaad, and O. S. Oubbati, "5G Network Slicing With Unmanned Aerial Vehicles: Taxonomy, Survey, and Future Directions," *Transactions on Emerging Telecommunications Technologies* 34 (2023): 1–48, https://doi.org/10.1002/ett.4721.

29. A. Rovira-Sugranes, A. Razi, F. Afghah, and J. Chakareski, "A Review of AI-Enabled Routing Protocols for UAV Networks: Trends, Challenges, and Future Outlook," *Ad Hoc Networks* 130 (2022): 102790, https://doi.org/10.1016/j.adhoc.2022.102790.

30. M. Dai, N. Huang, Y. Wu, J. Gao, and Z. Su, "Unmanned Aerial Vehicles Assisted Wireless Networks: Advancements, Challenges, and Solutions," *IEEE Internet of Things Journal* 10, no. 5 (2022): 4117–4147, https://doi.org/10.1109/JIOT.2022.3230786.

31. S. Taimoor, L. Ferdouse, and W. Ejaz, "Holistic Resource Management in UAV-Assisted Wireless Networks: An Optimization Perspective," *Journal of Network and Computer Applications* 205 (2022): 103439, https://doi.org/10.1016/j.jnca.2022.103439.

32. G. Alsuhli, A. Fahim, and Y. Gadallah, "A Survey on the Role of UAVs in the Communication Process: A Technological Perspective," *Computer Communications* 194 (2022): 86–123, https://doi.org/10.1016/j.comcom.2022.07.021.

33. Y. Liu, H. N. Dai, Q. Wang, M. K. Shukla, and M. Imran, "Unmanned Aerial Vehicle for Internet of Everything: Opportunities and Challenges," *Computer Communications* 155 (2020): 66–83, https://doi.org/10.1016/j.comcom.2020.03.017.

34. R. Shahzadi, M. Ali, H. Z. Khan, and M. Naeem, "UAV Assisted 5G and Beyond Wireless Networks: A Survey," *Journal of Network and Computer Applications* 189 (2021): 103114, https://doi.org/10.1016/j.jnca.2021.103114.

35. S. Zaidi, M. Atiquzzaman, and C. T. Calafate, "Internet of Flying Things (IoFT): A Survey," *Computer Communications* 165 (2021): 53–74, https://doi.org/10.1016/j.comcom.2020.10.023.

36. S. Hayat, E. Yanmaz, and R. Muzaffar, "Survey on Unmanned Aerial Vehicle Networks for Civil Applications: A Communications Viewpoint," *IEEE Communications Surveys and Tutorials* 18 (2016): 2624–2661, https://doi.org/10.1109/COMST.2016.2560343.

37. Z. Ding, Y. Ma, S. Chen, and F. Dan, "UAV-Assisted Adaptive Relay Nodes Placement for Network Connectivity Recovery in WSNs," in *2024 36th Chinese Control and Decision Conference (CCDC)* (Xi'an, China: IEEE, 2024): 1507–1514, https://doi.org/10.1109/CCDC62350.2024.10588364.

38. B. Alzahrani, O. S. Oubbati, A. Barnawi, M. Atiquzzaman, and D. Alghazzawi, "UAV Assistance Paradigm: State-of-the-Art in Applications and Challenges," *Journal of Network and Computer Applications* 166 (2020): 102706, https://doi.org/10.1016/j.jnca.2020.102706.

39. U. Baroudi, M. Aldarwbi, and M. Younis, "Energy-Aware Connectivity Restoration Mechanism for Cyber-Physical Systems of Networked Sensors and Robots," *IEEE Systems Journal* 14 (2020): 3093–3104, https://doi.org/10.1109/JSYST.2020.2970649.

40. R. Kumar and T. Amgoth, "Adaptive Cluster-Based Relay-Node Placement for Disjoint Wireless Sensor Networks," *Wireless Networks* 26 (2020): 651–666, https://doi.org/10.1007/s11276-019-02171-8.

41. Y. K. Joshi and M. Younis, "Restoring Connectivity in a Resource Constrained WSN," *Journal of Network and Computer Applications* 66 (2016): 151–165, https://doi.org/10.1016/j.jnca.2016.03.009.

42. V. Ranga, M. Dave, and A. K. Verma, "A Hybrid Timer Based Single Node Failure Recovery Approach for WSANs," *Wireless Personal Communications* 77 (2014): 2155–2182, https://doi.org/10.1007/s11277-014-1631-4.

43. K. Akkaya, F. Senel, A. Thimmapuram, and S. Uludag, "Distributed Recovery From Network Partitioning in Movable Sensor/Actor Networks via Controlled Mobility," *IEEE Transactions on Computers* 59 (2010): 258–271, https://doi.org/10.1109/TC.2009.120.

44. S. Lee and M. Younis, "Recovery From Multiple Simultaneous Failures in Wireless Sensor Networks Using Minimum Steiner Tree," *Journal of Parallel and Distributed Computing* 70 (2010): 525–536.

45. Y. K. Joshi and M. Younis, "Exploiting Skeletonization to Restore Connectivity in a Wireless Sensor Network," *Computer Communications* 75 (2016): 97–107.

46. R. Kumar, T. Amgoth, and D. Das, "Obstacle-Aware Connectivity Establishment in Wireless Sensor Networks," *IEEE Sensors Journal* 21 (2021): 5543–5552, https://doi.org/10.1109/JSEN.2020.3032144.

47. W. Lalouani, M. Younis, and N. Badache, "Interconnecting Isolated Network Segments Through Intermittent Links," *Journal of Network and Computer Applications* 108 (2018): 53–63, https://doi.org/10.1016/j.jnca.2018.02.010.



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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

48. A. A. Abbasi, K. Akkaya, and M. Younis, "A Distributed Connectivity Restoration Algorithm in Wireless Sensor and Actor Networks," in *32nd IEEE Conference on Local Computer Networks (LCN 2007)* (Dublin, Ireland: IEEE, 2007): 496–503. 10.1109/LCN.2007.19.

49. A. Abbas and M. Younis, "Establishing Connectivity Among Disjoint Terminals Using a Mix of Stationary and Mobile Relays," *Computer Communications* 36 (2013): 1411–1421, https://doi.org/10.1016/j.comcom.2013.06.004.

50. F. Senel and M. Younis, "Optimized Relay Node Placement for Establishing Connectivity in Sensor Networks," in *2012 IEEE Global Communications Conference (GLOBECOM)* (Anaheim, CA: IEEE, 2012): 512–517, https://doi.org/10.1109/GLOCOM.2012.6503164.

51. V. Sreejith, K. R. Anupama, L. J. Gudino, and R. Suriyadeepan, "Partition Discovery and Connectivity Restoration in WSN Using Mobile Relays," in *Proceedings of the 16th International Conference on Distributed Computing and Networking* (New York, NY: Association for Computing Machinery, 2015), https://doi.org/10.1145/2684464.2684487.

52. V. Ranga, M. Dave, and A. K. Verma, "Relay Node Placement to Heal Partitioned Wireless Sensor Networks," *Computers and Electrical Engineering* 48 (2015): 371–388, https://doi.org/10.1016/j.compeleceng.2015.09.014.

53. B. Kalyanasundaram and M. Younis, "Using Mobile Data Collectors to Federate Clusters of Disjoint Sensor Network Segments," in *2013 IEEE International Conference on Communications (ICC)* (Budapest, Hungary: IEEE, 2013): 1496–1500, https://doi.org/10.1109/ICC.2013.6654724.

54. F. Senel and M. Younis, "Optimized Interconnection of Disjoint Wireless Sensor Network Segments Using K Mobile Data Collectors," in *2012 IEEE International Conference on Communications (ICC)* (Ottawa, Canada: IEEE, 2012): 492–496. 10.1109/ICC.2012.6364467.

55. A. Zear and V. Ranga, "Path Planning of Unmanned Aerial Vehicles: Current State and Future Challenges," in *First International Conference on Sustainable Technologies for Computational Intelligence*. Advances in Intelligent Systems and Computing, vol. 1045, eds. A. Luhach, J. Kosa, R. Poonia, X. Z. Gao, and D. Singh (Singapore: Springer, 2020), https://doi.org/10.1007/978-981-15-0029-9_32.

56. D. Mishra and R. Kumar, "Cluster Based Fault Detection and Diagnosis Scheme for Wireless Sensor Network," *International Journal of Management, Technology And Engineering* 8 (2024): 240–248.

57. M. Imran, M. Younis, A. Md Said, and H. Hasbullah, "Localized Motion-Based Connectivity Restoration Algorithms for Wireless Sensor and Actor Networks," *Journal of Network and Computer Applications* 35 (2012): 844–856, https://doi.org/10.1016/j.jnca.2011.12.002.

58. F. M. Al-Turjman, H. S. Hassanein, W. M. Alsalih, and M. Ibnkahla, "Optimized Relay Placement for Wireless Sensor Networks Federation in Environmental Applications," *Wireless Communications and Mobile Computing* 11 (2011): 1677–1688, https://doi.org/10.1002/wcm.1211.

59. X. Han, X. Cao, E. L. Lloyd, and C. C. Shen, "Fault-Tolerant Relay Node Placement in Heterogeneous Wireless Sensor Networks," *IEEE Transactions on Mobile Computing* 9 (2010): 643–656, https://doi.org/10.1109/TMC.2009.161.

60. Z. Yun, X. Bai, D. Xuan, T. H. Lai, and W. Jia, "Optimal Deployment Patterns for Full Coverage and κ-Connectivity (κ≤6) Wireless Sensor Networks," *IEEE/ACM Transactions on Networking* 18 (2010): 934–947, https://doi.org/10.1109/TNET.2010.2040191.

61. K. Vaidya and M. Younis, "Efficient Failure Recovery in Wireless Sensor Networks Through Active, Spare Designation," in *DCOSS'10—International Conference on Distributed Computing in Sensor Systems, Adjunct Workshop Proceedings: IWSN, MobiSensors, Poster and Demo Sessions* (Santa Barbara, CA: IEEE, 2010): 1–6, https://doi.org/10.1109/DCOSSW.2010.5593284.

62. A. Zamanifar, O. Kashefi, and M. Sharifi, "Aom: An Efficient Approach to Restore Actor-Actor Connectivity in Wireless Sensor and Actor Networks," *International Journal of Computer Networks & Communications (IJCNC)* 1 (2009): 61–72.

63. X. Bai, Z. Yun, D. Xuan, T. H. Lai, and W. Jia, "Deploying Four-Connectivity and Full-Coverage Wireless Sensor Networks," in *IEEE INFOCOM 2008—The 27th Conference on Computer Communications* (Phoenix, AZ: IEEE, 2008), 296–300.

64. F. Dai and J. Wu, "On Constructing k-Connected k-Dominating set in Wireless Ad Hoc and Sensor Networks," *Journal of Parallel and Distributed Computing* 66 (2006): 947–958, https://doi.org/10.1016/j.jpdc.2005.12.010.

65. B. Hao, J. Tang, and G. Xue, "Fault-Tolerant Relay Node Placement in Wireless Sensor Networks: Formulation and Approximation," in *2004 Workshop on High Performance Switching and Routing, 2004. HPSR* (Phoenix, AZ: IEEE, 2004): 246–250, https://doi.org/10.1109/HPSR.2004.1303479.

66. M. Younis, S. Lee, and A. A. Abbasi, "A Localized Algorithm for Restoring Internode Connectivity in Networks of Moveable Sensors," *IEEE Transactions on Computers* 59 (2010): 1669–1682, https://doi.org/10.1109/TC.2010.174.

67. A. Alfadhly, U. Baroudi, and M. Younis, "Optimal Node Repositioning for Tolerating Node Failure in Wireless Sensor Actor Network," in *2010 25th Biennial Symposium on Communications, QBSC 2010* (Kingston, ON, Canada: IEEE, 2010): 67–71, https://doi.org/10.1109/BSC.2010.5473000.

68. A. Alfadhly, U. Baroudi, and M. Younis, "Least Distance Movement Recovery Approach for Large Scale Wireless Sensor and Actor Networks," in *IWCMC 2011-7th International Wireless Communications and Mobile Computing Conference* (Istanbul, Turkey: IEEE, 2011): 2058–2063, https://doi.org/10.1109/IWCMC.2011.5982851.

69. M. Imran, M. Younis, A. M. Said, and H. Hasbullah, "Volunteer-Instigated Connectivity Restoration Algorithm for Wireless Sensor and Actor Networks," in *IEEE International Conference on Wireless Communications, Networking and Information Security* (Beijing, China: IEEE, 2010): 679–683. 10.1007/978-3-642-20754-9_25.

70. M. L. Chiang, H. C. Hsieh, T. L. Lin, T. P. Chang, and H. W. Chen, "Dynamic Weight-Based Connectivity Recovery in Wireless Sensor and Actor Networks," *Journal of Supercomputing* 80 (2024): 734–760, https://doi.org/10.1007/s11227-023-05518-3.

71. Z. Shen and H. Jiang, "Mobile Relay Scheduling in Partitioned Wireless Sensor Networks," *IEEE Transactions on Vehicular Technology* 65 (2016): 5563–5578, https://doi.org/10.1109/TVT.2015.2455043.

72. Y. K. Joshi and M. Younis, "Distributed Approach for Reconnecting Disjoint Segments," in *2013 IEEE Global Communications Conference (GLOBECOM)* (Atlanta, GA: IEEE, 2013): 255–260, https://doi.org/10.1109/GLOCOM.2013.6831080.

73. I. F. Senturk, K. Akkaya, and F. Senel, "An Effective and Scalable Connectivity Restoration Heuristic for Mobile Sensor/Actor Networks," in *2012 IEEE Global Communications Conference (GLOBECOM)* (Anaheim, CA: IEEE, 2012): 518–523, https://doi.org/10.1109/GLOCOM.2012.6503165.

74. M. Y. Sir, I. F. Senturk, E. Sisikoglu, and K. Akkaya, "An Optimization-Based Approach for Connecting Partitioned Mobile Sensor/Actuator Networks," in *2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)* (Shanghai, China: IEEE, 2011): 525–530, https://doi.org/10.1109/INFCOMW.2011.5928869.

75. A. A. Abbasi, M. Younis, and U. Baroudi, "Restoring Connectivity in Wireless Sensor-Actor Networks With Minimal Node Movement," in *2011 7th International Wireless Communications and Mobile Computing Conference* (Istanbul, Turkey: IEEE, 2011): 2046–2051, https://doi.org/10.1109/IWCMC.2011.5982849.

76. N. Tamboli and M. Younis, "Coverage-Aware Connectivity Restoration in Mobile Sensor Networks," *Journal of Network and Computer Applications* 33 (2010): 363–374, https://doi.org/10.1016/j.jnca.2010.03.008.



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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

77. A. A. Abbasi, M. Younis, K. Akkaya, et al., "Movement-Assisted Connectivity Restoration in Wireless Sensor and Actor Networks," *IEEE Transactions on Parallel and Distributed Systems* 20 (2009): 1366–1379, https://doi.org/10.1109/TPDS.2010.102.

78. B. Chen, H. Chen, and C. Wu, "Obstacle-Avoiding Connectivity Restoration Based on Quadrilateral Steiner Tree in Disjoint Wireless Sensor Networks," *IEEE Access* 7 (2019): 124116–124127, https://doi.org/10.1109/ACCESS.2019.2938225.

79. G. K. Verma and V. Ranga, "Whale Optimizer to Repair Partitioned Heterogeneous Wireless Sensor Networks," *International Journal of Grid and Distributed Computing* 11 (2018): 11–28, https://doi.org/10.14257/ijgdc.2018.11.5.02.

80. F. Senel and M. Younis, "Novel Relay Node Placement Algorithms for Establishing Connected Topologies," *Journal of Network and Computer Applications* 70 (2016): 114–130, https://doi.org/10.1016/j.jnca.2016.04.025.

81. S. Lee, M. Younis, and M. Lee, "Optimized Bi-Connected Federation of Multiple Sensor Network Segments," *Ad Hoc Networks* 38 (2016): 1–18, https://doi.org/10.1016/j.adhoc.2015.12.003.

82. I. F. Senturk, K. Akkaya, and S. Yilmaz, "Relay Placement for Restoring Connectivity in Partitioned Wireless Sensor Networks Under Limited Information," *Ad Hoc Networks* 13 (2014): 487–503, https://doi.org/10.1016/j.adhoc.2013.09.005.

83. S. Lee and M. Younis, "Optimized Relay Node Placement for Connecting Disjoint Wireless Sensor Networks," *Computer Networks* 56 (2012): 2788–2804.

84. F. Senel and M. Younis, "Optimized Connectivity Restoration in a Partitioned Wireless Sensor Network," in *2011 IEEE Global Telecommunications Conference-GLOBECOM 2011* (Houston, TX: IEEE, 2011): 1–5.

85. A. Kashyap, S. Khuller, and M. Shayman, "Relay Placement for Higher Order Connectivity in Wireless Sensor Networks," in *Proceedings IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications* (Barcelona, Spain: IEEE, 2006), https://doi.org/10.1109/INFOCOM.2006.273.

86. Z. Wei, M. Zhu, N. Zhang, et al., "UAV Assisted Data Collection for Internet of Things: A Survey," *IEEE Internet of Things Journal* 9 (2022): 15460–15483, https://doi.org/10.1109/jiot.2022.3176903.

87. O. S. Oubbati, H. Badis, A. Rachedi, A. Lakas, and P. Lorenz, "Multi-UAV Assisted Network Coverage Optimization for Rescue Operations Using Reinforcement Learning," in *2023 IEEE 20th Consumer Communications & Networking Conference (CCNC)* (Las Vegas, NV: IEEE, 2023): 1003–1008.

88. İ. Baştürk, "Energy-Efficient Communication for UAV-Enabled Mobile Relay Networks," *Computer Networks* 213 (2022): 109071.

89. Z. Wang, Y. Cao, Z. Chang, T. Lv, and W. Ni, "Energy Efficiency Maximization in UAV Communication Networks With Nonlinear Energy Harvesting," *Computer Networks* 241 (2024): 110222.

90. S. Ravikumar, S. Sekar, P. Sirenjeevi, and R. Deepa, "Optimizing Resource Allocation in Ultra-Dense Networks With UAV Assistance: A Levy Flight-Based Approach," *Expert Systems With Applications* 235 (2024): 120954, https://doi.org/10.1016/j.eswa.2023.120954.

91. H. Li, X. Hao, J. Wen, F. Liu, and Y. Zhang, "Multi-Objective Deployment of UAVs for Multi-Hop FANET: UAV-Assisted Emergency Vehicular Network," *Drones* 8 (2024): 262, https://doi.org/10.3390/drones8060262.

92. X. Luo, J. Xie, L. Xiong, Z. Wang, and Y. Liu, "UAV-Assisted Fair Communications for Multi-Pair Users: A Multi-Agent Deep Reinforcement Learning Method," *Computer Networks* 242 (2024): 110277, https://doi.org/10.1016/j.comnet.2024.110277.

93. J. H. Kim, M. C. Lee, and T. S. Lee, "Generalized UAV Deployment Design for UAV-Assisted Wireless Networks," in *IEEE International Conference on Communications* (Rome, Italy: IEEE, 2023): 1394–1399, https://doi.org/10.1109/ICC45041.2023.10279344.

94. P. A. Karegar and A. Al-Anbuky, "UAV-Assisted Data Gathering From a Sparse Wireless Sensor Adaptive Networks," *Wireless Networks* 29 (2023): 1367–1384, https://doi.org/10.1007/s11276-022-03194-4.

95. P. Wang, D. Li, Y. Zhang, and X. Chen, "UAV-Assisted Vehicular Communication System Optimization With Aerial Base Station and Intelligent Reflecting Surface," *IEEE Transactions on Intelligent Vehicles* (2023): 1–12, 10.1109/TIV.2023.3324385.

96. A. Sonny, S. R. Yeduri, and L. R. Cenkeramaddi, "Autonomous UAV Path Planning Using Modified PSO for UAV-Assisted Wireless Networks," *IEEE Access* 11 (2023): 70353–70367, https://doi.org/10.1109/ACCESS.2023.3293203.

97. M. Hammami, C. Chaieb, W. Ajib, H. Elbiaze, and R. Glitho, "UAV-Assisted Wireless Networks for Stringent Applications: Resource Allocation and Positioning," in *2023 IEEE Wireless Communications and Networking Conference (WCNC)* (Glasgow, United Kingdom: IEEE, 2023): 1–6, 10.1109/WCNC55385.2023.10118877.

98. C. Lv, Y. Ren, X. Li, et al., "Unmanned Aerial Vehicle-Assisted Sparse Sensing in Wireless Sensor Networks," *IEEE Wireless Communications Letters* 12 (2023): 977–981, https://doi.org/10.1109/LWC.2023.3254580.

99. M. Kang and S. W. Jeon, "Energy-Efficient Data Aggregation and Collection for Multi-UAV-Enabled IoT Networks," *IEEE Wireless Communications Letters* 13 (2024): 1004–1008, https://doi.org/10.1109/LWC.2024.3355934.

100. E. P. De Freitas, T. Heimfarth, I. F. Netto, et al., "UAV Relay Network to Support WSN Connectivity," in *2010 International Congress on Ultra Modern Telecommunications and Control Systems and Workshops, ICUMT 2010* (Moscow, Russia: IEEE, 2010): 309–314, https://doi.org/10.1109/ICUMT.2010.5676621.

101. Y. Chen, W. Feng, and G. Zheng, "Optimum Placement of UAV as Relays," *IEEE Communications Letters* 22 (2018): 248–251, https://doi.org/10.1109/LCOMM.2017.2776215.

102. O. Esrafilian, R. Gangula, and D. Gesbert, "UAV-Relay Placement With Unknown User Locations and Channel Parameters," in *2018 52nd Asilomar Conference on Signals, Systems, and Computers* (Pacific Grove, CA: IEEE, 2019), 1075–1079, https://doi.org/10.1109/ACSSC.2018.8645508.

103. J. Chen, U. Mitra, and D. Gesbert, "Optimal UAV Relay Placement for Single User Capacity Maximization Over Terrain With Obstacles," in *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)* (Cannes, France: IEEE, 2019), 1–5, https://doi.org/10.1109/SPAWC.2019.8815496.

104. W. Gaofeng, G. Xiaoguang, Z. Kun, and F. Xiaowei, "Multi-Objective Placement of Unmanned Aerial Vehicles as Communication Relays Based on Clustering Method," in *2019 Chinese Control and Decision Conference (CCDC)* (Nanchang, China: IEEE, 2019): 1462–1467, https://doi.org/10.1109/CCDC.2019.8833348.

105. D. Orfanus, E. P. De Freitas, and F. Eliassen, "Self-Organization as a Supporting Paradigm for Military UAV Relay Networks," *IEEE Communications Letters* 20 (2016): 804–807, https://doi.org/10.1109/LCOMM.2016.2524405.

106. J. Polo, G. Hornero, C. Duijneveld, A. García, and O. Casas, "Design of a Low-Cost Wireless Sensor Network With UAV Mobile Node for Agricultural Applications," *Computers and Electronics in Agriculture* 119 (2015): 19–32, https://doi.org/10.1016/j.compag.2015.09.024.

107. X. Ma, R. Kacimi, and R. Dhaou, "Fairness-Aware UAV-Assisted Data Collection in Mobile Wireless Sensor Networks," in *2016 International Wireless Communications and Mobile Computing Conference (IWCMC)* (Paphos, Cyprus: IEEE, 2016): 995–1001, https://doi.org/10.1109/IWCMC.2016.7577194.



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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

108. I. Jawhar, N. Mohamed, and J. Al-Jaroodi, "UAV-Based Data Communication in Wireless Sensor Networks: Models and Strategies," in *2015 International Conference on Unmanned Aircraft Systems (ICUAS)* (Denver, CO: IEEE, 2015): 687–694, 10.1109/ICUAS.2015.7152351.

109. I. Jawhar, N. Mohamed, J. Al-Jaroodi, and S. Zhang, "Data Communication in Linear Wireless Sensor Networks Using Unmanned Aerial Vehicles," in *2013 International Conference on Unmanned Aircraft Systems (ICUAS)* (Atlanta, GA: IEEE, 2013): 492–499, https://doi.org/10.1109/ICUAS.2013.6564725.

110. N. Mohamed, H. Aldhaheri, K. Almurshidi, et al., "Using UAVs to Secure Linear Wireless Sensor Networks," in *2016 IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS)* (New York, NY: IEEE, 2016): 424–429, https://doi.org/10.1109/BigDataSecurity-HPSC-IDS.2016.85.

111. J. Lyu, Y. Zeng, R. Zhang, and T. J. Lim, "Placement Optimization of UAV-Mounted Mobile Base Stations," *IEEE Communications Letters* 21 (2017): 604–607, https://doi.org/10.1109/LCOMM.2016.2633248.

112. M. N. Soorki, M. Mozaffari, W. Saad, M. H. Manshaei, and H. Saidi, "Resource Allocation for Machine-To-Machine Communications With Unmanned Aerial Vehicles," in *2016 IEEE Globecom Workshops (GC Wkshps)* (Washington, DC: IEEE, 2016), 1–6, https://doi.org/10.1109/GLOCOMW.2016.7849026.

113. M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Mobile Internet of Things: Can UAVs Provide an Energy-Efficient Mobile Architecture?" in *2016 IEEE Global Communications Conference (GLOBECOM)* (Washington, DC: IEEE, 2016), 1–6, https://doi.org/10.1109/GLOCOM.2016.7841993.

114. M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Unmanned Aerial Vehicle With Underlaid Device-to-Device Communications: Performance and Tradeoffs," *IEEE Transactions on Wireless Communications* 15 (2016): 3949–3963.

115. O. I. Lysenko, I. V. Uriadnikova, S. V. Valuiskyi, S. M. Chumachenko, P. I. Kirchu, and A. I. Semenchenko, "Increasing of Manet Throughput Using Quasi-Optimal UAVs Placement Control," *Science & Military Journal* 8 (2013): 24.

116. J. Huo, Z. Xu, Y. Zhang, and X. Shan, "A UAV Mobile Strategy in Mobile Ad Hoc Networks," in *2011 International Conference on Electronics, Communications and Control (ICECC)* (Ningbo, China: IEEE, 2011): 686–690, https://doi.org/10.1109/ICECC.2011.6066589.

117. M. Gao, B. Zhang, and L. Wang, "A Dynamic Priority Packet Scheduling Scheme for Post-Disaster UAV-Assisted Mobile Ad Hoc Network," in *2021 IEEE Wireless Communications and Networking Conference (WCNC)* (Nanjing, China: IEEE, 2021): 6–11, https://doi.org/10.1109/WCNC49053.2021.9417537.

118. H. Menouar, I. Guvenc, K. Akkaya, A. S. Uluagac, A. Kadri, and A. Tuncer, "UAV-Enabled Intelligent Transportation Systems for the Smart City: Applications and Challenges," *IEEE Communications Magazine* 55 (2017): 22–28, https://doi.org/10.1109/MCOM.2017.1600238CM.

119. O. S. Oubbati, A. Lakas, N. Lagraa, and M. B. Yagoubi, "UVAR: An Intersection UAV-Assisted VANET Routing Protocol," in *2016 IEEE Wireless Communications and Networking Conference* (Doha, Qatar: IEEE, 2016): 2–7, https://doi.org/10.1109/WCNC.2016.7564747.

120. H. Seliem, R. Shahidi, M. H. Ahmed, and M. S. Shehata, "Drone-Based Highway-VANET and DAS Service," *IEEE Access* 6 (2018): 20125–20137, https://doi.org/10.1109/ACCESS.2018.2824839.

121. Y. Azzoug and A. Boukra, "Enhanced UAV-Aided Vehicular Delay Tolerant Network (VDTN) Routing for Urban Environment Using a Bio-Inspired Approach," *Ad Hoc Networks* 133 (2022): 102902, https://doi.org/10.1016/j.adhoc.2022.102902.

122. X. Xu, Y. Zeng, Y. L. Guan, and R. Zhang, "Overcoming Endurance Issue: UAV-Enabled Communications With Proactive Caching," *IEEE Journal on Selected Areas in Communications* 36 (2018): 1231–1244, https://doi.org/10.1109/JSAC.2018.2844979.

123. M. Chen, M. Mozaffari, W. Saad, C. Yin, M. Debbah, and C. S. Hong, "Caching in the Sky: Proactive Deployment of Cache-Enabled Unmanned Aerial Vehicles for Optimized Quality-of-Experience," *IEEE Journal on Selected Areas in Communications* 35 (2017): 1046–1061, https://doi.org/10.1109/JSAC.2017.2680898.

124. T. Fang, H. Tian, X. Zhang, X. Chen, X. Shao, and Y. Zhang, "Context-Aware Caching Distribution and UAV Deployment: A Game-Theoretic Approach," *Applied Sciences (Switzerland)* 8 (2018): 1959, https://doi.org/10.3390/app8101959.

125. H. Wang, J. Wang, G. Ding, L. Wang, T. A. Tsiftsis, and P. K. Sharma, "Resource Allocation for Energy Harvesting-Powered D2D Communication Underlaying UAV-Assisted Networks," *IEEE Transactions on Green Communications and Networking* 2 (2018): 14–24, https://doi.org/10.1109/TGCN.2017.2767203.

126. S. Zhang, H. Zhang, B. Di, and L. Song, "Resource Allocation and Trajectory Design for Cellular UAV-to-X Communication Networks in 5G," in *2018 IEEE Global Communications Conference (GLOBECOM)* (Abu Dhabi, United Arab Emirates: IEEE, 2018), https://doi.org/10.1109/GLOCOM.2018.8647671.

127. V. Sharma, K. Srinivasan, H.-C. C. Chao, K.-L. L. Hua, and W.-H. H. Cheng, "Intelligent Deployment of UAVs in 5G Heterogeneous Communication Environment for Improved Coverage," *Journal of Network and Computer Applications* 85 (2017): 94–105, https://doi.org/10.1016/j.jnca.2016.12.012.

128. N. H. Motlagh, M. Bagaa, and T. Taleb, "UAV-Based IoT Platform: A Crowd Surveillance Use Case," *IEEE Communications Magazine* 55 (2017): 128–134, https://doi.org/10.1109/MCOM.2017.1600587CM.

129. C. Zhan, H. Hu, Z. Liu, Z. Wang, and S. Mao, "Multi-UAV-Enabled Mobile-Edge Computing for Time-Constrained IoT Applications," *IEEE Internet of Things Journal* 8 (2021): 15553–15567.

130. H. Li, J. Li, M. Liu, Z. Ding, and F. Gong, "Energy Harvesting and Resource Allocation for Cache-Enabled UAV Based IoT NOMA Networks," *IEEE Transactions on Vehicular Technology* 70 (2021): 9625–9630, https://doi.org/10.1109/TVT.2021.3098351.

131. S. Y. Park, C. S. Shin, D. Jeong, and H. J. Lee, "DroneNetX: Network Reconstruction Through Connectivity Probing and Relay Deployment by Multiple UAVs in Ad Hoc Networks," *IEEE Transactions on Vehicular Technology* 67 (2018): 11192–11207, https://doi.org/10.1109/TVT.2018.2870397.

132. T. Heimfarth and J. P. De Araujo, "Using Unmanned Aerial Vehicle to Connect Disjoint Segments of Wireless Sensor Network," in *2014 IEEE 28th International Conference on Advanced Information Networking and Applications* (Victoria, BC, Canada: IEEE, 2014): 907–914, https://doi.org/10.1109/AINA.2014.110.

133. I. Jawhar, N. Mohamed, J. Al-Jaroodi, D. P. Agrawal, and S. Zhang, "Communication and Networking of UAV-Based Systems: Classification and Associated Architectures," *Journal of Network and Computer Applications* 84 (2017): 93–108, https://doi.org/10.1016/j.jnca.2017.02.008.

134. A. Mukherjee, S. Misra, and N. S. Raghuwanshi, "A Survey of Unmanned Aerial Sensing Solutions in Precision Agriculture," *Journal of Network and Computer Applications* 148 (2019): 102461, https://doi.org/10.1016/j.jnca.2019.102461.

135. R. Masroor, M. Naeem, and W. Ejaz, "Resource Management in UAV-Assisted Wireless Networks: An Optimization Perspective," *Ad Hoc Networks* 121 (2021): 102596, https://doi.org/10.1016/j.adhoc.2021.102596.

136. M. Basharat, M. Naeem, Z. Qadir, and A. Anpalagan, "Resource Optimization in UAV-Assisted Wireless Networks—A Comprehensive Survey," *Transactions on Emerging Telecommunications Technologies* 33 (2022): e4464.



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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

137. Y. Liu, S. Xie, and Y. Zhang, "Cooperative Offloading and Resource Management for UAV-Enabled Mobile Edge Computing in Power IoT System," *IEEE Transactions on Vehicular Technology* 69 (2020): 12229–12239.

138. H. Peng and X. Shen, "Multi-Agent Reinforcement Learning Based Resource Management in MEC-and UAV-Assisted Vehicular Networks," *IEEE Journal on Selected Areas in Communications* 39 (2020): 131–141.

139. N. P. Sharvari, D. Das, J. Bapat, and D. Das, "Connectivity and Collision Constrained Opportunistic Routing for Emergency Communication Using UAV," *Computer Networks* 220 (2023): 109468, https://doi.org/10.1016/j.comnet.2022.109468.

140. S. M. Mnif and L. A. Saidane, *Mobile, Secure, and Programmable Networking* (Cham: Springer International Publishing, 2016), https://doi.org/10.1007/978-3-319-50463-6.

141. H. Shi, Y. Lu, Z. Hou, and S. Huang, "3D Dubins Net-Recovery Path Planning for Fixed Wing UAV," in *2018 Chinese Control and Decision Conference (CCDC)* (Shenyang, China: IEEE, 2018): 604–610, https://doi.org/10.1109/CCDC.2018.8407203.

142. J. Chen, F. Ye, and Y. Li, "Travelling Salesman Problem for UAV Path Planning With Two Parallel Optimization Algorithms," in *2017 Progress in Electromagnetics Research Symposium—Fall (PIERS—FALL)* (Singapore: IEEE, 2017): 832–837, https://doi.org/10.1109/PIERS-FALL.2017.8293250.

143. Z. Yijing, Z. Zheng, Z. Xiaoyi, and L. Yang, "Q Learning Algorithm Based UAV Path Learning and Obstacle Avoidence Approach," in *2017 36th Chinese Control Conference (CCC)* (Dalian, China: IEEE, 2017): 3397–3402.

144. S. Razzaq, C. Xydeas, M. E. Everett, A. Mahmood, and T. Alquthami, "Three-Dimensional UAV Routing With Deconfliction," *IEEE Access* 6 (2018): 21536–21551, https://doi.org/10.1109/ACCESS.2018.2824558.

145. 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 (2019): 1863–1886, https://doi.org/10.1109/comst.2019.2923708.

146. A. Sharma, P. Vanjani, N. Paliwal, et al., "Communication and Networking Technologies for UAVs: A Survey," *Journal of Network and Computer Applications* 168 (2020): 102739, https://doi.org/10.1016/j.jnca.2020.102739.

147. M. M. U. Chowdhury, W. Saad, and I. Güvenç, "Mobility Management for Cellular-Connected UAVs: A Learning-Based Approach," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)* (Dublin, Ireland: IEEE, 2020): 1–6.

148. O. Dukkanci, B. Y. Kara, and T. Bektaş, "Minimizing Energy and Cost in Range-Limited Drone Deliveries With Speed Optimization," *Transportation Research Part C: Emerging Technologies* 125 (2021): 102985, https://doi.org/10.1016/j.trc.2021.102985.

149. X. Deng, M. Guan, Y. Ma, X. Yang, and T. Xiang, "Vehicle-Assisted UAV Delivery Scheme Considering Energy Consumption for Instant Delivery," *Sensors* 22 (2022): 2045, https://doi.org/10.3390/s22052045.

150. Y. Qin, M. A. Kishk, and M.-S. Alouini, "Stochastic Geometry-Based Analysis of Multi-Purpose UAVs for Package and Data Delivery," *IEEE Internet of Things Journal* 10 (2022): 4664–4676, https://doi.org/10.1109/jiot.2022.3218674.

151. S. W. Jackson, N. A. Riccoboni, A. H. A. Rahim, et al., "Autonomous Airborne Multi-Rotor UAS Delivery System," in *2020 International Conference on Unmanned Aircraft Systems (ICUAS)* (Athens, Greece: IEEE, 2020): 702–708, https://doi.org/10.1109/ICUAS48674.2020.9214011.

152. X. Ma, T. Liu, R. Kacimi, R. Dhaou, and S. Liu, "Duration-Aware Data Collection in UAV-Aided Mobile Sensor Networks," in *2021 International Wireless Communications and Mobile Computing (IWCMC)* (Harbin City, China, 2021): 394–399, https://doi.org/10.1109/IWCMC51323.2021.9498971.

153. T. Feng, L. Xie, J. Yao, and J. Xu, "UAV-Enabled Data Collection for Wireless Sensor Networks With Distributed Beamforming," *IEEE Transactions on Wireless Communications* 21 (2022): 1347–1361, https://doi.org/10.1109/TWC.2021.3103739.

154. K. K. Nguyen, T. Q. Duong, T. Do-Duy, H. Claussen, and L. Hanzo, "3D UAV Trajectory and Data Collection Optimisation via Deep Reinforcement Learning," *IEEE Transactions on Communications* 70 (2022): 2358–2371, https://doi.org/10.1109/TCOMM.2022.3148364.

155. W. Qi, Q. Song, X. Kong, and L. Guo, "A Traffic-Differentiated Routing Algorithm in Flying Ad Hoc Sensor Networks With SDN Cluster Controllers," *Journal of the Franklin Institute* 356 (2019): 766–790.

156. A. Rovira-Sugranes and A. Razi, "Predictive Routing for Dynamic UAV Networks," in *2017 IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE)* (Montreal, QC, Canada: IEEE, 2017): 43–47.

157. F. Aadil, A. Raza, M. F. Khan, M. Maqsood, I. Mehmood, and S. Rho, "Energy Aware Cluster-Based Routing in Flying Ad-Hoc Networks," *Sensors* 18 (2018): 1413.

158. S.-C. Choi, H. R. Hussen, J.-H. Park, and J. Kim, "Geolocation-Based Routing Protocol for Flying Ad Hoc Networks (FANETs)," in *2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN)* (Prague, Czech Republic: IEEE, 2018): 50–52.

159. G. Gankhuyag, A. P. Shrestha, and S.-J. Yoo, "Robust and Reliable Predictive Routing Strategy for Flying Ad-Hoc Networks," *IEEE Access* 5 (2017): 643–654.

160. I. Moon, L. T. Dung, and T. Kim, "Optimal 3D Placement of UAV-BS for Maximum Coverage Subject to User Priorities and Distributions," *Electronics (Basel)* 11 (2022): 1036.

161. A. B. M. Adam, M. S. A. Muthanna, A. Muthanna, T. N. Nguyen, and A. A. Abd El-Latif, "Toward Smart Traffic Management With 3D Placement Optimization in UAV-Assisted NOMA IIoT Networks," *IEEE Transactions on Intelligent Transportation Systems* 24 (2022): 15448–15458.

162. J. Sabzehali, V. K. Shah, Q. Fan, B. Choudhury, L. Liu, and J. H. Reed, "Optimizing Number, Placement, and Backhaul Connectivity of Multi-UAV Networks," *IEEE Internet of Things Journal* 9 (2022): 21548–21560.

163. Y. Liu, W. Huangfu, H. Zhou, H. Zhang, J. Liu, and K. Long, "Fair and Energy-Efficient Coverage Optimization for UAV Placement Problem in the Cellular Network," *IEEE Transactions on Communications* 70 (2022): 4222–4235.

164. M. Khelifi and I. Butun, "Swarm Unmanned Aerial Vehicles (SUAVs): A Comprehensive Analysis of Localization, Recent Aspects, and Future Trends," *Journal of Sensors* 2022 (2022): 8600674.

165. X. Wan, Y. Shao, S. Zhang, and S. Li, "Terrain Aided Planetary UAV Localization Based on Geo-Referencing," *IEEE Transactions on Geoscience and Remote Sensing* 60 (2022): 1–18.

166. A. Gupta and X. Fernando, "Simultaneous Localization and Mapping (SLAM) and Data Fusion in Unmanned Aerial Vehicles: Recent Advances and Challenges," *Drones* 6 (2022): 85, https://doi.org/10.3390/drones6040085.

167. A. Couturier and M. A. Akhloufi, "A Review on Absolute Visual Localization for UAV," *Robotics and Autonomous Systems* 135 (2021): 103666, https://doi.org/10.1016/j.robot.2020.103666.

168. N. Gyagenda, J. V. Hatilima, H. Roth, and V. Zhmud, "A Review of GNSS-Independent UAV Navigation Techniques," *Robotics and Autonomous Systems* 152 (2022): 104069, https://doi.org/10.1016/j.robot.2022.104069.

169. W. Khawaja, I. Guvenc, D. W. Matolak, U. C. Fiebig, and N. Schneckenburger, "A Survey of Air-to-Ground Propagation Channel



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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

Modeling for Unmanned Aerial Vehicles," *IEEE Communications Surveys and Tutorials* 21 (2019): 2361–2391, https://doi.org/10.1109/COMST.2019.2915069.

170. C. Yan, L. Fu, J. Zhang, and J. Wang, "A Comprehensive Survey on UAV Communication Channel Modeling," *IEEE Access* 7 (2019): 107769–107792, https://doi.org/10.1109/ACCESS.2019.2933173.

171. A. Shamsoshoara, F. Afghah, A. Razi, S. Mousavi, J. Ashdown, and K. Turk, "An Autonomous Spectrum Management Scheme for Unmanned Aerial Vehicle Networks in Disaster Relief Operations," *IEEE Access* 8 (2020): 58064–58079.

172. M. A. Jasim, H. Shakhatreh, N. Siasi, A. H. Sawalmeh, A. Aldalbahi, and A. Al-Fuqaha, "A Survey on Spectrum Management for Unmanned Aerial Vehicles (UAVs)," *IEEE Access* 10 (2022): 11443–11499, https://doi.org/10.1109/ACCESS.2021.3138048.

173. M. A. Akram, H. Ahmad, A. N. Mian, A. D. Jurcut, and S. Kumari, "Blockchain-Based Privacy-Preserving Authentication Protocol for UAV Networks," *Computer Networks* 224 (2023): 109638.

174. K. S. Jasim, K. M. Ali Alheeti, and A. K. A. Najem Alaloosy, "Intelligent Detection System for Spoofing and Jamming Attacks in UAVs," in *Advances in Cybersecurity, Cybercrimes, and Smart Emerging Technologies* (Riyadh, Saudi Arabia: Springer, 2023): 97–110.

175. Y. Mekdad, A. Aris, L. Babun, et al., "A Survey on Security and Privacy Issues of UAVs," *Computer Networks* 224 (2023): 109626.



10991131, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/dac.6107 by Delhi Technological University, Wiley Online Library on [15/01/2025]. 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

Physics of Plasmas

RESEARCH ARTICLE | JANUARY 07 2025

# Nonlinear effects associated with Alfvén ion cyclotron wave, and turbulence generation in presence of magnetic islands in solar wind plasmas

Ritu ; Jyoti ; Suresh C. Sharma ; R. P. Sharma

Check for updates

*Phys. Plasmas* 32, 012301 (2025)
https://doi.org/10.1063/5.0227102

View Online
Export Citation

## Articles You May Be Interested In

Preface for frontiers of magnetic reconnection research in heliophysical, astrophysical, and laboratory plasmas

*Phys. Plasmas* (July 2022)

Formation and evolution of coherent structures in 3D strongly turbulent magnetized plasmas

*Phys. Plasmas* (April 2023)

Turbulence-driven magnetic reconnection and the magnetic correlation length: Observations from Magnetospheric Multiscale in Earth's magnetosheath

*Phys. Plasmas* (January 2022)

15 January 2025 10:19:41

AIP Publishing

Physics of Plasmas

Special Topics Open for Submissions

Learn More

AIP Publishing

# Nonlinear effects associated with Alfvén ion cyclotron wave, and turbulence generation in presence of magnetic islands in solar wind plasmas

Cite as: Phys. Plasmas **32**, 012301 (2025); doi: 10.1063/5.0227102
Submitted: 6 July 2024 · Accepted: 23 November 2024 ·
Published Online: 7 January 2025

View Online Export Citation CrossMark

Ritu,[1] Jyoti,[1] Suresh C. Sharma,[1,a)] and R. P. Sharma[2]

## AFFILIATIONS

[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

[a)]**Author to whom correspondence should be addressed:** suresh321sharma@gmail.com

## ABSTRACT

In order to comprehend particle acceleration and transit in interplanetary space, nonlinear processes are of great importance. One process contributing to the heating of the plasmas might be the ponderomotive nonlinearity and magnetic islands that causes the localization of the parallel propagating left-handed circularly polarized dispersive Alfvén wave having a finite frequency. This study investigates the impact of nonlinearity on the creation of localized structures and current sheets formation in solar wind. Pseudo-spectral method and finite difference method with modified predictor-corrector approach are the mathematical models used for spatial integration and temporal integration, respectively, for studying the evolution of left-handed circularly polarized dispersive Alfvén wave. The findings of current study show that the nonlinearity in the system and the presence of magnetic islands are responsible for chaotic structures and turbulent state. Furthermore, to refine the physics behind the localization and current sheets formation the semi-analytical model has been used. The localized structures and current sheets obtained have the scale size of the order of ion inertial length.

*© 2025 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).* https://doi.org/10.1063/5.0227102

15 January 2025 10:19:41

## I. INTRODUCTION

Solar wind is magnetized plasma emerging from the outer surface of sun. Solar wind plasmas support numerous fluctuations including waves, turbulence, and instabilities over a wide range. Solar wind fluctuations are commonly alfvenic and due to nonlinear interactions manifest the turbulent cascade of energy from large length scales to small length scales.[1–3] Waves in solar wind have significant role in several processes, such as particles acceleration, transportation, and heating of plasma.[4] Waves can transfer energy to solar wind via wave particle interactions as well as wave—wave interactions.[5,6] Alfvén waves, Alfvén ion cyclotron waves, and whistler waves are some examples of waves in solar wind.

Alfvén waves are ubiquitous in solar wind. These waves in solar wind are considered to originate in region close to the sun. Alfvén waves are commonly detected in laboratory and astrophysical plasmas having the direction of propagation parallel to the background magnetic field.[7–11] Parallel propagating Alfvén waves are usually non-dispersive in nature but when their frequency become comparable to ion cyclotron frequency these waves show dispersive behavior.[12,13] In the present work, authors used the two-fluid model which is considered to be useful for studying large scale phenomena, such as turbulence, magnetic reconnection, and particle acceleration. The two-fluid model also accounts for hall effect and dispersive waves.[14,15] Furthermore, the dispersive waves are considered to be the most promising wave form in framework of magnetic reconnection.[16,17]

Magnetic reconnection is a non-linear phenomenon that takes place in astrophysical systems such as in solar wind, magnetopause, and magnetosphere.[16,18–20] A quickly changing inhomogeneous magnetic field within a relatively small area is the fundamental characteristic, which is common in these systems. Furthermore, reconnection involves a change in magnetic topology in two dimensions as well as the presence of magnetic X-type neutral point.[21,22] Magnetic reconnection is a remarkable component for understanding the various factors of astrophysical plasmas.[23–26] The dynamical mechanisms in magnetic


© Author(s) 2025

reconnection are considered to be responsible for the appearance of magnetic islands.[27,28] Magnetic reconnection plays a significant role in plasma turbulence.[29–34] Magnetic islands can exist in several forms, such as elongated and round. Magnetic islands play a key role in reconnection processes. The particle acceleration of solar wind might be significantly aided by these magnetic islands.[28,35,36]

Moreover, Alfvén waves are produced in solar wind due to the anisotropy in temperature instabilities.[4,10] Alfvén waves can be excited by electrostatic or electromagnetic waves, energetic charged particles, and non-uniform plasmas particles. The presence of Alfvén wave in solar wind shows significant impact on the several plasmas processes such as magnetic reconnection, turbulence, and heating.[37,38] Finite amplitude circularly polarized Alfvén waves are considered to be an exact solution to nonlinear magnetohydrodynamic equations. In the presence of density fluctuations, these waves are unstable to decay processes.[39]

The observation of Parker Solar Probe has supported the existence of ion scale waves in near the sun.[40–42] The electromagnetic waves having frequency close to the ion cyclotron frequency can have left or right-handed polarization. The left-handed polarized Alfvén waves having frequency in the vicinity of ion cyclotron frequency are also known as Alfvén ion cyclotron (AIC) waves.[43–45] AIC waves are generated locally in solar wind. AIC waves are generated locally in solar wind. As the ion thermal anisotropy approaches equilibrium, anisotropic temperature instabilities are a primary mechanism for producing AIC waves. This occurs when the surplus ion thermal energy is released, and the wave amplitude simultaneously achieves a saturation condition.[46] Moreover, the contribution of AIC wave in wave particle energy exchange has been detected by Parker Solar Probe.[47] The AIC wave plays a significant role in heating the plasma through wave particle interaction. AIC waves can couple with ions via resonance, due to their frequency being in the vicinity of proton gyro-frequency, and may cause the particles to accelerate and heat up. The distinct characteristics of AIC waves suggest that they could be viable means of transferring energy that result in turbulent dissipation and ambient heating.[44,48]

Magnetic reconnection or other processes, such as parametric instabilities that result in the creation of current sheets, coherent or localized structures, can produce turbulent energy transfer over small length scales.[36,49–51] The formation of current sheets at small length scale has been reported by Perri *et al.*[52] in dissipation region of solar wind and is considered to be responsible for energy transfer through magnetic reconnection. Heating, coherent structures, and turbulence are related to each other.[53,54] Many authors have reported the heating and acceleration of the particles of solar wind via AIC waves. Numerous investigations have been done to find out the relationship between magnetic reconnection and turbulence.[55–58] It is still uncertain how precisely AIC wave, magnetic islands, and turbulence relate to one another. The purpose of current work is to study in aforementioned context.

In magnetized plasmas, AIC waves have been studied widely.[59–61] To explain the AIC wave dynamics, the majority of these research have employed linear, quasilinear, or nonlinear models without field perturbation. Therefore, in current work authors study the impact of magnetic islands (field perturbation) and nonlinearity on system near reconnection region in solar wind.

To investigate the combined impact of the wave's nonlinearity and the chain of magnetic islands, we develop a two fluid model in which the coupling is taken into account due to ponderomotive nonlinearity. The nonlinear coupling of high frequency AIC wave and low frequency ion acoustic wave (IAW) results in localized structures, which is a source of transfer of energy to the particles in plasma.

This paper's study is arranged as follows: The dynamics of Alfvén ion cyclotron (AIC) wave and ion acoustic wave are covered in Sect. II. Methods of numerical simulation and the simulation's outcomes are examined in Sec. III. In addition, we have examined the semianalytical model in Sec. IV to comprehend the fundamental physics behind the localization. Section V provides a summary and conclusion of present study.

## II. AIC WAVE DYNAMICS

Let us assume that the magnetic field $B_0$ is aligned in background, in z-direction. The parallel propagating Alfvén wave has frequency near the ion cyclotron frequency. There is a previously existing chain of magnetic islands (field perturbation due to reconnection) in y-direction. The dependence of magnetic islands in y-direction is greater than their dependence in x-direction, i.e., $\delta B_y \gg \delta B_x$. The total ambient magnetic field experienced by Alfvén wave is $B = B_0\hat{z} + \delta B(x,y)\hat{y}$. On combining, Ampère's law and Faraday's law, one will get the following equation:[62]

$$\nabla^2\vec{E} - \nabla.(\nabla\vec{E}) = \frac{1}{c^2}\frac{\partial^2\vec{D}}{\partial t^2}, \tag{1}$$

where $\vec{D} = \tilde{\varepsilon}\vec{E}$, $\tilde{\varepsilon}$ is the dielectric constant. Now, take $\vec{\nabla}.\vec{D} = 0$ and $e^{-i\omega t}$ variations of electric field. One may write the above equation in component form as below

$$\frac{\partial^2 E_x}{\partial y^2} + \frac{\partial^2 E_x}{\partial z^2} - \frac{\partial^2 E_z}{\partial x\partial z} - \frac{\partial^2 E_y}{\partial x\partial y} = -\frac{\omega_0^2}{c^2}(\tilde{\varepsilon}_{xx}E_x + \tilde{\varepsilon}_{xy}E_y), \tag{2}$$

$$\frac{\partial^2 E_y}{\partial x^2} + \frac{\partial^2 E_y}{\partial z^2} - \frac{\partial^2 E_x}{\partial y\partial x} - \frac{\partial^2 E_z}{\partial y\partial z} = -\frac{\omega_0^2}{c^2}(\tilde{\varepsilon}_{yx}E_x + \tilde{\varepsilon}_{yy}E_y), \tag{3}$$

$$\frac{\partial E_z}{\partial z} = -\frac{1}{\tilde{\varepsilon}_{zz}}\left[\tilde{\varepsilon}_{xx}\frac{\partial E_x}{\partial x} + \tilde{\varepsilon}_{xy}\frac{\partial E_y}{\partial x} - \tilde{\varepsilon}_{xy}\frac{\partial E_x}{\partial y} + \tilde{\varepsilon}_{yy}\frac{\partial E_y}{\partial y}\right]. \tag{4}$$

Now, $E_+ = E_x + iE_y$ and $E_- = E_x - iE_y$ are defined as the left and right modes of circularly polarized Alfvén wave, respectively. Using Eq. (4) in Eqs. (2) and (3) and rearranging the terms one may get the coupled wave equation as follows:

$$\begin{aligned}&\frac{\partial^2 E_+}{\partial z^2} + \frac{1}{2}\left(1 + \frac{\tilde{\varepsilon}_{+00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial^2 E_+}{\partial x^2} + \frac{\partial^2 E_+}{\partial y^2}\right)\\ &+ \frac{1}{2}\left(-1 + \frac{\tilde{\varepsilon}_{-00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial^2 E_-}{\partial x^2} + \frac{\partial^2 E_-}{\partial y^2}\right)\\ &- i\left(1 - \frac{\tilde{\varepsilon}_{-00}}{\varepsilon_{0zz}}\right)\frac{\partial^2 E_-}{\partial x\partial y} = -\frac{\omega_0^2}{c^2}\tilde{\varepsilon}_{+0}E_+,\end{aligned} \tag{5}$$

$$\begin{aligned}&\frac{\partial^2 E_-}{\partial z^2} + \frac{1}{2}\left(1 + \frac{\tilde{\varepsilon}_{-00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial^2 E_-}{\partial x^2} + \frac{\partial^2 E_-}{\partial y^2}\right)\\ &+ \frac{1}{2}\left(-1 + \frac{\tilde{\varepsilon}_{+00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial^2 E_+}{\partial x^2} + \frac{\partial^2 E_+}{\partial y^2}\right)\\ &- i\left(1 - \frac{\tilde{\varepsilon}_{+00}}{\varepsilon_{0zz}}\right)\frac{\partial^2 E_+}{\partial x\partial y} = -\frac{\omega_0^2}{c^2}\tilde{\varepsilon}_{-0}E_-.\end{aligned} \tag{6}$$

15 January 2025 10:19:41


© Author(s) 2025

Furthermore, using $\vec{D} = \vec{E} + \frac{4\pi i}{\omega}\vec{j}$ we can write the values of constants as $\varepsilon_{0zz} = 1 - \omega_{pe}^2/\omega^2$ $\tilde{\varepsilon}_{+0} = \frac{\omega_{pe}^2}{\omega_{ce}(\omega-\omega_{ci})}\left(1 - 2\frac{\delta B_y}{B_0} + \frac{\delta n}{n_0}\right)$, $\tilde{\varepsilon}_{-0} = \frac{\omega_{pe}^2}{\omega_{ce}(\omega+\omega_{ci})}\left(1 - 2\frac{\delta B_y}{B_0} + \frac{\delta n}{n_0}\right)$, $\tilde{\varepsilon}_{+00} = \frac{\omega_{pe}^2}{\omega_{ce}(\omega-\omega_{ci})}$ and $\tilde{\varepsilon}_{-00} = \frac{\omega_{pe}^2}{\omega_{ce}(\omega+\omega_{ci})}$. The plasma frequency and gyro-frequency for species 'f' can be defined as $\omega_{pf} = \sqrt{4\pi n_0 e^2/m_f}$ and $\omega_{ci} = eB_0/m_f c$, respectively (where, f = e for electrons and i for ions).

Here, in this work focus will be only on left mode. The left mode of circularly polarized Alfvén wave is also known as Alfvén ion cyclotron (AIC) wave or simply ion cyclotron wave.[59,63,64] These waves have frequency near the ion cyclotron frequency, strongly interact with ions through cyclotron resonance and damped by transferring their energy to plasma particles. Hence, AIC waves are considered to be a viable mean of transferring energy and heating. In the current study, the role of AIC waves is shown in heating and acceleration of particles due to the presence of magnetic islands and ponderomotive nonlinearity in system,

$$\frac{\partial^2 E_+}{\partial z^2} + \frac{1}{2}\left(1 + \frac{\tilde{\varepsilon}_{+00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial^2 E_+}{\partial x^2} + \frac{\partial^2 E_+}{\partial y^2}\right) = -\frac{\omega^2}{c^2}\tilde{\varepsilon}_{+0}E_+. \quad (7)$$

Using the identity of magnetic field and vector potential, i.e., $\vec{B} = \nabla \times \vec{A}$, Eq. (7) can be re-written as follows:

$$\frac{\partial^2 A_+}{\partial z^2} + \frac{1}{2}\left(1 + \frac{\tilde{\varepsilon}_{+00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial^2 A_+}{\partial x^2} + \frac{\partial^2 A_+}{\partial y^2}\right) = -\frac{\omega^2}{c^2}\frac{\omega_{pe}^2}{\omega_{ce}(\omega-\omega_{ci})}\left(1 - 2\frac{\delta A_z}{A_{0y}} + \frac{\delta n}{n_0}\right)A_+. \quad (8)$$

### A. Ion acoustic wave dynamics

Consider the ion acoustic wave (IAW) having low frequency and propagation parallel to the magnetic field present in background. The fundamental equations governing the dynamics of wave are as follows:

(i) Equation of motion

$$\frac{\partial v_f}{\partial t} = \frac{q_f \vec{E}}{m_f} + \frac{q_f}{cm_f}(\vec{v}_f \times \vec{B}) + \frac{\vec{F}_f}{m_f} - \frac{T_f}{m_f}\nabla\frac{n_f}{n_0}. \quad (9)$$

(ii) Equation of Continuity

$$\frac{\partial n_f}{\partial t} + \nabla.(n_f \vec{v}_f) = 0. \quad (10)$$

(iii) Ponderomotive Force

$$\vec{F}_f = \frac{q_f}{c}\left(\vec{v}_f^{\,*} \times \vec{B}\right) - m_f\left(\vec{v}_f^{\,*}.\vec{\nabla}\right)\vec{v}_f, \quad (11)$$

where $n_f$ is the density, $m_f$ is the mass, $\vec{v}_f$ is the velocity, $\vec{v}_f^*$ is the complex conjugate of velocity, and $q_f$ is the charge of the species 'f' of plasma (f = e for electrons & i for ions) and c is the speed of light in vacuum. Furthermore, in terms of scalar potential electric field is defined as $\vec{E} = -\vec{\nabla}\phi$. The thermal velocity of electron has been considered to be very much large than the phase velocity of wave,[65] the scalar potential can be written as

$$\phi = \frac{T_e \delta n_s}{en_0} - \frac{F_{ez}}{ik_z e}. \quad (12)$$

Upon using Eq. (12) and the components of equation of motion, we will get the following form:

$$\frac{\partial v_{ix}}{\partial t} = \frac{\Omega^2 \rho_s^2}{n_0}\frac{\partial n}{\partial x} - \frac{\Omega^2 k_x F_{ez}}{\omega_{ci}^2 m_i k_z} - \frac{\Omega^2 F_{ix}}{\omega_{ci}^2 m_i} - \frac{i\Omega F_{iy}}{\omega_{ci} m_i}, \quad (13)$$

$$\frac{\partial v_{iz}}{\partial t} = -\frac{c_s^2}{n_0}\frac{\partial n}{\partial z} + \frac{(F_{ez} + F_{iz})}{m_i}. \quad (14)$$

Now, using Eqs. (13) and (14) in equation of continuity, the equation of continuity takes the following form:

$$\frac{\partial^2 n}{\partial t^2} - \rho_s^2\frac{\partial^2}{\partial t^2}\frac{\partial^2 n}{\partial x^2} - c_s^2\frac{\partial^2 n}{\partial z^2} = \frac{\Omega^2 n_0}{\omega_{ci}^2 m_i}\frac{\partial F_{ix}}{\partial x} + \frac{i\Omega n_0}{m_i \omega_{ci}}\frac{\partial F_{iy}}{\partial x} + \frac{\Omega^2 n_0 k_x}{\omega_{ci}^2 m_i k_z}\frac{\partial F_{ez}}{\partial x} - \frac{n_0}{m_i}\frac{\partial (F_{ez} + F_{iz})}{\partial z}. \quad (15)$$

Here, Ω is the frequency of IAW and its value is much less than the frequency of AIC wave. The constants are defined as $c_s = \sqrt{(T_e + T_i)/m_i}$ is the ion sound speed and $\rho_s = (c_s/\omega_{ci})$ is ion acoustic gyro-radius. The ponderomotive force components (i.e., $F_{ix}, F_{iy}, F_{iz}$ and $F_{ez}$) are obtained using Eq. (11) and have the form

$$F_{ix} = -\frac{e^2\omega^2}{8m_i(\omega^2 - \omega_{ci}^2)^2}\left[2 - \frac{\omega_{ci}}{\Omega}\left(3 - \frac{\omega_{ci}^2}{\Omega^2}\right)\right]\frac{\partial |A_z|^2}{\partial x}, \quad (16)$$

$$F_{iy} = -\frac{ie^2}{16m_i(\omega^2 - \omega_{ci}^2)^2}\left[(\omega^2 - \omega_{ci}^2)^2 + \left(1 + \frac{\omega_{ci}}{\omega}\right)(\omega^2 - \omega_{ci}^2)\right]\frac{\partial |A_z|^2}{\partial z}, \quad (17)$$

$$F_{iz} = -\frac{e^2}{8m_i(\omega^2 - \omega_{ci}^2)}\left(1 - \frac{\omega_{ci}}{\omega}\right)\frac{\partial |A_z|^2}{\partial z}, \quad (18)$$

and

$$F_{ez} = -\frac{e^2}{8m_e(\omega^2 - \omega_{ce}^2)}\left(1 - \frac{\omega_{ce}}{\omega}\right)\frac{\partial |A_z|^2}{\partial z}, \quad (19)$$

where A is the field of AIC wave. Using ponderomotive force components [i.e., Eq. (16)–(19)] in Eq. (15), density fluctuation will take the following form: $\frac{\partial^2 n}{\partial t^2} - \rho_s^2\frac{\partial^2}{\partial t^2}\frac{\partial^2 n}{\partial t^2} - c_s^2\frac{\partial^2 n}{\partial z^2} = -\frac{e^2\Omega^2}{16m_i^2\omega_{ci}^2}\left[\frac{2(1+\omega_{ce}/\omega)m_i}{(\omega^2-\omega_{ce}^2)m_e} + \frac{\omega^2}{(\omega^2-\omega_{ci}^2)^2}\left\{2 - \frac{\omega_{ci}}{\omega}\left(3 - \frac{\omega_{ci}^2}{\omega^2}\right)\right\}\right]\frac{\partial^2 |A_z|^2}{\partial x^2} - \frac{e^2(1 - \omega/\omega_{ci})}{8m_i^2(\omega_{ci}^2 - \omega^2)}\frac{\partial^2 |A_z|^2}{\partial z^2}$. (20)

The above equation represents the time-dependent variation of the acoustic density perturbation for the effect of ponderomotive nonlinearity, which is determined by the AIC wave.

Now, taking the steady state approximation and following Jyoti *et al.*,[20] density perturbation will take the form:

$$\frac{\delta n}{n_0} = \exp(-\alpha_0 |A_z|^2) - 1, \quad (21)$$

where $\alpha_0 = \frac{\omega_{pi}^2}{32\pi n_0 T(\omega_{ci}^2 - \omega^2)}\left(1 - \frac{\omega}{\omega_{ci}}\right)$.

15 January 2025 10:19:41


© Author(s) 2025

Furthermore, Eq. (8) is solved semi-analytically by considering the wave solution[66] as follows:

$$A_+ = A(x,y,z,t)e^{-i(\omega t - k_+ z)}, \tag{22}$$

where A is the complex amplitude of AIC wave and $k_+\left(=\frac{\omega}{c}\tilde{\varepsilon}_{+00}^{1/2}\right)$ is the wave number of wave. Substituting, Eqs. (21) and (22) in Eq. (8) and further, assuming $\partial_z A \ll k_+ A$, we will have

$$\frac{2i}{\omega}\frac{\partial A}{\partial t} + \frac{2i}{k_+}\frac{\partial A}{\partial z} + \frac{1}{2k_+^2}\left(1+\frac{\tilde{\varepsilon}_{+00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial^2 A}{\partial x^2}+\frac{\partial^2 A}{\partial y^2}\right) - 2\frac{\delta A_z}{A_{0y}}A + (\exp(\alpha_0|A_z|^2-1))A = 0. \tag{23}$$

The dimensionless form of Eq. (23) can be written as

$$i\frac{\partial A}{\partial t} + iC_1\frac{\partial A}{\partial z} + C_2\left(\frac{\partial^2 A}{\partial x^2}+\frac{\partial^2 A}{\partial y^2}\right) + \left(2\frac{\delta A_z}{A_0}-\frac{\delta n}{n_0}\right)A = 0, \tag{24}$$

where $C_1$&$C_2$ 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 chains of magnetic islands, i.e., magnetic reconnection has been considered within the framework of magnetic islands. The magnetic islands profile[67] is considered as

$$\frac{\delta A_z}{A_0} = -\frac{x^2}{2} + 0.5\cos ky. \tag{25}$$

Using Eq. (25) in Eq. (24), the dynamical equation of the AIC wave is obtained. Here, the magnetic perturbation will dominate over the background density perturbation. This will be useful in basic understanding of localized structures and current sheet formation, under the influence of magnetic islands. Finally, we have

$$i\frac{\partial A}{\partial t} + iC_i\frac{\partial A}{\partial z} + C_2\left(\frac{\partial^2 A}{\partial x^2}+\frac{\partial^2 A}{\partial y^2}\right) + \left(2\left(-\frac{x^2}{2}+0.5\cos ky\right)-\frac{\delta n}{n_0}\right)A = 0. \tag{26}$$

Equation (26) represents the nonlinear dynamical equation of AIC wave, which will be solved further by numerical simulations. This equation is further solved by using the solar wind parameters given by Alexandrova *et al.*[68]

It should also be emphasized that the value of ion inertial length $(\lambda_i = c/\omega_{pi})$ and ion gyroradius $(\rho_s = c_s/\omega_{ci})$ are approximately equal (i.e., the value of plasma beta is nearly one).

## III. NUMERICAL SIMULATION

A modified non-linear Schrödinger model is developed for solving Eq. (26). Equation (26) is solved numerically, in 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 $(5\times10^{-5})$. The spatial domain size in this simulation results in the appearance of more wave modes than what is standard in Biskamp's or similar studies. This setup provides a better resolution for smaller wavenumbers. Power spectrum for smaller wavenumbers may become more accurate and representative of real physical systems because it encompasses a broader range of wave numbers. The perturbations in wave number of AIC wave 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. The numerical methodology that is occupied for time evolution making use of finite difference method along with predictor–corrector scheme,[69] and for spatial integration portion pseudo spectral method is utilized. The simplicity of the finite difference method over other approaches, including the finite element and finite volume methods, makes it significantly beneficial. The finite difference method takes less time for a given grid. Computationally, this method is very accurate, and efficient, 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 non-linear Schrödinger equation and then the accuracy of code has been 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 non-linear Schrödinger equation and in present study we consider the initial condition for numerical simulation as

$$\tilde{A}_1(x,y) = \cos(2x+2.3) + \cos(y+4.1). \tag{27}$$

Numerical simulations are performed for solar wind at 1 AU. The typical plasma parameters of solar wind,[68] used in simulation and calculations, are as follows: $T_i \approx 5.8\times10^5 K, T_e \approx 1.4\times10^5 K, n_0 \approx 3cm^{-3}$, $B_{0z} \approx 6.2\times10^{-5}G, \omega_{ci} \approx 0.594Hz, v_A \approx 7.815\times10^6 cm/s, \rho_s \approx 1.298\times10^7 cm, \lambda_i \approx 1.315\times10^7 cm, \omega_{pe} \approx 9.767\times10^4 Hz$ and the other parameters are as follows: $\omega = 1.2\omega_{ci}, \omega_{ci} = 0.594Hz, k_+ = 2.0398\times10^{-7}cm^{-1}$, and $A_n = 3.029\times10^{-8}$ statV/cm.

### A. Results and discussion

Numerical simulations have been performed for Eq. (26) by imposing initial conditions as given in Eq. (27).

**Case I: Only ponderomotive nonlinearity is present**

The projection plots, contour plots, and current sheets have been studied for different values of time. Figure 1 represents the localized structures of AIC wave at different times. The figure shows that there is symmetry in structures at time t = 0. At later period, the structures become chaotic, indicating turbulence. The modification occurs in the structures due to density perturbation.

**Case II: Ponderomotive nonlinearity as well as magnetic islands are present**

The projection plots shown in Fig. 2 also depict symmetry at the early stage of time, and as the time evolves this symmetry breakdown and, we obtain irregular structures at later period of time. It can be clearly observed from Figs. 1 and 2, the localization of AIC wave is more intense in case when both preexisting chain of magnetic islands as well as density modification are present. The localized structures are of the order of ion inertial length.

From Figs. 1 and 2, it is clear that basic evolution of the system is similar but it is not entirely independent of the presence of magnetic islands, rather they play a minor role showing the change in $|A_Z|^2$ is not very large. Figure 3 (left panel) represents the contour plots of

15 January 2025 10:19:41


© Author(s) 2025

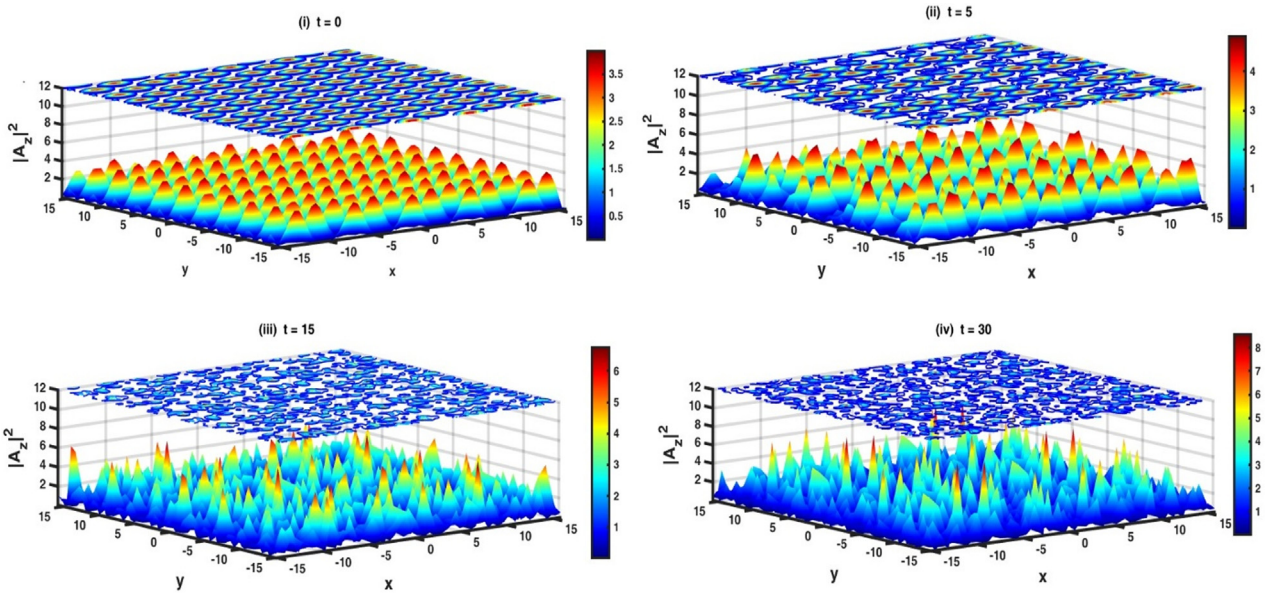



**FIG. 1.** Representation of projection plots of square amplitude of vector potential, $|A_z|2$ (in normalized units) of AIC wave for different time at (i) t = 0, (ii) t = 5, (iii) t = 15, and (iv) t = 30 in the presence of ponderomotive nonlinearity only.

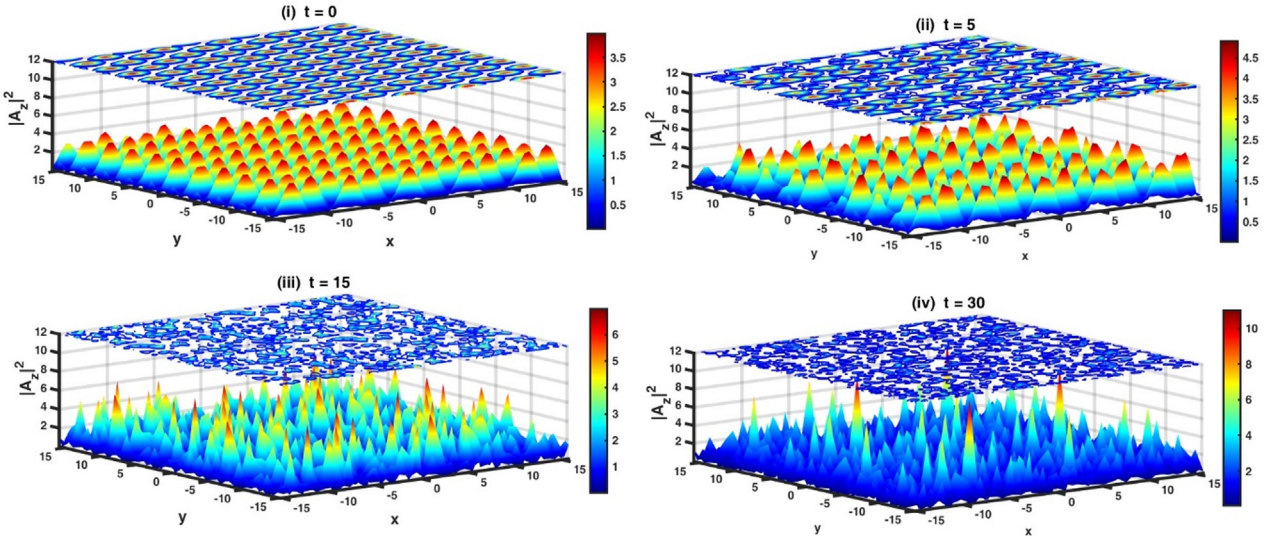



**FIG. 2.** Representation of projection plots of square amplitude of vector potential, $|A_z|2$ (in normalized units) of AIC wave for different time at (i) t = 0, (ii) t = 5, (iii) t = 15, and (iv) t = 30 in the presence of ponderomotive nonlinearity as well as magnetic islands.

vector potential at different times. The contour plots are utilized to depict the presence of magnetic islands. Here, the size of structures at time t = 0 is nearly $3\lambda_i$ and we see that the structures modified with time showing high intensity. As shown in Fig. 3 (right panel), with the time evolution, spatial evolution of current density, i.e., $j = -\nabla^2 A_z$, has been studied. One can also see from Fig. 3 (right panel), the structures of X-O points begin to deform with increase in time and result in chaotic structures. The formation of the chaotic structures is signature of the generation of turbulence. However, $|A_Z|^2$ is not very large the right panels of Fig. 3 would appear more or less similar for the case without initial magnetic islands also.

### B. Quasi-steady state of system

To observe the turbulent power spectrum, we will take quasi-steady state into account (i.e., $\partial/\partial t \approx 0$). So, Eq. (26) will become

15 January 2025 10:19:41


© Author(s) 2025

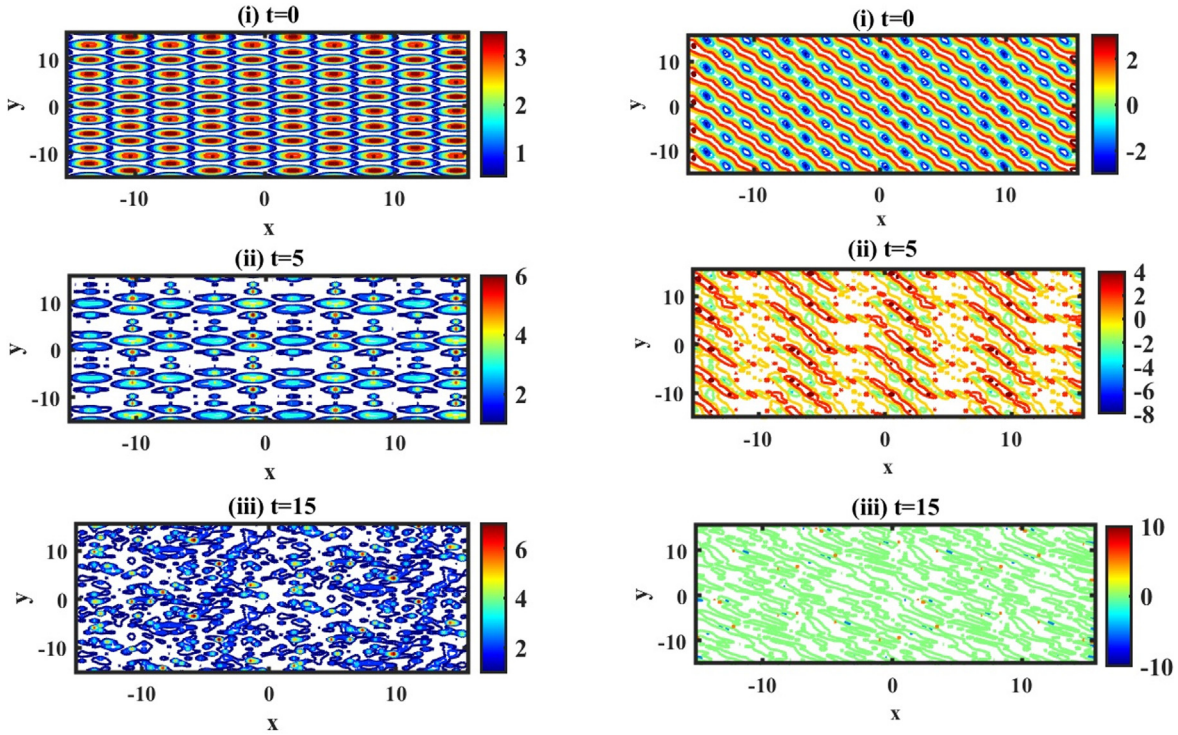



**FIG. 3.** Representation of contour plots of magnetic vector potential in left panel (i)-(iii) and current density in right panel (i)-(iii), of AIC wave under the effect of modified density and magnetic islands in background at different time.

$$iC_1\frac{\partial A}{\partial z}+C_2\left(\frac{\partial^2 A}{\partial x^2}+\frac{\partial^2 A}{\partial y^2}\right)+\left(2\left(-\frac{x^2}{2}+0.5\cos ky\right)-\frac{\delta n}{n_0}\right)A=0. \tag{28}$$

For solving the above equation, the finite difference method has been used in z- direction and pseudo-spectral method along x- and y-directions. Figure 4 shows the power spectra, obtained by plotting the results obtained from separate numerical simulation with $(512\times512)$ grid points to get better resolution. Furthermore, the simulation results obtained for vector potential are Fourier transformed and the square amplitude is calculated. There are various methods for plotting wave spectrum, such as (a) finite $k_\perp$ and $k_{||}=0$ (b) finite $k_{||}$ and $k_\perp=0$ (c) finite

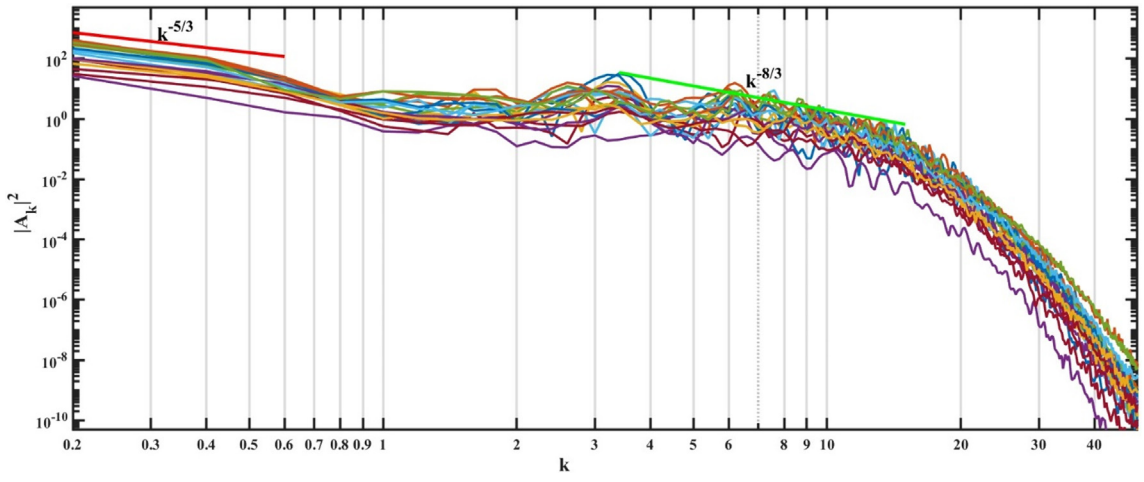



**FIG. 4.** Power spectra $|A_k|^2$ vs wave number k of AIC wave between z = 30 to z = 50 in the presence of background density perturbation (ponderomotive) and magnetic islands (in normalized units).

15 January 2025 10:19:41


© Author(s) 2025

$k_\perp$ or $k_{||}$ and the others. We have employed the first approach. The power spectrum is obtained for studying the distribution of energy over different length scales. The steepness of power spectrum at small length scales ($k\lambda_i \gg 1$) reveals 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 purpose. The resultant spectra indicate an approximate scaling of $k^{-5/3}$ in inertial range, and a scaling of $k^{-8/3}$ is followed in dissipation range as shown by Bowen *et al.*[48] The dissipation range where the spectra become steeper is a signature of turbulence.

### C. Particle acceleration & heating

The power spectrum becomes steep at smaller scales due to dispersive properties of wave and indicates transfer of energy to smaller length scales from larger ones. The Fokker–Planck equation can cause energization when charged particles contact with localized structures repeatedly. This approach uses turbulence to accelerate charged particles through repeated interactions with magnetic clouds. Various methods have been used to inspect the acceleration of charged particles. The present study examines the localized structures of nonlinear AIC waves through nonlinear coupling and turbulence generation. As a result of their repeated interactions with localized structures, charged particles in this situation accumulate energy and accelerate. The following formula yields a quasi-linear diffusion equation:[70]

$$\frac{\partial g(v,t)}{\partial t} = \frac{\partial}{\partial v}\left\{D(v)\frac{\partial g(v,t)}{\partial v}\right\}.$$

Here, $g(v,t)$ indicates the velocity distribution function and $D(v)$ stands for the velocity space diffusion coefficient. The diffusion equation given above is used to represent the recurring interaction of charged particles with the localized structures. In case the ponderomotive nonlinearity setting up time is smaller than observation time, the velocity distribution function might remain constant with respect to time. The distribution function, $g(v)$ and spectral index, $\mu$[71,72] in this context are related in the following way:

$$g(v) \propto v^{2-\mu}.$$

In this study, $\mu \sim 2.7$. Spectral index promotes the enhancement of thermal tails in energetic particles, leading to their acceleration. A steeper spectral index (larger $\mu$) generally indicates more energy dissipation at smaller scales, while a shallower index (smaller $\mu$) suggests efficient energy transfer across scales or acceleration of particles to higher energies.

## IV. SEMI-ANALYTICAL MODEL

The semianalytical analysis is done for the better understanding of the physics behind the localization of the AIC wave. Taking, quasi steady state of Eq. (23), one will obtain Eq. (29). This approach will be helpful in the estimation of the size of current sheets, localized structures, and the influence of field, and density perturbation on these structures. For steady state of Eq. (23), one will have

**Case I:** When the Hall term (i.e., $\omega/\omega_{ci}$) is taken

$$2ik_+\frac{\partial A}{\partial z} + \frac{1}{2}\left(1+\frac{\tilde{\varepsilon}_{+00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial^2 A}{\partial x^2}+\frac{\partial^2 A}{\partial y^2}\right) - 2\left(-\frac{x^2}{2}+0.5\cos ky\right)k_+^2 A + (\exp(\alpha_0|A|^2-1))k_+^2 A = 0. \quad (29)$$

For solving above Eq. (29), the solution of vector potential is considered[73]

$$A(x,y,z) = A_0(x,y,z)e^{ik_+ S(x,y,z)}. \quad (30)$$

Using above equation in Eq. (29), the real and imaginary parts can be written as given by the following Eqs. (31) and (32), respectively:

$$2\frac{\partial S}{\partial z} - \frac{1}{2}\left(1+\frac{\tilde{\varepsilon}_{+00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial^2 A_0}{\partial x^2}+\frac{\partial^2 A_0}{\partial y^2}\right)\frac{1}{k_+^2 A_0} + \frac{1}{2}\left(1+\frac{\tilde{\varepsilon}_{+00}}{\varepsilon_{0zz}}\right)\left(\left(\frac{\partial S}{\partial x}\right)^2+\left(\frac{\partial S}{\partial y}\right)^2\right) = \left(\exp(\alpha|A_0|^2)-1\right)+2\left(\frac{x^2}{2}-0.5\cos ky\right), \quad (31)$$

$$\frac{\partial A_0^2}{\partial z} + \frac{A_0^2}{2}\left(1+\frac{\tilde{\varepsilon}_{+00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial^2 S}{\partial x^2}+\frac{\partial^2 S}{\partial y^2}\right) + \frac{1}{2}\left(1+\frac{\tilde{\varepsilon}_{+00}}{\varepsilon_{0zz}}\right)\left(\frac{\partial A_0^2}{\partial x}\frac{\partial S}{\partial x}+\frac{\partial A_0^2}{\partial y}\frac{\partial S}{\partial y}\right) = 0. \quad (32)$$

Furthermore, the solution of above equations is given by assuming the Gaussian profiles as follows:

$$A_0^2(x,y,z) = \frac{E_{00}^2(x,y,z)}{f_1 f_2}\exp\left(-\frac{x^2}{r_{01}^2 f_1^2}-\frac{y^2}{r_{02}^2 f_2^2}\right), \quad (33)$$

$$S(x,y,z) = \beta_1(z)\frac{x^2}{2}+\beta_2(z)\frac{y^2}{2}+\phi(z). \quad (34)$$

Here, $r_{01}, r_{02}$ represent the transverse scale size of AIC wave and S represents the eikonal of wave. Also, $\beta_1(z)$ and $\beta_2(z)$ vary slowly with respect to z and their values are given by following equations:

$$\beta_1(z) = \frac{2}{\left(1+\frac{\tilde{\varepsilon}_{+00}}{\varepsilon_{0zz}}\right)}\frac{1}{f_1}\frac{df_1}{dz}, \quad \text{and} \quad \beta_1(z) = \frac{2}{\left(1+\frac{\tilde{\varepsilon}_{+00}}{\varepsilon_{0zz}}\right)}\frac{1}{f_2}\frac{df_2}{dz}, \quad (35)$$

where $f_1$, $f_2$ represent the dimensionless beam width parameters of AIC wave.

Now, the beam width parameters of AIC wave are obtained by solving Eqs. (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.,

$$\frac{d^2 f_1}{dZ^2} = \frac{1}{4}\left(1+\frac{\tilde{\varepsilon}_{+00}}{\varepsilon_{0zz}}\right)^2\left(\frac{1}{f_1^3}-\frac{2R_{d1}^2 f_1}{\lambda_i^2(1+\tilde{\varepsilon}_{+00}/\varepsilon_{0zz})} - \frac{2|E_{00}|^2 R_{d1}^2\alpha_0}{r_{01}^2 f_1^2 f_2(1+\tilde{\varepsilon}_{+00}/\varepsilon_{0zz})}\exp\left(\frac{\alpha_0|E_{00}|^2}{f_1 f_2}\right)\right), \quad (36)$$

$$\frac{d^2 f_2}{dZ^2} = \frac{1}{4}\left(1+\frac{\tilde{\varepsilon}_{+00}}{\varepsilon_{0zz}}\right)^2\left(\frac{1}{f_2^3}-\frac{2R_{d2}^2 0.5k^2 f_2}{\lambda_i^2(1+\tilde{\varepsilon}_{+00}/\varepsilon_{0zz})} - \frac{2|E_{00}|^2 R_{d2}^2\alpha_0}{r_{02}^2 f_1 f_2^2(1+\tilde{\varepsilon}_{+00}/\varepsilon_{0zz})}\exp\left(\frac{\alpha_0|E_{00}|^2}{f_1 f_2}\right)\right), \quad (37)$$

where $R_{d1} = k_+ r_{01}^2$ and $R_{d2} = k_+ r_{02}^2$. We have solved Eqs. (34) and (35) for plane wavefront, using the initial conditions $f_1 = f_2 = 1$ at

15 January 2025 10:19:41


© Author(s) 2025

$z = 0$ and $\frac{df_1}{dz} = \frac{df_2}{dz} = 0$ at $z = 0$. The last two terms of Eqs. (36) and (37) are due to the presence of background density perturbation and magnetic islands. Both these terms have opposite sign from the first term and acts in converging manner. There continues a competition between the first terms and last two terms, due to which AIC wave gets localized. The converging and diverging effects are not same in x and y directions. Hence, the rate of change of $f_1, f_2$ along x and y direction will be different and results in the dissimilarity in localized structures along x and y directions,

$$r_{01} = \left((\lambda_i^2/2k_+^2)\left(1 + \frac{\varepsilon_{+00}}{\varepsilon_{0zz}}\right)\right)^{1/4} = 0.518\,\lambda_i, \tag{38}$$

$$r_{02} = \left((\lambda_i^2/k^2k_+^2)\left(1 + \frac{\varepsilon_{+00}}{\varepsilon_{0zz}}\right)\right)^{1/4} = 2.727\,\lambda_i. \tag{39}$$

The values of $r_{01}$ and $r_{02}$ are obtained by the comparison of converging and diverging terms, taking ponderomotive nonlinearity to be zero.

Now, take the nonlinearity to be present in the system due to the ponderomotive force and magnetic islands. In this case, the scale size of localized structures can be obtained by solving Eqs. (36) and (37). These equations can be re-written as follows:

$$\frac{1}{\lambda_i^2} - \frac{2k_+^2}{(1 + \varepsilon_{+00}/\varepsilon_{0zz})}\frac{r_{01}^4}{\lambda_i^4} - \frac{2|E_{00}|^2\alpha_0 k_+^2}{(1 + \varepsilon_{+00}/\varepsilon_{0zz})}\exp(\alpha_0|E_{00}|^2)\left(\frac{r_{01}^2}{\lambda_i^2}\right) = 0 \tag{40}$$

$$\frac{1}{\lambda_i^2} - \frac{k_+^2k^2}{(1 + \varepsilon_{+00}/\varepsilon_{0zz})}\frac{r_{02}^4}{\lambda_i^4} - \frac{2|E_{00}|^2\alpha_0 k_+^2}{(1 + \varepsilon_{+00}/\varepsilon_{0zz})}\exp(\alpha_0|E_{00}|^2)\left(\frac{r_{02}^2}{\lambda_i^2}\right) = 0 \tag{41}$$

As the above equations are quadratic in nature and the root of Eqs. (40) and (41) can be written in the following form:

$$\left(\frac{r_{01}^2}{\lambda_i^2}\right) = \frac{1}{2}\left(-\alpha_0|E_{00}|^2\exp(\alpha_0|E_{00}|^2) + \left(|E_{00}|^4\alpha_0^2\exp(2\alpha_0|E_{00}|^2) + \frac{2(1 + \varepsilon_{+00}/\varepsilon_{0zz})}{k_+^2\lambda_i^2}\right)^{1/2}\right), \tag{42}$$

$$\left(\frac{r_{02}^2}{\lambda_i^2}\right) = \left(-\frac{\alpha_0|E_{00}|^2}{k^2}\exp(\alpha_0|E_{00}|^2) + \left(\frac{|E_{00}|^4\alpha_0^2\exp(2\alpha_0|E_{00}|^2)}{k^2} + \frac{(1 + \varepsilon_{+00}/\varepsilon_{0zz})}{k^2k_+^2\lambda_i^2}\right)^{1/2}\right). \tag{43}$$

One may get insight into the scale size by plotting $\alpha_0|E_{00}|^2$ vs $r_{01}^2/\lambda_i^2$ and $\alpha_0|E_{00}|^2$ vs $r_{02}^2/\lambda_i^2$ (i.e., AIC field strength V s scale size). It is clear from Fig. 5, when the ponderomotive nonlinearity is also taken into account the scale size decrease with increase in nonlinearity. Here, it should also be noticed that the resolution size $\left(\frac{10\pi\times\lambda_i}{128}\right)$ of the numerical model under consideration is approximately equals to $0.245\,\lambda_i$.

**Case II:** When Hall term is neglected (i.e., $\omega/\omega_{ci} \to 0$), dispersive Alfvén wave tends to non-dispersive Alfvén wave. The dispersion relation for non-dispersive Alfvén wave $(i.e., \omega = k_z v_A)$ can be obtained from the x-component of wave equation. Taking only the x components of wave equation, we have

$$\frac{\partial^2 E_x}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2 D}{\partial t^2}. \tag{44}$$

Using the identity $\vec{D} = \vec{E} + \frac{4\pi i}{\omega}\vec{j}$ (where, $\vec{j} = N_0 e(\vec{v}_i - \vec{v}_e)$) in above equation, this equation becomes

$$\frac{\partial^2 E_x}{\partial z^2} = \frac{4\pi N_0 e^2}{m_i\omega_{ci}^2 c^2}\frac{\partial^2 E_x}{\partial t^2}, \tag{45}$$

where $\frac{m_i\omega_{ci}^2c^2}{4\pi N_0 e^2} = v_A^2$. This equation gives $\omega = k_z v_A$, which is the dispersion relation for non-dispersive Alfvén wave. However, if we include $E_z$ then

$$\frac{\partial^2 E_x}{\partial z^2} - \frac{\partial^2 E_z}{\partial x \partial z} = \frac{1}{c^2}\frac{\partial^2 D}{\partial t^2}, \tag{46}$$

$$\frac{\partial E_z}{\partial z} = -\frac{1}{\varepsilon_{zz}}\left(\varepsilon_{xx}\frac{\partial E_x}{\partial x}\right). \tag{47}$$

Again, using the same identity $\vec{D} = \vec{E} + \frac{4\pi i}{\omega}\vec{j}$ and putting value of Eq. (47) in Eq. (46), one may obtain the following equation:

$$\frac{\partial^2 E_x}{\partial z^2} + \frac{\varepsilon_{xx}}{\varepsilon_{zz}}\frac{\partial^2 E_x}{\partial x^2} = \frac{\partial^2}{\partial t^2}\left\{\frac{1}{c^2}\left(E_x + \frac{4\pi i N_0 e}{\omega}(v_{ix} - v_{ex})\right)\right\}, \tag{48}$$

$$\frac{\partial^2 E_x}{\partial z^2} + \frac{\varepsilon_{xx}}{\varepsilon_{zz}}\frac{\partial^2 E_x}{\partial x^2} = -\frac{1}{v_A^2}\left(1 - 2\frac{\delta B_y}{B_{0z}} + \frac{\delta n}{n_0}\right)\frac{\partial^2 E_x}{\partial t^2}. \tag{49}$$

In Eq. (49) when $\omega/\omega_{ci} \to 0$, dispersive Alfvén wave tends to be non-dispersive in nature. Furthermore, we want to see its effect on the formation of current sheets. Now, in terms of vector potential, above equation can be rewritten as

$$\frac{\partial^2 A}{\partial z^2} + \frac{\varepsilon_{xx}}{\varepsilon_{zz}}\frac{\partial^2 A}{\partial x^2} = -\frac{1}{v_A^2}\left(1 - 2\frac{\delta A_z}{A_{0y}} + \frac{\delta n}{n_0}\right)\frac{\partial^2 A}{\partial t^2}. \tag{50}$$

Take the solution of Eq. (50) in form of an envelope as follows:

$$A_+ = A(x, y, z, t)e^{-i(\omega t - k_+ z)}.$$

$$-\frac{2i}{\omega}\frac{\partial A}{\partial t} + 2ik\frac{\partial A}{\partial z} + \frac{\varepsilon_{xx}}{\varepsilon_{0zz}}\frac{\partial^2 A}{\partial x^2} - 2\left(-\frac{x^2}{2} + 0.5\cos ky\right)k^2 A + (\exp(\alpha_0|A|^2 - 1))k^2 A = 0. \tag{51}$$

In the dimensionless form, Eq. (51) will be written in following form:

$$i\frac{\partial A}{\partial t} + i\frac{\partial A}{\partial z} + C\frac{\partial^2 A}{\partial x^2} + \frac{\delta n}{n_0}A + 2\frac{\delta A}{A_{0y}}A = 0. \tag{52}$$

Taking the solution of Eq. (52) as given by Eq. (30) and proceeding similarly as done in case I one may get

$$A_0^2(x, y, z) = \frac{E_{00}^2(x, z)}{f_1}\exp\left(-\frac{x^2}{r_{01}^2 f_1^2}\right),$$

15 January 2025 10:19:41


© Author(s) 2025

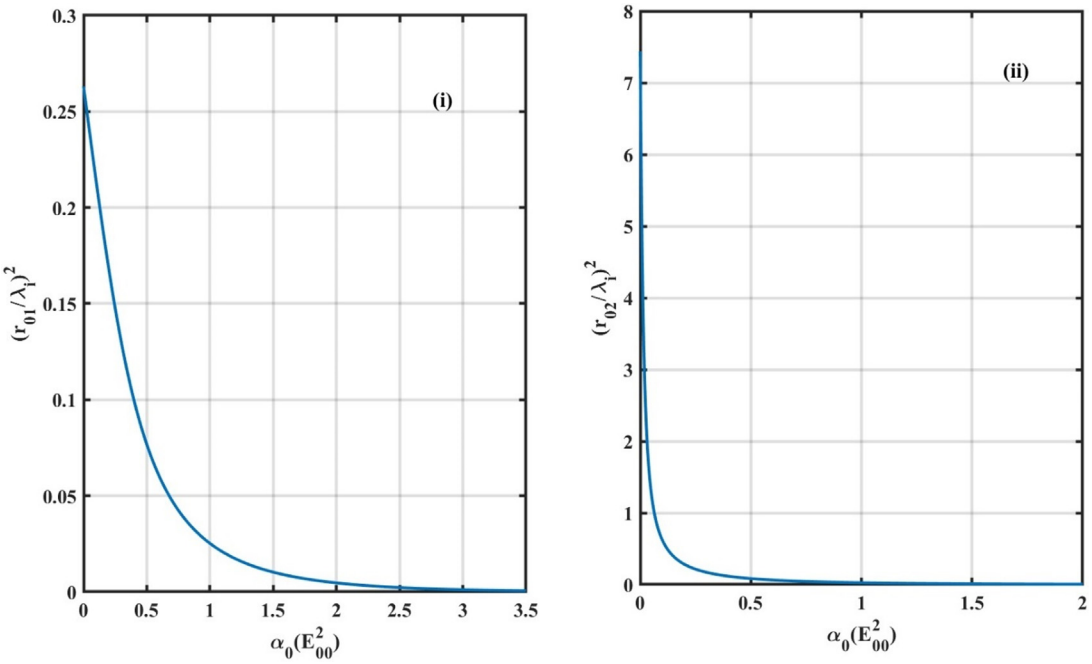



**FIG. 5.** (i) Variation of transverse scale size $(r_{01}/\lambda_i)^2$ with field strength $\alpha_0|E_{00}|^2$ (ii) Variation of transverse scale size $(r_{02}/\lambda_i)^2$ with field strength $\alpha_0|E_{00}|^2$.

$$S(x, y, z) = \beta_1(z)\frac{x^2}{2} + \phi(z).$$

Here, $f_1$ is identical as defined above in Eq. (33). By solving these equations within paraxial limits, the localized region will not be obtained. Liu *et al.*[74] have shown localization with hall effect and Sweet-Parker reconnection without hall effect. We have illustrated it in a different way to get physical insight.

## V. SUMMARY AND CONCLUSION

In the present paper, we studied the localized structures and current sheets in the presence of

(i) a fully developed chain of magnetic islands
(ii) modification in background density in solar wind at 1AU.

The dynamical equation of the AIC wave has been established using two fluid model. The nonlinearity in system arises due to the nonlinear interaction between AIC wave and IAW. Simulation outcomes of AIC wave reveal that initially the localization structures are highly coherent but as the time evolves these structures becomes random/chaotic. We have also studied the power spectra at different times. The conclusion of simulation results can be summarized as below:

i) The numerical simulations outcomes show that the AIC wave gets localized with increasing amplitude. The localization of vector potential may provide a lane for the cascading of energy from high to low scale.
ii) From Figs. 1 and 2, it is concluded that localization is weak in the case when only ponderomotive nonlinearity is considered. When ponderomotive nonlinearity and field perturbation (magnetic islands) both are considered, then localization become relatively intense, indicating relatively strong turbulence. Furthermore, current sheets of the order of sub ion scale are formed and grow into turbulent state with evolution of time.
iii) The power spectra illustrate turbulence, showing an approximate scaling of -5/3 in inertial range (i.e., in large length scale) and -8/3 in steeper range (i.e., in small length scale). It shows the energy distribution over different length scales.
iv) By using semi-analytical model, we have calculated the transverse scale size of localized structures, which is of the order of ion inertial length and greater than the system resolution size. We also observed that the transverse scale size of coherent structures is modified due to the presence of magnetic islands and background density perturbation.

In conclusion, it is found that as the nonlinearity in the system increases, we get more random current sheets or localized structures. As suggested by some observations, the formation of the localized structures or current sheets reveals the turbulence in system. Furthermore, current sheets formation is affected by hall term also. In the studied parameter regime, kinetic effects have not been taken into account and in collisionless plasma, kinetic effects may be essential for phenomena, such as particle acceleration, turbulence, and magnetic reconnection. The kinetic models characterize plasma at the particle level, including individual particle behavior, velocity distribution, and non-thermal effects. Therefore, the authors are planning to study in this direction in their future work for studying turbulence, magnetic islands, and energy dissipation in more detail.

15 January 2025 10:19:41


© Author(s) 2025

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

### Author Contributions

**Ritu:** Formal analysis (equal); Writing – original draft (equal); Writing – review & editing (equal). **Jyoti:** Validation (equal); Visualization (equal); Writing – review & editing (equal). **Suresh C. Sharma:** Conceptualization (equal); Resources (equal); Software (equal). **R. P. Sharma:** Conceptualization (equal); Data curation (equal).

## DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## REFERENCES


[1]R. Bruno and D. Telloni, "Spectral analysis of magnetic fluctuations at proton scales from fast to slow solar wind," Astrophys. J. Lett. **811**(2), L17 (2015).
[2]D. Verscharen, K. G. Klein, and B. A. Maruca, "The multi-scale nature of the solar wind," Living Rev. Sol. Phys. **16**(1), 5 (2019).
[3]W. H. Matthaeus, M. Wan, S. Servidio, A. Greco, K. T. Osman, S. Oughton, and P. Dmitruk, "Intermittency, nonlinear dynamics and dissipation in the solar wind and astrophysical plasmas," Philos. Trans. R. Soc. A **373**(2041), 20140154 (2015).
[4]L. Chen and F. Zonca, "On nonlinear physics of shear Alfvén waves," Phys Plasmas **20**(5), 055402 (2013).
[5]Y. Narita, S. P. Gary, S. Saito, K. Glassmeier, and U. Motschmann, "Dispersion relation analysis of solar wind turbulence," Geophys. Res. Lett. **38**(5), L05101, https://doi.org/10.1029/2010GL046588 (2011).
[6]P. C. Liewer, M. Velli, and B. E. Goldstein, "Alfvén wave propagation and ion cyclotron interactions in the expanding solar wind: One-dimensional hybrid simulations," J. Geophys. Res. **106**(A12), 29261–29281, https://doi.org/10.1029/2001JA000086 (2001).
[7]H. Huang, Y. Yu, L. Dai, and T. Wang, "Kinetic Alfvén waves excited in two-dimensional magnetic reconnection," JGR. Space Phys. **123**(8), 6655–6669 (2018).
[8]N. F. Cramer, *The Physics of Alfvén Waves* (John Wiley & Sons, 2011).
[9]J. W. Belcher and L. Davis, Jr., "Large-amplitude Alfvén waves in the interplanetary medium, 2," J. Geophys. Res. **76**(16), 3534–3563, https://doi.org/10.1029/JA076i016p03534 (1971).
[10]W. Gekelman, "Review of laboratory experiments on Alfvén waves and their relationship to space observations," J. Geophys. Res. **104**(A7), 14417–14435, https://doi.org/10.1029/98JA00161 (1999).
[11]A. Hasegawa and C. Uberoi, "The Alfvén wave: Prepared for the Office of Fusion Energy, Office of Energy Research, US Department of Energy," (Technical Information Center, US Department of Energy, 1982).
[12]R. Meyrand and S. Galtier, "Spontaneous chiral symmetry breaking of Hall magnetohydrodynamic turbulence," Phys. Rev. Lett. **109**(19), 194501 (2012).
[13]V. Krishan and S. M. Mahajan, "Magnetic fluctuations and Hall magnetohydrodynamic turbulence in the solar wind," J. Geophys. Res. **109**(A11), A11105, https://doi.org/10.1029/2004JA010496 (2004).
[14]D. Biskamp, E. Schwarz, and J. F. Drake, "Two-fluid theory of collisionless magnetic reconnection," Phys Plasmas **4**(4), 1002–1009 (1997).
[15]M. Yamada, Y. Ren, H. Ji, J. Breslau, S. Gerhardt, R. Kulsrud, and A. Kuritsyn, "Experimental study of two-fluid effects on magnetic reconnection in a laboratory plasma with variable collisionality," Phys Plasmas **13**(5), 052119 (2006).
[16]Y.-H. Liu, W. Daughton, H. Karimabadi, H. Li, and S. P. Gary, "Do dispersive waves play a role in collisionless magnetic reconnection?," Phys Plasmas **21**(2), 022113 (2014).
[17]B. N. Rogers, R. E. Denton, J. F. Drake, and M. A. Shay, "Role of dispersive waves in collisionless magnetic reconnection," Phys. Rev. Lett. **87**(19), 195004 (2001).
[18]K. T. Osman, W. H. Matthaeus, J. T. Gosling, A. Greco, S. Servidio, B. Hnat, S. C. Chapman, and T. D. Phan, "Magnetic reconnection and intermittent turbulence in the solar wind," Phys. Rev. Lett. **112**(21), 215002 (2014).
[19]S. S. Cerri and F. Califano, "Reconnection and small-scale fields in 2D-3V hybrid-kinetic driven turbulence simulations," New J. Phys. **19**(2), 025007 (2017).
[20]J. Jyoti, S. Sharma, and R. P. Sharma, "Localization of beam generated whistler wave and turbulence generation in reconnection region of magnetopause," Phys Plasmas **31**(2), 022902 (2024).
[21]D. Sundkvist, A. Retinò, A. Vaivads, and S. D. Bale, "Dissipation in turbulent plasma due to reconnection in thin current sheets," Phys. Rev. Lett. **99**(2), 025004 (2007).
[22]K. Schindler, M. Hesse, and J. Birn, "General magnetic reconnection, parallel electric fields, and helicity," J. Geophys. Res. **93**(A6), 5547–5557, https://doi.org/10.1029/JA093iA06p05547 (1988).
[23]V. M. Vasyliunas, "Theoretical models of magnetic field line merging," Rev. Geophys. **13**(1), 303–336, https://doi.org/10.1029/RG013i001p00303 (1975).
[24]F. Guo, X. Li, H. Li, W. Daughton, B. Zhang, N. Lloyd-Ronning, Y.-H. Liu, H. Zhang, and W. Deng, "Efficient production of high-energy nonthermal particles during magnetic reconnection in a magnetically dominated ion–electron plasma," Astrophys. J. Lett. **818**(1), L9 (2016).
[25]J. T. Gosling, "Magnetic reconnection in the solar wind," Space Sci. Rev. **172**, 187–200 (2012).
[26]L. Comisso and L. Sironi, "The interplay of magnetically dominated turbulence and magnetic reconnection in producing nonthermal particles," Astrophys. J. **886**(2), 122 (2019).
[27]S. Eriksson, D. L. Newman, G. Lapenta, and V. Angelopoulos, "On the signatures of magnetic islands and multiple X-lines in the solar wind as observed by ARTEMIS and WIND," Plasma Phys. Controlled Fusion **56**(6), 064008 (2014).
[28]O. V. Khabarova, G. P. Zank, G. Li, J. A. le Roux, G. M. Webb, O. E. Malandraki, and V. V. Zharkova, "Dynamical small-scale magnetic islands as a source of local acceleration of particles in the solar wind," J. Phys. Conf. Ser. **642**(1), 012033 (2015).
[29]C. Huang, Q. Lu, S. Lu, P. Wang, and S. Wang, "The effect of a guide field on the structures of magnetic islands formed during multiple X line reconnections: Two-dimensional particle-in-cell simulations," JGR. Space Phys. **119**(2), 798–807 (2014).
[30]E. G. Zweibel and M. Yamada, "Magnetic reconnection in astrophysical and laboratory plasmas," Annu. Rev. Astron. Astrophys. **47**(1), 291–332 (2009).
[31]A. Ishizawa and N. Nakajima, "Turbulence driven magnetic reconnection causing long-wavelength magnetic islands," Phys Plasmas **17**(7), 072308 (2010).
[32]S. Servidio, W. H. Matthaeus, M. A. Shay, P. A. Cassak, and P. Dmitruk, "Magnetic reconnection in two-dimensional magnetohydrodynamic turbulence," Phys. Rev. Lett. **102**(11), 115003 (2009).
[33]M. J. Choi, "Interaction between a magnetic island and turbulence," Rev. Mod. Plasma Phys. **5**(1), 9 (2021).
[34]N. F. Loureiro and S. Boldyrev, "Role of magnetic reconnection in magnetohydrodynamic turbulence," Phys. Rev. Lett. **118**(24), 245101 (2017).
[35]S. Servidio, W. H. Matthaeus, M. A. Shay, P. Dmitruk, P. A. Cassak, and M. Wan, "Statistics of magnetic reconnection in two-dimensional magnetohydrodynamic turbulence," Phys Plasmas **17**(3), 032315 (2010).
[36]P. Dmitruk, W. H. Matthaeus, and N. Seenu, "Test particle energization by current sheets and nonuniform fields in magnetohydrodynamic turbulence," Astrophys. J. **617**(1), 667 (2004).
[37]J. T. Gosling, R. M. Skoug, D. J. McComas, and C. W. Smith, "Direct evidence for magnetic reconnection in the solar wind near 1 AU," J. Geophys. Res. **110**(A1), A11105, https://doi.org/10.1029/2004JA010809 (2005).
[38]A. A. Van Ballegooijen and M. Asgari-Targhi, "Heating and acceleration of the fast solar wind by Alfvén wave turbulence," Astrophys. J. **821**(2), 106 (2016).
[39]A. Verdini and M. Velli, "Alfvén waves and turbulence in the solar atmosphere and solar wind," Astrophys. J. **662**(1), 669 (2007).
[40]J. L. Verniero, D. E. Larson, R. Livi, A. Rahmati, M. D. McManus, P. S. Pyakurel, K. G. Klein, T. A. Bowen, J. W. Bonnell, and B. L. Alterman, "Parker solar probe observations of proton beams simultaneous with ion-scale waves," Astrophys. J. **248**(1), 5 (2020).


15 January 2025 10:19:41


© Author(s) 2025

[41]D. Perrone, R. Bruno, R. D'amicis, D. Telloni, R. D. Marco, M. Stangalini, S. Perri, O. Pezzi, O. Alexandrova, and S. D. Bale, "Coherent events at ion scales in the inner heliosphere: Parker solar probe observations during the first encounter," Astrophys. J. **905**(2), 142 (2020).
[42]D. Duan, J. He, T. A. Bowen, L. D. Woodham, T. Wang, C. H. K. Chen, A. Mallet, and S. D. Bale, "Anisotropy of solar wind turbulence in the inner heliosphere at kinetic scales: PSP observations," Astrophys. J. Lett. **915**(1), L8 (2021).
[43]C. Shi, M. Velli, R. Lionello, N. Sioulas, Z. Huang, J. S. Halekas, A. Tenerani, V. Réville, J.-B. Dakeyo, and M. Maksimović, "Proton and electron temperatures in the solar wind and their correlations with the solar wind speed," Astrophys. J. **944**(1), 82 (2023).
[44]O. Dhamane, V. Pawaskar, A. Raghav, Z. Shaikh, R. D'Amicis, K. Ghag, K. Kumbhar, D. Telloni, G. Nicolaou, and P. Tari, "Observation of Alfvén ion cyclotron waves in ICME magnetic clouds at 1 au," Astrophys. J. **957**(1), 38 (2023).
[45]L. Ofman, S. A. Boardsen, L. K. Jian, P. Mostafavi, J. L. Verniero, R. Livi, M. McManus, A. Rahmati, D. Larson, and M. L. Stevens, "Observations and modeling of unstable proton and α particle velocity distributions in Sub-Alfvénic solar wind at parker solar probe perihelia," Astrophys. J. **954**(2), 109 (2023).
[46]L.-L. Zhao, G. P. Zank, M. Nakanotani, and L. Adhikari, "Observations of waves and structures by frequency–wavenumber spectrum in solar wind turbulence," Astrophys. J. **944**(1), 98 (2023).
[47]D. Vech, M. M. Martinović, K. G. Klein, D. M. Malaspina, T. A. Bowen, J. L. Verniero, K. Paulson, T. D. de Wit, J. C. Kasper, and J. Huang, "Wave-particle energy transfer directly observed in an ion cyclotron wave," Astron. Astrophys. **650**, A10 (2021).
[48]T. A. Bowen, S. D. Bale, B. D. G. Chandran, A. Chasapis, C. H. K. Chen, T. Dudok de Wit, A. Mallet, R. Meyrand, and J. Squire, "Mediation of collisionless turbulent dissipation through cyclotron resonance," Nat. Astron. **8**, 482 (2024).
[49]G. G. Howes, "A dynamical model of plasma turbulence in the solar wind," Phil. Trans. R. Soc. A **373**(2041), 20140145 (2015).
[50]V. Zhdankin, D. A. Uzdensky, J. C. Perez, and S. Boldyrev, "Statistical analysis of current sheets in three-dimensional magnetohydrodynamic turbulence," Astrophys. J. **771**(2), 124 (2013).
[51]A. Lazarian, G. L. Eyink, A. Jafari, G. Kowal, H. Li, S. Xu, and E. T. Vishniac, "3D turbulent reconnection: Theory, tests, and astrophysical implications," Phys Plasmas **27**(1), 012305 (2020).
[52]S. Perri, M. L. Goldstein, J. C. Dorelli, and F. Sahraoui, "Detection of small-scale structures in the dissipation regime of solar-wind turbulence," Phys. Rev. Lett. **109**(19), 191101 (2012).
[53]P. Wu, S. Perri, K. Osman, M. Wan, W. H. Matthaeus, M. A. Shay, M. L. Goldstein, H. Karimabadi, and S. Chapman, "Intermittent heating in solar wind and kinetic simulations," Astrophys. J. **763**(2), L30 (2013).
[54]M. Nakanotani, G. P. Zank, and L. Zhao, "Particle acceleration in an MHD-scale system of multiple current sheets," Front. Astron. Space Sci. **9**, 954040 (2022).
[55]S. Lion, O. Alexandrova, and A. Zaslavsky, "Coherent events and spectral shape at ion kinetic scales in the fast solar wind turbulence," Astrophys. J. **824**(1), 47 (2016).
[56]D. Smith, S. Ghosh, P. Dmitruk, and W. H. Matthaeus, "Hall and turbulence effects on magnetic reconnection," Geophys Res Lett **31**(2), L02805, https://doi.org/10.1029/2003GL018689 (2004).
[57]C. Dong, L. Wang, Y.-M. Huang, L. Comisso, T. A. Sandstrom, and A. Bhattacharjee, "Reconnection-driven energy cascade in magnetohydrodynamic turbulence," Sci. Adv. **8**(49), eabn7627, (2022).
[58]L. Comisso and L. Sironi, "Particle acceleration in relativistic plasma turbulence," Phys. Rev. Lett. **121**(25), 255101 (2018).
[59]J. He, E. Marsch, C. Tu, S. Yao, and H. Tian, "Possible evidence of Alfvén-cyclotron waves in the angle distribution of magnetic helicity of solar wind turbulence," Astrophys. J. **731**(2), 85 (2011).
[60]J. Ambrosiano and S. H. Brecht, "A simulation study of the Alfvén ion-cyclotron instability in high-beta plasmas," Phys Fluids **30**(1), 108–114 (1987).
[61]N. F. Loureiro, A. A. Schekochihin, and S. C. Cowley, "Instability of current sheets and formation of plasmoid chains," Phys Plasmas **14**(10), 100703 (2007).
[62]T. H. Stix, *Waves in Plasmas AIP* (AIP, 1992).
[63]H. Breuillard, E. Yordanova, A. Vaivads, and O. Alexandrova, "The effects of kinetic instabilities on small-scale turbulence in Earth's magnetosheath," Astrophys. J. **829**(1), 54 (2016).
[64]M. Heuer and E. Marsch, "Diffusion plateaus in the velocity distributions of fast solar wind protons," J. Geophys. Res. **112**(A3), A03102, https://doi.org/10.1029/2006JA011979 (2007).
[65]P. K. Shukla and L. Stenflo, "Nonlinear propagation of electromagnetic ion-cyclotron Alfvén waves," Phys. Fluids **28**(5), 1576–1578 (1985).
[66]M. S. Sodha, A. K. Ghatak, and V. K. Tripathi, "V self focusing of laser beams in plasmas and semiconductors," in *Progress in Optics* (Elsevier, 1976), pp. 169–265.
[67]R. Fitzpatrick and F. L. Waelbroeck, "Two-fluid magnetic island dynamics in slab geometry. II. Islands interacting with resistive walls or resonant magnetic perturbations," Phys Plasmas **12**(2), 022308 (2005).
[68]O. Alexandrova, V. Carbone, P. Veltri, and L. Sorriso-Valvo, "Small-scale energy cascade of the solar wind turbulence," Astrophys. J. **674**(2), 1153 (2008).
[69]J. Gazdag, "Time-differencing schemes and transform methods," J. Comput. Phys. **20**(2), 196 (1976).
[70]H. D. Singh and R. P. Sharma, "Generation of coherent wave packets of kinetic Alfvén waves in solar plasmas," Phys Plasmas **14**(10), 102304 (2007).
[71]N. H. Bian and P. K. Browning, "Particle acceleration in a model of a turbulent reconnecting plasma: A fractional diffusion approach," Astrophys. J. **687**(2), L111 (2008).
[72]W. Rozmus, R. P. Sharma, J. C. Samson, and W. Tighe, "Nonlinear evolution of stimulated Raman scattering in homogeneous plasmas," Phys Fluids **30**(7), 2181–2193 (1987).
[73]S. A. Akhmanov, A. P. Sukhorukov, and R. V. Khokhlov, "Self-focusing and diffraction of light in a nonlinear medium," Sov. Phys. Usp. **10**(5), 609 (1968).
[74]Y.-H. Liu, P. Cassak, X. Li, M. Hesse, S.-C. Lin, and K. Genestreti, "First-principles theory of the rate of magnetic reconnection in magnetospheric and solar plasmas," Commun. Phys. **5**(1), 97 (2022).

15 January 2025 10:19:41


© Author(s) 2025

2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS) | 979-8-3503-8749-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICTACS62700.2024.10840499

# Object Shape Classification Using 4x4 Ultrasonic Sensor Array

Sambhav Jain[1], Kuldeep Singh[2], Qasim Murtaza[3]
[1,3]*Department of Mechanical Engineering, Delhi Technological University, Delhi–110042, India*
[2]*Combat Vehicles Research and Development Establishment, DRDO, Chennai–600054, India*
*E-mail: [1]sambhavjain_pe22b14_41@dtu.ac.in, [2]kuldeep.cvrde@gov.in, [3]qasimmurtaza@dce.ac.in*

***Abstract*—This research paper presents a comprehensive study on an ultrasonic sensor-based system designed for real-time object shape recognition using machine learning (ML) algorithms grounded in echolocation principles. Echolocation, a sound-based sensory system employed by animals such as bats and dolphins, enables precise identification of object shape, location and size through the interpretation of emitted and received echoes. This paper details the design and implementation of the sensor array system and evaluates the performance of various deep learning and machine learning algorithms for this application. The machine learning algorithm XGBoost classified object shapes with an accuracy of 99.1%.**

***Keywords: Object Shape Recognition, Machine Learning, Sensor Array, Ultrasonic Sensor***

## I. Introduction

Sensors of all kinds have proven to be instrumental in the field of robotics. They are what a robot uses to take input from the environment around it. Sensors convert various physical quantities, such as distance, temperature, and humidity, into electrical pulses, which are then transmitted to a control system [1]. Ultrasonic sensors are sound based. They send out a burst of sound waves from their transmitter at a specific frequency which is above that of the listening frequency of human beings. This sound wave reflects back from the nearest object in its path, which is then caught by the receiver [2]. The sensor reports back the time of flight (TOF), from when a sound wave is emitted, to when it is received [3].

A sensor array is a spatial arrangement of multiple sensors used to gather data. Ultrasonic sensor-based arrays have been utilized in various configurations to collect and process TOF data [4], [5], [6], [7], [8], [9]. A two-dimensional sensor array with a total of 16 HC-SR04 sensors mounted in a 4x4 square-shaped arrangement is proposed in this paper. The sensors are spatially placed in a manner to minimize the interference of their respective ultrasonic bursts with each other.

As mentioned before, a sensor produces an electrical pulse based on the physical parameters in its surroundings. The pulse is sent to a controlling device like a microcontroller for processing. An Arduino is used to collect data from the 16 sensors. The multiple distance readings arranged in a table are used to train machine learning (ML) models for classification of different shapes. The study conducted in this paper uses two shapes-triangle and circle, hence performing binary classification.

The methods and algorithms discussed in the paper are applicable to a broad range of use cases. Ultrasonic sensors have already been implemented in vehicles to provide assistance in tasks like parking [11], [12] as well as autonomous driving [13]. In the field of robotics, [14] especially in drones, to detect, recognize and avoid an object is a vital operation that has been carried out with the assistance of ultrasonic sensors and artificial intelligence [15], [16]. Ultrasonic sensor arrays offer a well-documented, cost-effective approach to developing assistive technologies for the blind [17], [18], [19], [20]. The proposed system design and the high accuracy object shape classifier significantly enhance the applicability of ultrasonic sensors across the mentioned diverse fields.

## II. Methods and Materials

The complete setup consists of two major components – the Ultrasonic Sensor Array and the Control System.

### *A. Ultrasonic Sensor Array*

As previously mentioned, HC-SR04 was used to build the sensor array in this study. It has a transmitter, receiver and auxiliary nodes for the functioning of the module. HC-SR04 can be considered as a rangefinder. The sound waves produced are at a frequency of 40 kHz, which is twice that of the maximum frequency a human ear can normally hear (20 kHz). Hence, this sound wave is inaudible to humans. The sound wave reflects back from the nearest object in range and returns to the receiver. The sensor reports the TOF back [21]. The sensors emit the sound wave in a conical manner with an angle of 15°. This process is shown in Fig. 1. (a)

HC-SR04 has four pins which connect it to the microcontroller:

- Vcc – Voltage Supply (+5V)
- Gnd – Ground
- Trig – Sends Ultrasonic Burst
- Echo – Takes TOF Measurement The sensor, along with the 4 pins, is shown in Fig. 1 (b). The

Copyright © IEEE–2024 ISBN: 979-8-3503-8749-0


Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:59:36 UTC from IEEE Xplore. Restrictions apply.

Vcc and Gnd pins power the sensor by providing a power supply of 5V and completing the circuit. The Trig pin is on (or set to HIGH) for 10 µs. After 10 µs, the sensor sends out an ultrasonic burst for 8 complete cycles at 40 kHz. After the sonic burst is complete, the Echo pin is turned on (or set to HIGH). It stays at HIGH until the receiver detects the returning of the sound waves. At that point, the pin is set to LOW. The timing diagram is shown in Fig. 2. The time for which the Echo pin is set to HIGH is the TOF. This time is the output of the sensor.

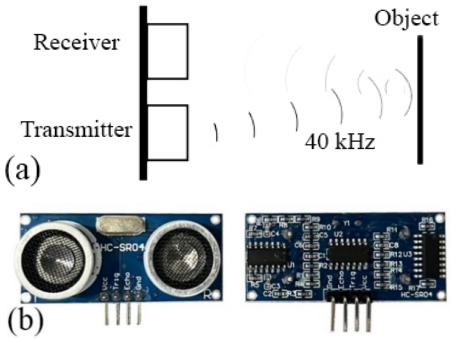



Fig. 1 (a) Ultrasonic Burst Being Emitted from Transmitter and Returning to Receiver. (b) Front and Back Side of HC-SR04 Sensor Module

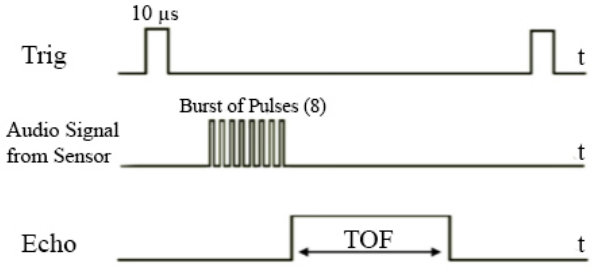



Fig. 2: HC-SR04 Trig and Echo Timing Diagram

For this study, the sensor array had 16 sensors set in 4 rows and 4 columns. The base which holds the array was made of a standard MDF board of thickness 4 mm. The length and breadth of this board were 75 cm x 70 cm. The sensors were placed in a square spatial arrangement. The distance between the left side of the first sensor and the right side of the last sensor in a row was 50 cm as shown in Fig. 3 (a). Similarly, the distance between the top of the first sensor and the bottom of the last sensor in a column was also 50 cm. The trapezium stands were also made of the same sheet of MDF. The dimensions of the stand are shown in Fig. 3 (b). A cut of 4 cm x 0.40 cm was made in the center of the smaller side of the trapezium and at the bottom of the rectangular board as well. This created an interlocking joint. The relative descent of the board was 8 cm.

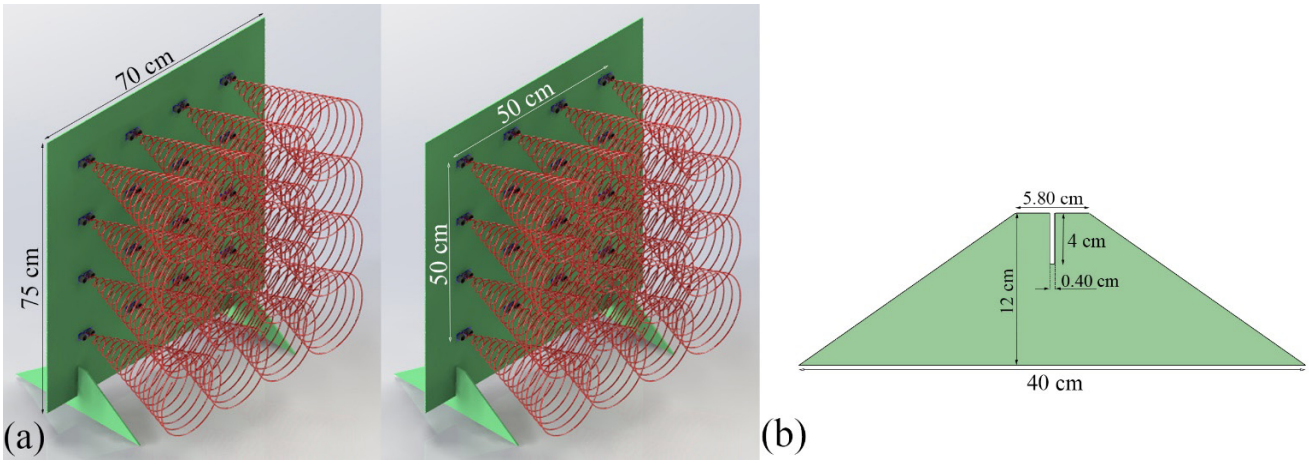



Fig. 3: (a) Dimensions of Sensor Array Board and Placement of Sensors (b) Dimensions of the Trapezium Stand

## *B. Control System*

A Control System is required to control the data collection and classification process. An Arduino was an ideal choice, given the specific components involved. An Arduino Mega 2560 was used to collect data from the sensor array after an interval of every 1 second. Arduino Mega offers 54 digital input/output pins which were suitable to collect data from 16 ultrasonic sensors requiring 2 connections (Trig and Echo) each. The calculation of distance from TOF occurred in Arduino Mega itself before it sent the values further.

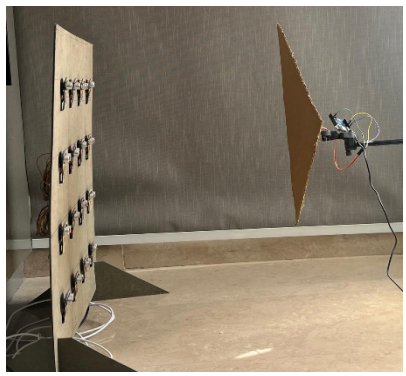


Fig. 4: Data Collection Process Using a Triangle Cutout


Copyright © IEEE–2024 ISBN: 979-8-3503-8749-0

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:59:36 UTC from IEEE Xplore. Restrictions apply.

The Arduino Mega was connected to Python running on Windows operating system. The data was collected through the means of serial communication. Pyserial library in Python was used to facilitate the same. Once the set of 16 distance values reached Python, they were classified into a triangle or circle using an already trained ML model.

## III. Data Collection and Machine Learning

The data collected to train different ML models was done with two shapes – triangle and circle. The readings from the array were initialized in a manner that the readings outside 70 cm range were reported as 0. In 2024, Mexias et al. [22] published a comparative study which suggested that the HC-SR04 sensor functions accurately in the 0-100 cm range, albeit having a maximum range of 400 cm. The distance data collected from triangles was labelled as 1, whereas data collected from circles was labelled as 0. The dataset contained 17 columns, distance measurements from 16 sensors and 1 column containing class. A total of 3101 rows were present in the data. Hence, the dimension of the dataset was 3101 × 17. Four different shapes of triangles and three diametrically varying circles were taken to collect the data. The shapes were rotated slowly at a constant speed using the MG995 servo motor while collecting the data. The servo motor was mounted to a tripod hanging perpendicular to the sensor array, as shown in Fig. 4. Distance between the array and the shapes, and the vertical distance between floor and the tripod was also varied. The ratio between the amount of data for classes 1 and 0 was approximately 3:2. Number of rows for class 0 was 1259 and for class 1 was 1842. Distribution of different types of triangles and circles in the dataset is shown in the pie-chart in Fig. 5. The nomenclature used in marking the sensors is given in Fig. 6.

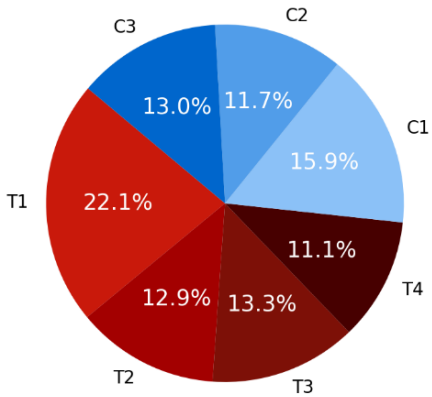



Fig. 5: Pie-chart Distribution of Different Dimensions of Shapes in the Dataset. T1-T4 are the Triangles and C1-C3 are the Circles

To classify the objects, present in front of the array, different ML models were trained on the data collected. XGBoost (XGB), Random Forest (RF) Classifier, Multilayer Perceptron (MLP) and Support Vector Machines (SVM) were used in the study. A comparative analysis was conducted between these models to determine the model performing most accurately.

### A. XGBoost

Extreme Gradient Boost or XGBoost is a scalable ensemble ML model used to perform tree boosting. Gradient tree boosting gives excellent results on many standard classification benchmarks [23]. It is a framework that trains decision trees sequentially. Each iteration updates the current prediction of the model by training a new decision tree. The aim of the model is to minimize the loss function which measures the model's fit to the data. This is the "boosting" process in which a stronger predictive model is created by combining weak classifiers (Decision Trees). This model performs with higher accuracy comparatively when classifying between triangle and circle on the collected data set. Metrics of the model's performances are given in Table 1.

Table 1: Metrics for Xgboost

| Accuracy | Precision | Recall | F1 Score |
|---|---|---|---|
| 0.991 | 0.989 | 0.997 | 0.993 |

The reason for high performance on the calculated standard metrics was due to the nature of the collected dataset, which was highly compatible with the features offered by this model. Some of these features are explained here. XGB uses max-depth technique to control the complexity of the model which essentially terminates the splitting of the decision trees when additional splits stop providing meaningful gain. Multicollinearity also occurs in the data collected by adjacent sensors. Decision trees only split according to the most suitable feature at each node. Therefore, the multicollinearity of different features did not interfere with the training of the model. Furthermore, XGB includes L1 and L2 regularization, preventing overfitting by reducing the importance of similar sets of readings. XGB employs out-of-core computation enabling processing of large amounts of data. Even if the size of the sensor array was increased, XGB would have performed optimally.

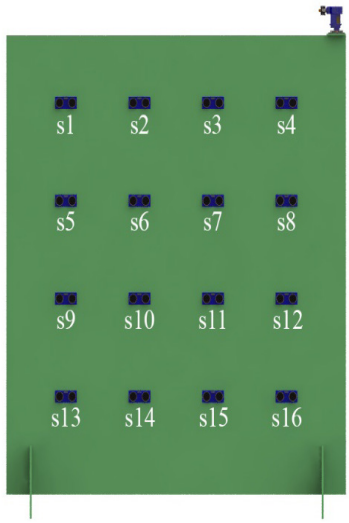



Fig. 6: Naming of the Sensors from s1-s16 when Viewing the Sensor Array from the Front

Copyright © IEEE–2024 ISBN: 979-8-3503-8749-0


Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:59:36 UTC from IEEE Xplore. Restrictions apply.

### B. Random Forest Classifier

RF is an ensemble tree-based model similar to XGB. However, the major difference is that instead of "boosting", it uses "bagging". Bagging or bootstrap aggregation is a method in which weaker or base models (Decision Trees) are trained on different subsets of the initial training dataset and the final prediction is made by combining the result of each base model. The classification is done by majority voting. The class that is predicted by the greatest number of base models is the final prediction [24]. Problem of multicollinearity is also tackled by this model by introducing randomness in the selected datasets. The model used here has a total of 100 individual decision trees. The quality of the splits is evaluated by the Gini impurity criterion. The metrics are given in Table 2.

Table 2: Metrics for Random Forest Classifier

| Accuracy | Precision | Recall | F1 Score |
|---|---|---|---|
| 0.977 | 0.963 | 0.998 | 0.981 |

### C. Multilayer Perceptron

One of the most popular Artificial Neural Networks (ANNs) is MLP. The layers in this neural network can be divided into three - input layer, hidden layers, and output layer. It is a type of feedforward ANN; the data enters through the input layer and the final classification is done by the output layer which consists of a single neuron. The data flows unidirectionally through the hidden layers. To perform binary classification, the neuron in the final layer uses sigmoid activation function [25]. A total of 6 layers including the input and the output layer were present in the MLP model. The number of neurons in each hidden layers were progressively decreasing: 50, 40, 30, 10. Each of the hidden layers had the relu activation function. The metrics for this model are shown in Table 3.

Table 3: Metrics for Multilayer Perceptron

| Accuracy | Precision | Recall | F1 Score |
|---|---|---|---|
| 0.925 | 0.930 | 0.945 | 0.938 |

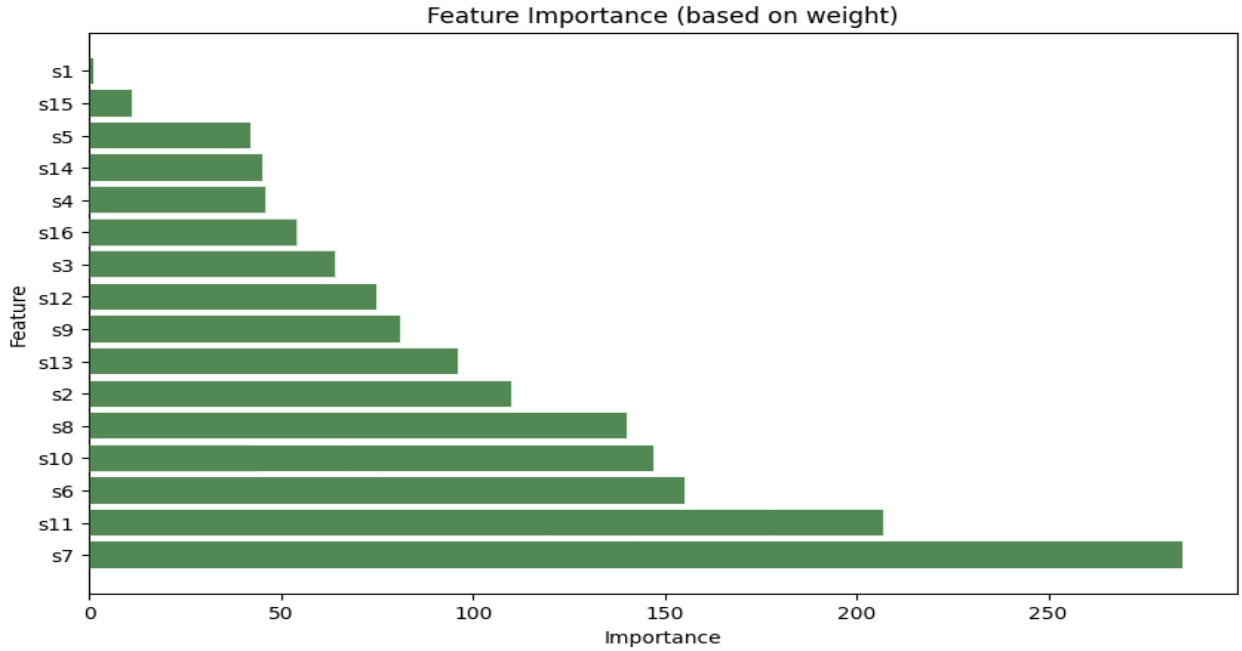



Fig. 7: Feature Importance Calculated from XGBoost, Based on Weight

The MLP model is considered as one of the simplest ANN. It is highly sensitive to noisy data and tends to overfit. L2 regularization has been used in the model to prevent this. Neural Networks classify with a higher accuracy when trained on a large dataset. When trained on a noisy and small dataset, the classification accuracy can be relatively low.

### D. Support Vector Machines

In recent years, SVMs have become widely-used classifiers across both academia and industry. It is informally considered a binary classification algorithm. Input vectors are mapped in a high-dimension feature space. A linear decision surface that separates the patterns found in data is formed in this high dimension space. On the basis of this decision surface, the SVM model generalizes itself on the input data. This decision surface is known as a "hyperplane" [26]. An SVM decision function is essentially an optimal hyperplane that classifies observations into different classes based on patterns of information. Training an SVM decision function involves determining this hyperplane that maximizes the margin between the support vectors of different class labels [27]. The metrics for this model is shown in Table 4.


Copyright © IEEE–2024 ISBN: 979-8-3503-8749-0

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:59:36 UTC from IEEE Xplore. Restrictions apply.

Table 4: Mmetrics for Support Vector Machines

| Accuracy | Precision | Recall | F1 Score |
|---|---|---|---|
| 0.925 | 0.917 | 0.962 | 0.939 |

## IV. Results

The study conducted compares different ML algorithms to determine which algorithm presents the highest accuracy while classifying between the shapes - triangle and circle. XGBoost was the model that performed classification with the highest accuracy of 99.1%. This model performed well on standard metrics as it classified with a precision of 98.9%, recall of 99.7% and F1 score of 99.3%. In 2019, Barua et al. [5] presented a comparative study in which they used a 3x3 ultrasonic sensor array. They classified three shapes (triangle, circle and rectangle), hence, performing multi-class classification. An overall maximum accuracy of 98.7% was achieved using Fitnet, a neural network. This research paper indicates that using a larger sensor array network tends to improve the classification accuracy.

Distance readings from each sensor were a different input feature. XGBoost can calculate the feature importance depending on various parameters. Feature importance based on weight ranks the features according to their frequency and significance in splitting the decision trees. The most important features based on weight are shown in Fig. 7. Sensor 7, 11, 6 and 10 are the four sensors placed in the center of the array. The readings collected by these sensors were recognized by the model as the most appropriate to split the nodes in the decision trees. Data collected by sensors placed on the corners and the bottom edges were given the least importance by the model (Sensor 1, 15, 5 and 14).

It is a standard practice amongst data scientists to split the data into training and testing sets picked on a random basis at the start of the process. The model generalises on the training data set. The final metrics are calculated when the trained model is used to classify the testing dataset. A similar practice has been followed while conducting research for this study. Since only four triangles and three circles were used to create the dataset, the testing and training data is split amongst these seven objects only. A better representation of the model's performance will be when it classifies distance values from the shape dimensions it has not been trained on. Hence, a new dataset of 200 rows was created using a triangle and a circle (100 rows each) which vary in their dimensions from the initial 7 objects. Table 5 shows the accuracy of the different classifiers when classifying data collected from the new dataset.

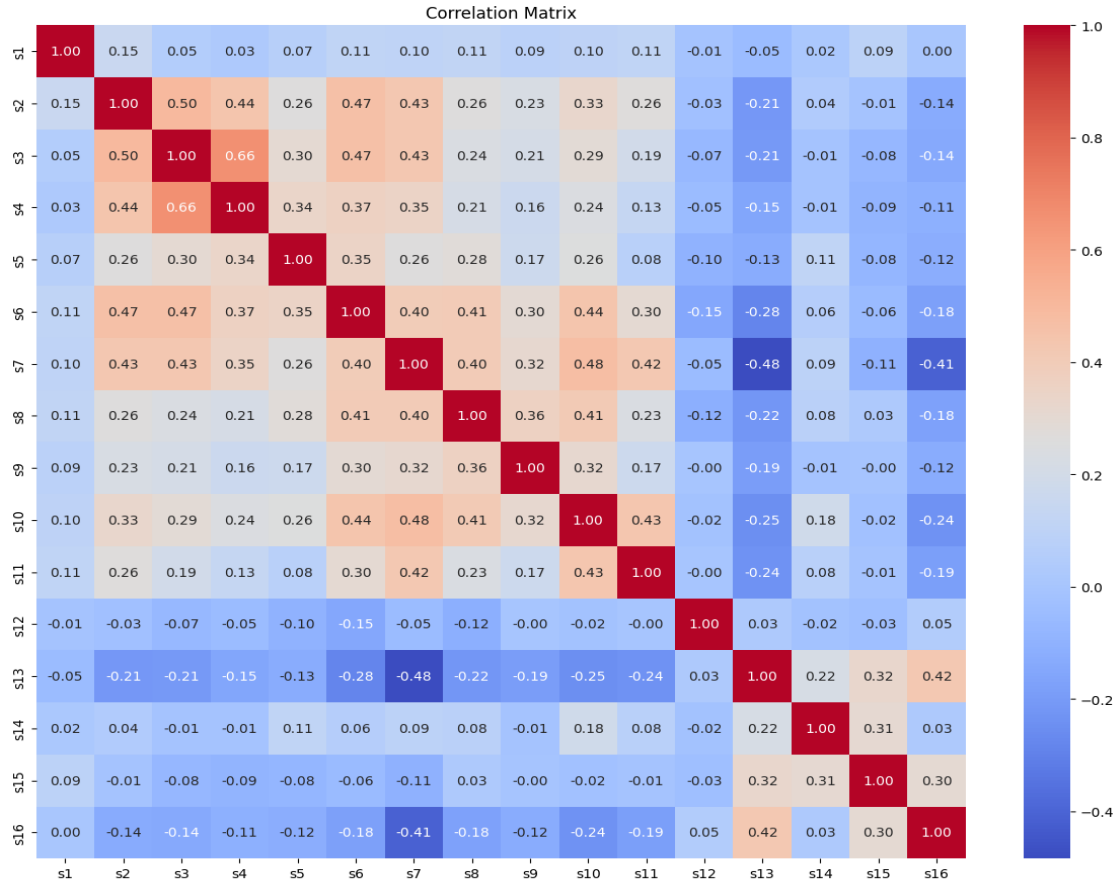



Fig. 8: Correlation Matrix of all Sensor Nodes

Copyright © IEEE–2024 ISBN: 979-8-3503-8749-0


Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:59:36 UTC from IEEE Xplore. Restrictions apply.

Table 5: Accuracy of Trained Models on Dataset Collected from New Shapes

| Model Name | Triangle (100) | Circle (100) |
|---|---|---|
| XGB | 96% | 97% |
| RF | 90% | 96% |
| MLP | 93% | 97% |
| SVM | 89% | 93% |

All the models classify more accurately for circles than for triangles. This is due to the fact that variations among circles are only limited to different diameters. The data collected from the new circle will be very similar to the circle data which the model has been trained on, even though the diameter for this circle is different.

Multicollinearity between different features can be visualised through a correlation matrix. A number closer to 1 shows direct proportionality between features, whereas a number closer to -1 shows inverse proportionality between features. The diagonal elements represent the correlation of each feature with itself. These values are always 1. The correlation matrix is shown in Fig. 8. This plot of correlation computes Pearson correlation coefficient, hence measuring the linear relationship between the features. Both the axes of the matrix represent the 16 sensor nodes, the nomenclature followed is shown in Fig. 6. There was only one instance of multicollinearity being outside the range of , between s4 and s3. The maximum positive (and overall maximum) collinearity was 0.66 (between s4 and s3). The minimum collinearity was between s7 and s13 being -0.48. Improved dataset with an overall lower correlation between features may obtain an even higher classification accuracy.

## V. Conclusion

The methods and algorithms demonstrated in this research paper present a viable alternative to a camera-based object shape recognition and classifying mechanism. A high accuracy of 99.1% was achieved using XGBoost algorithm while classifying between the two objects using a 4x4 ultrasonic sensor array network. The complete system is economically viable, computationally inexpensive and easy to reproduce at larger scales. Using an ultrasonic sensor array to detect and classify objects has military applications and can be efficacious when applied in autonomous vehicles, marine surveillance, robots, drones and aid technologies for the visually impaired.

## References

[1] M. Javaid, A. Haleem, S. Rab, R. P. Singh, and R. Suman, "Sensors for daily life: A review," 2021, doi: 10.1016/j.sintl.2021.100121.

[2] V. Parihar, Y. Rohilla, and K. Kumari, "Ultrasonic Sensor based Smart Cap as Electronic Travel Aid for Blind People," in *Proceedings of the Third International Conference on Smart Systems and Inventive Technology (ICSSIT 2020)*, 2020.

[3] K. A. Raza and W. Monnet, *Moving objects detection and direction-finding with HC-SR04 ultrasonic linear array*. 2019.

[4] S. Davila-Montero, S. Parsnejad, Y. Gtat, and A. J. Mason, "Improving Object Localization Resolution using a Wearable Ultrasonic Sensor Array," *Midwest Symposium on Circuits and Systems*, vol. 2020-August, pp. 237–240, Aug. 2020, doi: 10.1109/MWSCAS48704.2020.9184439.

[5] S. Barua, A. Saha, A. A. Shahjamal Khan, and R. H. Chowdhury, "Comparative Study of Object Shape Recognition using Ultrasonic Sensor Arrays with Artificial Neural Network," in *ICIET 2019 - 2nd International Conference on Innovation in Engineering and Technology*, Institute of Electrical and Electronics Engineers Inc., Dec. 2019. doi: 10.1109/ICIET48527.2019.9290607.

[6] C. Nadee and K. Chamnongthai, "An Ultrasonic-Based Sensor System for Elderly Fall Monitoring in a Smart Room," *J Healthc Eng*, vol. 2022, no. 1, p. 2212020, Jan. 2022, doi: 10.1155/2022/2212020.

[7] A. R. Patkar and P. P. Tasgaonkar, "Object recognition using horizontal array of ultrasonic sensors," *International Conference on Communication and Signal Processing, ICCSP 2016*, pp. 983–986, Nov. 2016, doi: 10.1109/ICCSP.2016.7754294.

[8] C. Nadee and K. Chamnongthai, "Ultrasonic array sensors for monitoring of human fall detection," *ECTI-CON 2015 - 2015 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, Aug. 2015, doi: 10.1109/ECTICON.2015.7207097.

[9] A. Patankar and T. Nikoubin, "Wearable system for obstacle detection and human assistance using ultrasonic sensor array," *ACM International Conference Proceeding Series*, vol. 06-08-July-2016, Jul. 2016, doi: 10.1145/2967878.2967893.

[10] X. Li, H. Sun, W. Gao, Y. Shi, G. Liu, and Y. Wu, "Wind speed and direction measurement based on arc ultrasonic sensor array signal processing algorithm," 2016, doi: 10.1016/j.isatra.2016.07.010.

[11] T.F. Wu, P.S. Tsai, N.T. Hu, and J.Y. Chen, "Research and implementation of auto parking system based on ultrasonic sensors," *Proceedings of the IEEE International Conference on Advanced Materials for Science and Engineering: Innovation, Science and Engineering, IEEE-ICAMSE 2016*, pp. 643–645, Feb. 2017, doi: 10.1109/ICAMSE.2016.7840267.

[12] Y. Shao, P. Chen, and T. Cao, "A grid projection method based on ultrasonic sensor for parking space detection," *International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 2018-July, pp. 3378–3381, Oct. 2018, doi: 10.1109/IGARSS.2018.8519022.

[13] E. N. Budisusila, B. Arifin, S. A. D. Prasetyowati, B. Y. Suprapto, and Z. Nawawi, "Artificial neural network algorithm for autonomous vehicle ultrasonic multi-sensor system," *EECCIS 2020 - 2020 10th Electrical Power, Electronics, Communications, Controls, and Informatics Seminar*, pp. 128–131, Aug. 2020, doi: 10.1109/EECCIS49483.2020.9263459.

[14] J. N. Yasin, S. A. S. Mohamed, M. H. Haghbayan, J. Heikkonen, H. Tenhunen, and J. Plosila, "Low-cost ultrasonic based object detection and collision avoidance method for autonomous robots," *International Journal of Information Technology (Singapore)*, vol. 13, no. 1, pp. 97–107, Feb. 2021, doi: 10.1007/S41870-020-00513-W/FIGURES/13.

[15] S. Ramakrishna Pillutla, N. Neelima, A. Sai Kumar, B. Narasimha Rao, and N. Suresh, "An Efficient two Dimensional Mapping of External Surroundings for Robotics and Drones," *2022 13th International Conference on Computing Communication and Networking Technologies, ICCCNT 2022*, 2022, doi: 10.1109/ICCCNT54827.2022.9984274.

[16] S. Suherman, R. A. Putra, and M. Pinem, "Ultrasonic Sensor Assessment for Obstacle Avoidance in Quadcopter-based Drone System," *MECnIT 2020 - International Conference on Mechanical, Electronics, Computer, and Industrial Technology*, pp. 50–53, Jun. 2020, doi: 10.1109/MECNIT48290.2020.9166607.

 Copyright © IEEE–2024 ISBN: 979-8-3503-8749-0

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:59:36 UTC from IEEE Xplore. Restrictions apply.

[17] S. Karmakar, S. K. Kundu, S. Sengupta, G. Priya, and A. Roy, "Third Eye: Obstacle Avoidance Device for Visually Impaired Persons," *Interdisciplinary Research in Technology and Management*, pp. 619–626, May 2024, doi: 10.1201/9781003430469-73.

[18] K. Goel, M. Nagar, M. Tiwari, A. R. Mishra, and S. S. Chauhan, "Enhanced Ultrasonic Cane for Visually Impaired People," *2024 2nd International Conference on Disruptive Technologies, ICDT 2024*, pp. 500–505, 2024, doi: 10.1109/ICDT61202.2024.10489694.

[19] C. Ramisetti *et al.*, "An Ultrasonic Sensor-based blind stick analysis with instant accident alert for Blind People," *2022 International Conference on Computer Communication and Informatics, ICCCI 2022*, 2022, doi: 10.1109/ICCCI54379.2022.9740786.

[20] M. Yasir, I. N. Lestari, C. Setiawan, Ulfiah, M. R. Effendi, and E. A. Z. Hamidi, "Design and Implementation of the Blind Navigation Aids Using Ultrasonic Sensor," *Proceeding of 2021 7th International Conference on Wireless and Telematics, ICWT 2021*, 2021, doi: 10.1109/ICWT52862.2021.9678455.

[21] V. A. Zhmud, N. O. Kondratiev, K. A. Kuznetsov, V. G. Trubin, and L. V Dimitrov, "Application of ultrasonic sensor for measuring distances in robotics," *IOP Conf. Series: Journal of Physics: Conf. Series*, vol. 1015, p. 32189, 2018, doi: 10.1088/1742-6596/1015/3/032189.

[22] S. Mexias, V. Fotopoulos, and K. Giannakopoulos, "Comparison of HC-SR04 and TF-LC02 Distance Sensors for Indoor Mapping Applications," in *2024 Panhellenic Conference on Electronics and Telecommunications, PACET 2024 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/PACET60398.2024.10497028.

[23] P. Li, "Robust LogitBoost and Adaptive Base Class (ABC) LogitBoost," *Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence, UAI 2010*, pp. 302–311, Mar. 2012, Accessed: Jun. 08, 2024. [Online]. Available: https://arxiv.org/abs/1203.3491v1

[24] G. Zachmann, I. Udo, F. Author, and F. A. Ihle, "Random Forests for Tracking on Ultrasonic Images," 2017.

[25] H. Taud and J. F. Mas, "Multilayer Perceptron (MLP)," in *Geomatic Approaches for Modeling Land Change Scenarios*, M. T. Camacho Olmedo, M. Paegelow, J.-F. Mas, and F. Escobar, Eds., Cham: Springer International Publishing, 2018, pp. 451–455. doi: 10.1007/978-3-319-60801-3_27.

[26] Q. Wang, "Support Vector Machine Algorithm in Machine Learning," *2022 IEEE International Conference on Artificial Intelligence and Computer Applications, ICAICA 2022*, pp. 750–756, 2022, doi: 10.1109/ICAICA54878.2022.9844516.

[27] D. A. Pisner and D. M. Schnyer, "Support vector machine," *Machine Learning: Methods and Applications to Brain Disorders*, pp. 101–121, Jan. 2020, doi: 10.1016/B978-0-12-815739-8.00006-7.

Copyright © IEEE–2024 ISBN: 979-8-3503-8749-0



Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:59:36 UTC from IEEE Xplore. Restrictions apply.

2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS) | 979-8-3503-8749-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICTACS62700.2024.10840803

# PATANLS: Performance Analysis of Trust-Aware Node Localization in WSNs

Yogendra Mohan[1], Rajesh Kumar Yadav[2], Manisha Manjul[3]
[1,2]*Department of Computer Science & Engineering, Delhi Technological University, Delhi, India*
[3]*Department of Computer Science & Engineering, G.B. Pant DSEU Okhala-1, New Delhi, India*
*E-mail: [1]yogendra.mohan@gmail.com, [2]rkyadav@dtu.ac.in, [3]bhavikamanisha@gmail.com*

***Abstract*—Finding the spatial coordinate of any unknown node using anchor nodes is Node Localization (NL). Anchor nodes (ANs) are GPS-enabled and know their locations. A malicious node (MN) may be involved in the localization process. So, Trust in the anchor nodes is essential for secure and reliable localization. A novel Sybil-free (SF) RSSI-based security scheme (PATANLS) is proposed to unravel the issues of Sybil attacks in localization schemes. The proposed scheme requires nominal overhead for the devices and works efficiently based on the received signal. PATANLS employs a trust-based technique to precisely determine the location of target nodes while limiting the influence of compromised or malevolent nodes on localization accuracy. The highly trustworthy ANs are later identified as legitimate nodes for the localization process. The paper addresses the limitations of existing localization approaches in WSNs. To further enhance the localization performance, the Coati Optimization Algorithm (COA) is designed by incorporating appropriate fitness functions. The performance of COA is then evaluated by comparing it with the butterfly optimization algorithm (BOA) and particle swarm optimization (PSO). The simulations are conducted utilizing the COA (Coati Optimization Algorithm), BOA (Butterfly Optimization Algorithm), and PSO (Particle Swarm Optimization). The comparative analysis of the COA, BOA, and PSO reveals that the COA surpasses the other algorithms in terms of node localization, node localization calculation time, and node localization error.**

***Keywords: Node Localization, Sybil Attack, Malicious Node, WSNs***

## I. Introduction

A Sybil attack [1] in WSNs is a security threat where a malicious node fantasizes to be multiple nodes. The attack can interrupt network protocols and, operations such as routing, data aggregation, and resource allocation, identity verification, resource testing, random key pre-distribution, trust management systems, and location verification are prevention and detection techniques. The model being suggested consists of three distinct categories of sensor nodes. The nodes consist of anchor nodes (ANs), non-GPS sensor nodes or unknown node, also referred to as target node (TN), and malicious node (MN) that are haphazardly distributed in two-dimensional area. Not all TNs possess accurate positional data. All sensors have consistent technological properties, particularly regarding the transmission range within their specific categories. All sensors with adequate energy can perform a node localization [2,3] procedure. Evaluating the actions of the nodes is crucial in determining their legitimacy or evil intent. The cryptographic-based and trust-based are the two basic security approaches to measure the trust behaviors of the node in WSNs. The cryptographic-based security mechanism is static and suffer from computational overheads and the trust-based security mechanism [5] is dynamic which mitigates internal threats in the distributed system. The legitimate nodes ($AN_1$, $AN_2$, $AN_3$, $AN_4$, $AN_5$) create a network as shown in Fig.1. The fifteen *TNs* ($TN_1$, *to* $TN_{15}$) are deployed in the sensor field. the malicious nodes (MN) create sybil identities ($SN_1$, $SN_2$, $SN_3$) to deceive various authentic nodes. The Sybil [4] identities can disrupt network operations by interacting with the legitimate nodes.

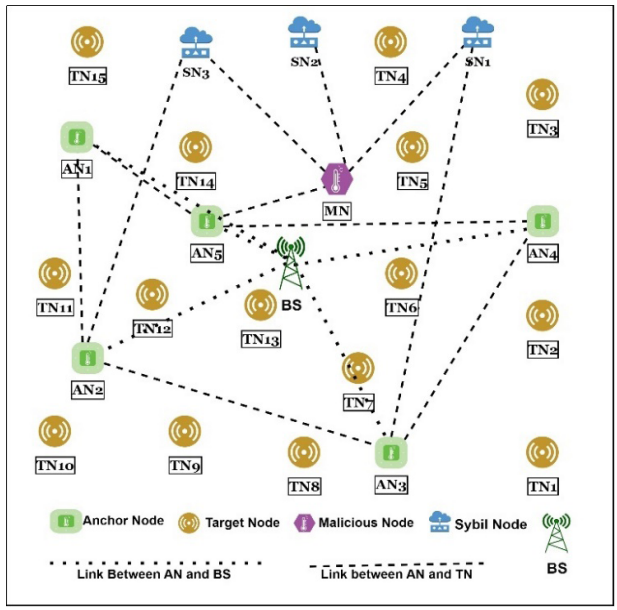



Fig. 1: Illustration of a Sybil attack in the proposed network model a WSNs

*Research Contributions:* Trust-aware[7] localization has been suggested as a remedy for the issues associated with existing localization methods. The suggested scheme has the following main contributions:

Copyright © IEEE–2024 ISBN: 979-8-3503-8749-0


Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 07:02:02 UTC from IEEE Xplore. Restrictions apply.

- The purpose is to allocate a distinct identifier to each anchor node to authenticate its authenticity for the purpose of secure localization.
- To tackle the difficulties presented by a trust-aware node localization strategy, the technique employs a reputation-based trust evaluation model to assess the likelihood of nodes being hostile or compromised. This information is then included in the localization process to enhance accuracy.
- The location of an unidentified node is determined by utilizing data from the most trustworthy anchor nodes.
- To enhance the reliability of localization in WSNs, an efficient solution is proposed, particularly when faced with malicious attacks,
- To apply the COA for further enhancement of the performance parameters.
- To evaluate and compare the proposed scheme (COA_Proposed) with the existing schemes (BOA, PSO).

The rest of the paper is structured as follows: The related work on optimal trust-aware node localization is presented in Section 2. The suggested model, which shows the secure and efficient localization procedure as well as the trust assessment scheme, is presented in Section 3. The evaluation of the results and a comparison of the COA, BOA[6], and PSO are presented in Section 4. Section 5 concludes with suggestions for further development.

## II. Related Work

The authors propose various trust-aware NL schemes for WSNs that considers trustworthy nodes to improve the NL accuracy.

Yali Yuan et al. [8] proposed an SF-APIT technique to address Sybil attacks in APIT localization schemes. The SF-APIT detects and protects distributed wireless localization techniques from Sybil attacks with high detection rates. The communication and storage are cost-effective.

Jingsha He et al. [9] presented a reputation-based security scheme for non-GPS localization to enhance the security and accuracy of localization in hostile environments. The reputation of each beacon is estimated based on a reputation evaluation model. The normal sensor nodes can acquire correct node position information from highly trustworthy beacons to perform localization. The simulation result shows that the proposed scheme can improve the security in the network and accuracy of localization in the hostile environments.

Shi Dong et al. [10] suggested a detection and defensive strategy against Sybil's attack based on the distance vector-hop (DV-Hop) method is described. Experimental findings suggest that the approach may effectively increase the security of localization in WSNs. When the beacon nodes are 50, the proposed model reduces the average localization error by 3% than the DV-Hop.

Tu N. Nguyen et al. [11] proposed a PSO-based enhanced Attack-Resistant Secure Localization Algorithm to minimize independent and collusion attacks. The proposed model demonstrate that our proposed localization scheme performance is better compared to the existing schemes.

Guangjie Han et al. [12] proposed a collaborative model for secure localization based on the trust value. Simulation findings reveal that the proposed method performs better than the compared related works in terms of average localization, location security, and localization ratio.

Rathindra Nath Biswas et al.[13] proposed a resilient localization scheme for WSNs that suits well for localization in hostile environments. the node position is estimated using the AoA method. The beacon nodes estimate the AoA information using the ESPRIT algorithm. Simulation results have a high localization accuracy over the existing schemes.

XIAOYIN LI et al. [14] presented a highly effective and secure method for range-free localization, utilizing outlier elimination and vector refining techniques. A method for eliminating outliers involves filtering out faulty beacon nodes when there are malicious anchor nodes present, to prevent attacks and enhance localization. A mechanism for refining vectors is modelled to enhance localization. The simulation results demonstrate a considerable improvement in localization compared to the DV-hop[14] and vector refinement DV-hop methods, both in the presence and absence of malicious beacon nodes.

G.G. Gebremariam et al.[19] proposed a secure node localization scheme for IoT-based WSNs. The blockchain is used against malicious attacks. The cascaded encryption is used for trust evaluation. Federated machine learning is used for data security and transmission. The malicious node mean detection accuracy of XGBoost and CLK means are 99.95% and 100%.

G.G. Gebremariam et. al[20] proposed hybrid machine-learning models for malicious node detection and secure localization against routing attacks. the mean detection of malicious node accuracy is 100% and the mean localization error is 0.191. the proposed system has improved security and reduced energy consumption.

Rameez Asif et. al[21] proposed a DV-Hop-based robust and secure localization scheme in WSNs. The cryptography algorithm is used for security and robustness. The localization accuracy is improved from 77% to 99% as the anchor nodes increase from 5% to 30% of the target nodes. The localization error decreases from 0.7 to 0.5 when anchor node density increases from 5% to 30%. The localization error is 0.5 when there is no malicious node and increases to 1.5 when malicious nodes are 30% of the target nodes.

Xiao Hai Et. al[22] proposed a Sybil attack-based secure localization scheme that uses a GA-PSO-BP hybrid algorithm. The maximum localization error of the scheme is


Copyright © IEEE–2024 ISBN: 979-8-3503-8749-0

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 07:02:02 UTC from IEEE Xplore. Restrictions apply.

0.23 meters and the root mean square error is 0.72 meters.

Ruchi Garg et. al[23] proposed a Farkas' Lemma-based range-free secure localization against wormhole attacks in WSNs. The localization error is more when the wormhole attack is introduced in the environment field. The localization error decreases when the anchor node density is increased.

G. Hemanth Kumar et. al[24] proposed a secure localization that detects malicious nodes in WSNs and improves end-to-end delay. The model protects against network failures. The proposed Buffalo optimization algorithm enhances the residual energy, and packet forwarding rate, and minimizes packet loss.

The authors have designed blockchains-based and cryptographic algorithm based localization models which suffers from computational and communicational overheads that is a big challenge for sensors because sensors have limited battery. The objective of the proposed scheme is to design a sybil free node localization in WSNs.

## III. Proposed Model

*Procedure for Sybil Detection:* Sybil attack detection is presented in the flowchart as shown in Fig. 3 and detection procedure is as follows:

*Initialization:* The Sybil attack detection aims to identify and mitigate Sybil nodes within the environment field of WSNs using a trust-based approach. Deploy sensor nodes (anchor, target, and malicious nodes) randomly within the network area. Every node in the network discovers the nodes of its neighbors by transmitting hello messages to all other nodes. The trust value of each anchor node is set to 0.5 Jules initially.

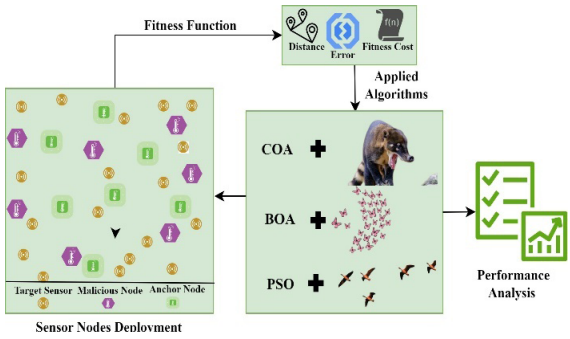



Fig. 2: Network Structure for Trust Aware Localization

*Trust Score Evaluation:* Monitor interactions between nodes over a particular period. Gather statistics on packet transmission, consistency of reported positions, and acknowledgement receiving.

*Trust Metrics:* The trust metrics are based on the interaction behavior of the nodes participating in the localization process. Communication Reliability (CR), Data Consistency (DC), and Behavioral Consistency (BC) are the performance parameters to define the trust metrics. The CR measures the success rate of the packets transmitted and acknowledged.

$$CR_{A,B} = \frac{Successful\ transmissions\ from\ B\ to\ A}{total\ transnmisiion\ from\ B\ to\ A} \quad (1)$$

The DC measures the consistency of node positions over time.

$$DC_{A,B} = 1 - \frac{variation\ in\ reported\ positions\ of\ B}{total\ variations\ observed} \quad (2)$$

The BC measures the consistency of a sensor node's behavior relative to past data.

$$BC_{A,B} = \frac{consistent\ behaviors\ of\ B\ over\ time}{total\ observed\ behaviors} \quad (3)$$

The total trust score ($TT_{A,B}$) is as follows:

$$TT_{A,B} = \omega_1 * CR_{A,B} + \omega_2 * DC_{A,B} + \omega_3 * BC_{A,B} \quad (4)$$

*Detection of Sybil Attack:* To detect the Sybil attack, set a threshold value of the trust ($TT_{th}$) to identify the malicious nodes in the network as follows:

$$TT_{A,B} = \begin{cases} Trusted_B, & if\ TT_{A,B} > TT_{th} \\ Untrusted_B, & Otherwise \end{cases} \quad (5)$$

The trust value of the nodes is decreased by 0.5 and nodes are flagged as a malicious node. Share information about detected malicious nodes with neighboring nodes to enhance security.

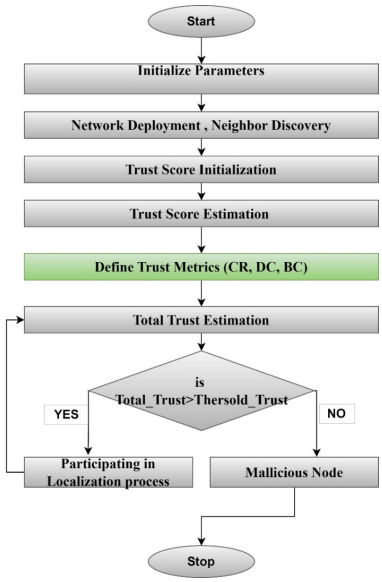



Fig. 3: Flowchart of Trust-Aware Node localization

Suppose a total of 100 packets are sent from node B to node A and 95 packets are successfully received by A then

$$CR_{A,B} = \frac{95}{100} = 0.95$$

This data suggests that 95% of the communications sent from node B to node A were successfully received. Assume that the recorded positions of node B exhibited minor discrepancies, resulting in a variation metric of 0.3. Node B's stated positions have a consistency rate of 70% as shown below.

$$DC_{A,B} = 1 - 0.3 = 0.70$$

Copyright © IEEE–2024 ISBN: 979-8-3503-8749-0


Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 07:02:02 UTC from IEEE Xplore. Restrictions apply.

The behavior consistency (BC): The sensor node B has exhibited consistently in 7 out of 10 observed behaviour is as follows:

$$BC_{A,B} = \frac{7}{10} = 0.70$$

*Trust Score* ($\mathbf{TT}_{\mathbf{AB}}$) *Estimation:* The weights are assigned as follows: $w_1 = 0.5$, $w_2 = 0.3$, *and* $w_3 = 0.2$ representing the significance of communication dependability, data consistency, and behavior consistency, respectively.

$$TT_{A,B} = \omega_1 * CR_{A,B} + \omega_2 * DC_{A,B} + \omega_3 * BC_{A,B}$$
$$TT_{A,B} = 0.50 * 0.95 + 0.30 * 0.70 + 0.20 * 0.70$$
$$TT_{A,B} = 0.475 + 0.210 + 0.140$$
$$TT_{A,B} = 0.825$$

The trust score 0.825, signifies that sensor node B is seen as very trustworthy from node A's perspective. If the value of $T_{AB}$ falls below a predetermined threshold value (0.5), node B could be identified as suspicious, possibly indicating it as a Sybil node.

*Trust-aware Node Localization:* The trust-aware RSSI-based localization [15,16] is modeled for estimating the position of the target nodes. Localization of the target node is an NP-hard problem. The highly trustworthy beacon node deployed randomly as shown in Fig 5 for all six cases. The trilateration method is used to estimate the location of the unknown node with the help of a minimum of 3 trustworthy beacon nodes. The beacon and unknown nodes are deployed randomly in the environment field of area 200x200 square meters. The position of the beacon node (AN) deployed in the environment field are as follows:

$$X_{AN}: x_{i,j} = \{ax_{1,1}, ax_{2,2}, ax_{3,3}, ax_{4,4}, ax_{5,5}, .. ax_{An,An},\} \quad (6)$$

The position randomly deployed unknown nodes which will be discovered by trustworthy beacon nodes are as follows:

$$X_{TN}: x_{i,j} = \{tx_{1,1}, tx_{2,2}, tx_{3,3}, tx_{4,4}, tx_{5,5}, .. tx_{Tn,Tn}\} \quad (7)$$

RSSI is the received signal power at the target node and the power be used to estimate the distance from the beacon node using the Free Space Path Loss model as follows:

$$RSSI_d = RSSI_0 - 10 * n * log_{10}\left(D_{t,r}^{Est}\right) \quad (8)$$

Where $\left(D_{t,r}^{Est}\right)$ is the estimated distance between beacon node and unknown nodes, $RSSI_d$ is the received signal at distance $\left(D_{t,r}^{Est}\right)$ dBm, $RSSI_0$ is signal strength at reference distance which is assumed 1[16].

Table 1: Path-loss component(n) in different environments [16]

| Environment | Path-loss Component(n) |
|---|---|
| Alley | 2.10 to 3.00 |
| Balconies and stairs | 1.40 to 2.40 |
| Corridors | 1.90 to 2.50 |
| Fences | 4.60 to 5.10 |
| Office | 1.40 to 2.50 |
| Parks | 2.70 to 3.40 |

The path loss component(n) is varying between 2 to 5 depends on environment. We chose 2.2 for the simulation. The target nodes send signals to beacon nodes deployed in the environment fields. The beacon nodes receive the signals and send their RSSI values to the target node. The target node estimates the distances using equation-(8) to each beacon nodes. If the coordinates of the at least three non-collinear beacon nodes are $(x_{A1}, y_{A1})$, $(x_{A2}, y_{A2})$, $(x_{A3}, y_{A3})$ , and the coordinate of the target node (TN) is $(x_T, y_T)$, then Euclidian distance from TN to other AN are as follows:

$$\begin{cases} (x_{A1} - x_T)^2 + (y_{A1} - y_T)^2 = {r_1}^2 \\ (x_{A2} - x_T)^2 + (y_{A2} - y_T)^2 = {r_2}^2 \\ (x_{A3} - x_T)^2 + (y_{A3} - y_T)^2 = {r_3}^2 \end{cases} \quad (9)$$

After applying substation {3rd equation of set (9) - 1st equation of set (9) and similarly 3rd equation of (9)-1st equation of (9)}, the new equation set are as follows:

$$\begin{cases} 2(x_{A1} - x_{A3})x_T + 2(y_{A1} - y_{A3})y_T + x_{A3}^2 - x_{A1}^2 + y_{A3}^2 - y_{A1}^2 = r_3^2 - r_1^2 \\ 2(x_{A2} - x_{A3})x_T + 2(y_{A2} - y_{A3})y_T + y_{A3}^2 - y_{A2}^2 + y_{A3}^2 - y_{A2}^2 = r_3^2 - r_2^2 \end{cases} \quad (10)$$

The equation (10) is equivalent to A.X=B, where

$$A = \begin{vmatrix} 2(x_{N1} - x_{N3}) & 2(y_{N1} - y_{N3}) \\ 2(x_{N2} - x_{N3}) & 2(y_{N2} - y_{N3}) \end{vmatrix}$$

$$X = \begin{vmatrix} x_N \\ y_N \end{vmatrix}$$

$$B = \begin{vmatrix} r_3^2 - r_1^2 + x_{N1}^2 - x_{A3}^2 + y_{A1}^2 - y_{A3}^2 \\ r_3^2 - r_2^2 + x_{A2}^2 - x_{A3}^2 + y_{A2}^2 - y_{A3}^2 \end{vmatrix} \quad (11)$$

The position of the TN (X $(x_T, y_T)$) is as follows:

$$A.X = B, Or \ A^T A.X = A^T.B$$
$$X = (A^T.A)^{-1}.A^T.B \quad or$$

The main aim of the optimization is to minimize the localization error. The fitness function plays a vital role in determining how meta-heuristic optimization algorithms work. The fitness function is a key factor in problem-solving since its design and execution can impact how well the optimization process goes. The difference between the actual distance and the estimated distance is estimated for the localization which is a fitness function in the proposed model as shown in equation-12. It is a mean square of the distance between the estimated node (Xest, Yest) and actual node $(X_{act}, Y_{act})$The mean square error is the obj_fun $((F_i(x)))$ as follows:

$$Fitness_i = F_i: F_i(x, y) = \frac{\sqrt{\sum_{i=1}^{N}((X_{act} - X_{est})^2 + (Y_{act} - Y_{est})^2)}}{N} \quad (12)$$

Where N is the number of TN. This fitness function is used in the following MHOA[18] to enhance the performance.

*Trust Aware Node Localization using COA, and PSO:* Each localized node unknown node runs COA, BOA, and PSO independently using the fitness function shown in equation-(12) to minimize the localization error. Three performance parameters are considered to evaluation and


Copyright © IEEE–2024 ISBN: 979-8-3503-8749-0

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 07:02:02 UTC from IEEE Xplore. Restrictions apply.

analysis the performance as follows:

***Average Node localization Errors ($ANL_e$):*** The $ANL_e$ is the mean distance error between the estimated position (Xest, Yest) and the actual position (Xac, Yac) of the unknown node.

$$ANL_e(\%) = \frac{\sqrt{\sum_{i=1}^{N}\left(\left(X_i^{est} - X_i^{act}\right)^2 + \left(Y_i^{est} - Y_i^{act}\right)^2\right)}}{N_t} \quad (13)$$

***Average Computation Time ($ANL_t(s)$):*** The $ANL_t(s)$ is average computation time of localization is as follows:

$$ANL_t(s) = \frac{\sum_{i=1}^{N} NL_t}{N_t} \quad (14)$$

Where $N_t$ is the total beacon, normal, and malicious nodes deployed in the sensor filed.

*Average Node Localization (ATNL):* ATNL is the average of the total nodes localized in $N_p$ of trails(pass) is as follows.

$$\text{ATNL} = \frac{\sum_{i=1}^{N} TNL_t}{N_p} \quad (15)$$

## IV. Result and Disscusion

Table 3 present the simulation results of the PSO, BOA, and COA Proposed scheme.

Table 2: Simulation parameters

| Deployment field | 200x200 square meters |
|---|---|
| Target Node (TN) | 30, 60, 90, 120, 150, 180 |
| Anchor Node (AN) | 10, 20, 30, 40, 50, 60 |
| Total Base Station (BS) | One |
| Transmission Range | 10 to 30 meters |
| Total no of Pass (Trails) | 10 |
| PSO: Cognitive factor [25] | 1.494 |
| PSO: Social factor [25] | 1.494 |
| PSO: inertial weight(w) [25] | 0.7 |
| BOA: modality (c)[25] | 0.01 |
| BOA: power exponent (a)[25] | 0.1 |

Table 3: Performance metrics of COA_Proposed, PSO, and BOA

| TN | AN | MN | Trail Pass | PSO | | |
|---|---|---|---|---|---|---|
| | | | | $NL_e$(mt) | $NL_t$(s) | TNL |
| 25 | 10 | 5 | 1 | 40.05 | 0.35 | 16 |
| | | | 2 | 39.25 | 0.34 | 17 |
| | | | 3 | 34.70 | 0.35 | 15 |
| | | | 4 | 33.60 | 0.35 | 16 |
| | | | 5 | 36.70 | 0.34 | 17 |
| | | | 6 | 35.80 | 0.34 | 15 |
| | | | 7 | 34.70 | 0.36 | 16 |
| | | | 8 | 33.50 | 0.34 | 15 |
| | | | 9 | 34.55 | 0.36 | 16 |
| | | | 10 | 36.35 | 0.35 | 17 |
| 50 | 20 | 10 | 1 | 32.35 | 0.44 | 35 |
| | | | 2 | 31;35 | 0.43 | 37 |
| | | | 3 | 31.45 | 0.45 | 34 |
| | | | 4 | 30.85 | 0.44 | 36 |
| | | | 5 | 30.95 | 0.45 | 34 |
| | | | 6 | 32.50 | 0.40 | 35 |
| | | | 7 | 31.00 | 0.40 | 34 |
| | | | 8 | 32.00 | 0.45 | 34 |
| | | | 9 | 28.00 | 0.42 | 37 |
| | | | 10 | 30.50 | 0.45 | 36 |
| 75 | 30 | 15 | 1 | 28.75 | 0.61 | 53 |
| | | | 2 | 25.65 | 0.62 | 54 |
| | | | 3 | 27.85 | 0.62 | 54 |
| | | | 4 | 27.35 | 0.63 | 55 |
| | | | 5 | 24.80 | 0.60 | 53 |
| | | | 6 | 27.80 | 0.62 | 54 |
| | | | 7 | 25.25 | 0.65 | 54 |
| | | | 8 | 25.35 | 0.62 | 55 |
| | | | 9 | 30.55 | 0.62 | 54 |
| | | | 10 | 27.90 | 0.65 | 55 |
| 100 | 40 | 20 | 1 | 24.25 | 0.70 | 74 |
| | | | 2 | 24.45 | 0.71 | 75 |
| | | | 3 | 23.55 | 0.70 | 74 |
| | | | 4 | 23.55 | 0.70 | 75 |
| | | | 5 | 24.86 | 0.71 | 74 |
| | | | 6 | 23.55 | 0.72 | 75 |
| | | | 7 | 24.45 | 0.72 | 76 |
| | | | 8 | 23.55 | 0.70 | 75 |
| | | | 9 | 23.25 | 0.72 | 77 |
| | | | 10 | 22.35 | 0.74 | 75 |
| 125 | 50 | 25 | 1 | 19.45 | 0.80 | 102 |
| | | | 2 | 18.95 | 0.78 | 101 |
| | | | 3 | 20.50 | 0.76 | 102 |
| | | | 4 | 10.25 | 0.75 | 103 |
| | | | 5 | 20.40 | 0.75 | 102 |
| | | | 6 | 20.00 | 0.80 | 100 |
| | | | 7 | 19.45 | 0.85 | 101 |
| | | | 8 | 20.25 | 0.80 | 101 |
| | | | 9 | 18.25 | 0.85 | 103 |
| | | | 10 | 19.25 | 0.75 | 101 |
| 150 | 60 | 30 | 1 | 18.75 | 1.15 | 123 |
| | | | 2 | 19.55 | 1.10 | 120 |
| | | | 3 | 18.50 | 1.15 | 120 |
| | | | 4 | 19.55 | 1.18 | 120 |
| | | | 5 | 18.75 | 1.15 | 120 |
| | | | 6 | 19.45 | 1.15 | 119 |
| | | | 7 | 19.45 | 1.18 | 120 |
| | | | 8 | 18.75 | 1.15 | 122 |
| | | | 9 | 18.75 | 1.15 | 123 |
| | | | 10 | 18.55 | 1.16 | 123 |
| | | | | BOA | | |

Copyright © IEEE–2024 ISBN: 979-8-3503-8749-0


Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 07:02:02 UTC from IEEE Xplore. Restrictions apply.

| | | | | $NL_e$(mt) | $NL_t$(s) | TNL |
|---|---|---|---|---|---|---|
| 25 | 10 | 5 | 1 | 29.00 | 0.26 | 17 |
| | | | 2 | 27.50 | 0.27 | 18 |
| | | | 3 | 27.10 | 0.28 | 17 |
| | | | 4 | 32.25 | 0.28 | 16 |
| | | | 5 | 27.95 | 0.26 | 17 |
| | | | 6 | 27.85 | 0.28 | 16 |
| | | | 7 | 27.85 | 0.27 | 15 |
| | | | 8 | 28.25 | 0.26 | 17 |
| | | | 9 | 28.95 | 0.28 | 18 |
| | | | 10 | 29.35 | 0.29 | 19 |
| 50 | 20 | 10 | 1 | 26.80 | 0.37 | 37 |
| | | | 2 | 25.20 | 0.36 | 35 |
| | | | 3 | 26.55 | 0.39 | 36 |
| | | | 4 | 28.45 | 0.36 | 37 |
| | | | 5 | 25.20 | 0.35 | 37 |
| | | | 6 | 26.50 | 0.35 | 34 |
| | | | 7 | 24.20 | 0.38 | 38 |
| | | | 8 | 26.50 | 0.36 | 37 |
| | | | 9 | 28.40 | 0.36 | 36 |
| | | | 10 | 24.25 | 0.37 | 38 |
| 75 | 30 | 15 | 1 | 22.85 | 0.42 | 59 |
| | | | 2 | 21.85 | 0.38 | 57 |
| | | | 3 | 24.45 | 0.37 | 58 |
| | | | 4 | 22.65 | 0.37 | 55 |
| | | | 5 | 23.75 | 0.45 | 54 |
| | | | 6 | 22.65 | 0.38 | 58 |
| | | | 7 | 22.25 | 0.36 | 57 |
| | | | 8 | 24.45 | 0.35 | 56 |
| | | | 9 | 22.58 | 0.40 | 59 |
| | | | 10 | 22.75 | 0.40 | 58 |
| 100 | 40 | 20 | 1 | 22.55 | 0.52 | 78 |
| | | | 2 | 22.35 | 0.54 | 82 |
| | | | 3 | 21.75 | 0.54 | 77 |
| | | | 4 | 21.35 | 0.52 | 78 |
| | | | 5 | 20.55 | 0.52 | 79 |
| | | | 6 | 22.85 | 0.53 | 80 |
| | | | 7 | 22.51 | 0.54 | 79 |
| | | | 8 | 21.80 | 0.52 | 80 |
| | | | 9 | 20.85 | 0.53 | 79 |
| | | | 10 | 20.95 | 0.54 | 79 |
| 125 | 50 | 25 | 1 | 17.50 | 0.75 | 102 |
| | | | 2 | 17.95 | 0.75 | 101 |
| | | | 3 | 16.75 | 0.70 | 102 |
| | | | 4 | 18.00 | 0.75 | 102 |
| | | | 5 | 17.20 | 0.70 | 102 |
| | | | 6 | 17.25 | 0.75 | 104 |
| | | | 7 | 16.85 | 0.70 | 105 |
| | | | 8 | 17.65 | 0.70 | 105 |
| | | | 9 | 17.50 | 0.72 | 102 |
| | | | 10 | 17.50 | 0.75 | 105 |
| 150 | 60 | 30 | 1 | 16.50 | 0.85 | 125 |
| | | | 2 | 15.75 | 0.85 | 130 |
| | | | 3 | 15.85 | 0.80 | 125 |
| | | | 4 | 15.75 | 0.85 | 125 |
| | | | 5 | 16.85 | 0.80 | 124 |
| | | | 6 | 16.00 | 0.82 | 125 |
| | | | 7 | 15.50 | 0.85 | 125 |
| | | | 8 | 15.55 | 0.85 | 125 |
| | | | 9 | 15.45 | 0.85 | 124 |
| | | | 10 | 16.55 | 0.85 | 125 |
| | | | | COA_Proposed | | |
| | | | | $NL_e$(mt) | $NL_t$(s) | TNL |
| 25 | 10 | 5 | 1 | 24.25 | 0.25 | 18 |
| | | | 2 | 23.65 | 0.24 | 18 |
| | | | 3 | 22.75 | 0.22 | 19 |
| | | | 4 | 20.20 | 0.23 | 20 |
| | | | 5 | 21.75 | 0.22 | 20 |
| | | | 6 | 24.10 | 0.22 | 18 |
| | | | 7 | 23.75 | 0.23 | 20 |
| | | | 8 | 23.25 | 0.24 | 18 |
| | | | 9 | 22.25 | 0.22 | 20 |
| | | | 10 | 24.00 | 0.21 | 18 |
| 50 | 20 | 10 | 1 | 20.50 | 0.25 | 40 |
| | | | 2 | 21.75 | 0.26 | 39 |
| | | | 3 | 22.50 | 0.25 | 39 |
| | | | 4 | 23.50 | 0.26 | 40 |
| | | | 5 | 22.50 | 0.26 | 41 |
| | | | 6 | 23.50 | 0.23 | 38 |
| | | | 7 | 22.75 | 0.28 | 41 |
| | | | 8 | 20.10 | 0.27 | 41 |
| | | | 9 | 20.11 | 0.25 | 40 |
| | | | 10 | 22.21 | 0.26 | 38 |
| 75 | 30 | 15 | 1 | 18.75 | 0.36 | 60 |
| | | | 2 | 16.85 | 0.36 | 61 |
| | | | 3 | 17.57 | 0.37 | 60 |
| | | | 4 | 18.85 | 0.35 | 62 |
| | | | 5 | 19.95 | 0.36 | 60 |
| | | | 6 | 18.85 | 0.36 | 62 |
| | | | 7 | 17.35 | 0.35 | 60 |
| | | | 8 | 17.45 | 0.35 | 62 |
| | | | 9 | 18.25 | 0.35 | 62 |
| | | | 10 | 17.45 | 0.36 | 62 |
| 100 | 40 | 20 | 1 | 16.45 | 0.45 | 84 |
| | | | 2 | 15.85 | 0.42 | 82 |
| | | | 3 | 16.70 | 0.44 | 82 |
| | | | 4 | 17.70 | 0.45 | 85 |
| | | | 5 | 17.00 | 0.40 | 82 |
| | | | 6 | 15.25 | 0.42 | 83 |
| | | | 7 | 15.25 | 0.45 | 85 |
| | | | 8 | 14.45 | 0.40 | 85 |
| | | | 9 | 15.65 | 0.44 | 84 |
| | | | 10 | 16.35 | 0.45 | 84 |


Copyright © IEEE–2024 ISBN: 979-8-3503-8749-0

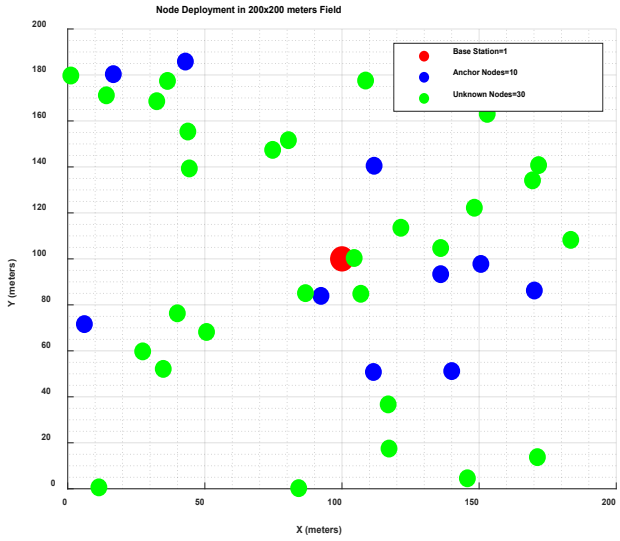
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 07:02:02 UTC from IEEE Xplore. Restrictions apply.

| | | | | | | |
|---|---|---|---|---|---|---|
| 125 | 50 | 25 | 1 | 11.25 | 0.55 | 110 |
| | | | 2 | 12.00 | 0.50 | 112 |
| | | | 3 | 12.20 | 0.55 | 110 |
| | | | 4 | 11.25 | 0.55 | 112 |
| | | | 5 | 11.00 | 0.52 | 110 |
| | | | 6 | 11.45 | 0.55 | 112 |
| | | | 7 | 11.45 | 0.55 | 112 |
| | | | 8 | 10.50 | 0.56 | 110 |
| | | | 9 | 11.25 | 0.55 | 112 |
| | | | 10 | 10.45 | 0.57 | 110 |
| 150 | 60 | 30 | 1 | 09.75 | 0.65 | 145 |
| | | | 2 | 08.55 | 0.65 | 148 |
| | | | 3 | 07.45 | 0.65 | 148 |
| | | | 4 | 06.65 | 0.65 | 145 |
| | | | 5 | 06.65 | 0.62 | 145 |
| | | | 6 | 08.45 | 0.64 | 148 |
| | | | 7 | 08.45 | 0.67 | 145 |
| | | | 8 | 07.45 | 0.65 | 146 |
| | | | 9 | 06.70 | 0.66 | 146 |
| | | | 10 | 06.75 | 0.70 | 145 |

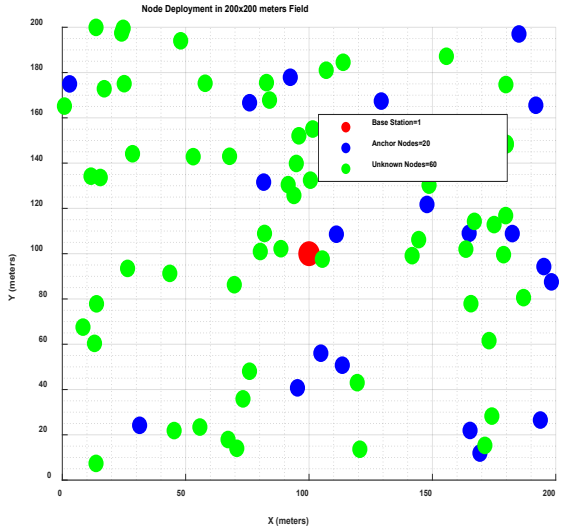
Table 4: Result summery of PSO, BOA, AND COA_Proposed

| **TN** | **AN** | **MN** | **PSO** | | |
|---|---|---|---|---|---|
| No of pass:10 | | | ANLe(m) | $ANL_t$(s) | ATNL (%) |
| 25 | 10 | 5 | 33.955 | 0.355 | 65.25 |
| 50 | 20 | 10 | 30.250 | 0.460 | 68.50 |
| 75 | 30 | 15 | 27.050 | 0.625 | 70.95 |
| 100 | 40 | 20 | 24.550 | 0.670 | 75.85 |
| 125 | 50 | 25 | 20.050 | 0.725 | 81.05 |
| 150 | 60 | 30 | 18.050 | 1.190 | 82.10 |
| **TN** | **AN** | **MN** | **BOA** | | |
| No of pass:10 | | | ANLe(m) | $ANL_t$(s) | ATNL(%) |
| 25 | 10 | 5 | 28.855 | 0.295 | 71.00 |
| 50 | 20 | 10 | 25.985 | 0.365 | 72.95 |
| 75 | 30 | 15 | 22.980 | 0.380 | 77.05 |
| 100 | 40 | 20 | 20.655 | 0.505 | 79.05 |
| 125 | 50 | 25 | 15.950 | 0.695 | 81.95 |
| 150 | 60 | 30 | 14.795 | 0.900 | 83.85 |
| **TN** | **AN** | **MN** | **Proposed_COA** | | |
| No of pass:10 | | | ANLe(m) | $ANL_t$(s) | ATNL(%) |
| 25 | 10 | 5 | 22.915 | 0.220 | 78.00 |
| 50 | 20 | 10 | 21.745 | 0.260 | 80.25 |
| 75 | 30 | 15 | 17.950 | 0.375 | 82.05 |
| 100 | 40 | 20 | 16.025 | 0.500 | 84.85 |
| 125 | 50 | 25 | 13.000 | 0.575 | 88.85 |
| 150 | 60 | 30 | 8.705 | 0.695 | 96.74 |

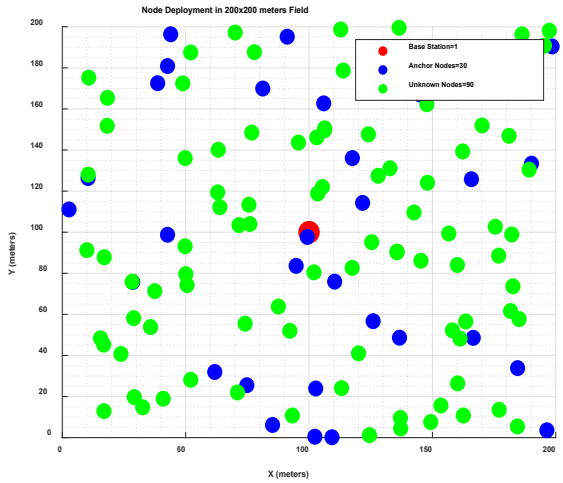
Each case of simulation is run 10 times and mean of NLe*(mt), NLt(s),* and *TNL* is estimated and the scummy of this mean value is presented in Table.4.




(a)



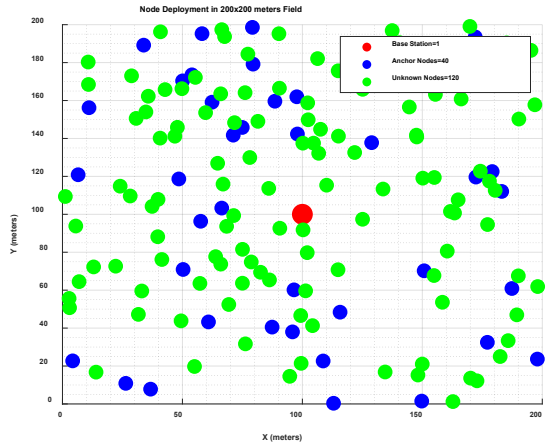

(b)



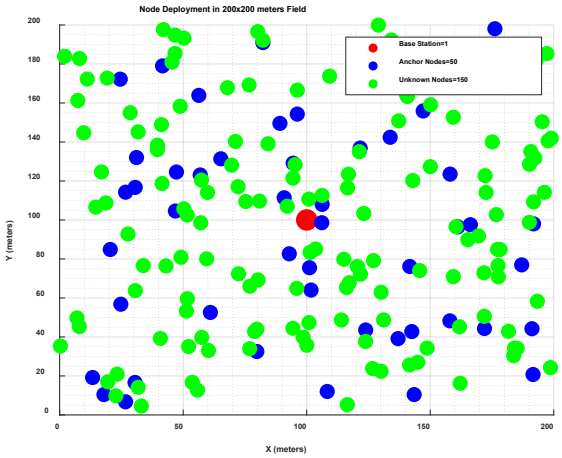

(c)

Copyright © IEEE–2024 ISBN: 979-8-3503-8749-0

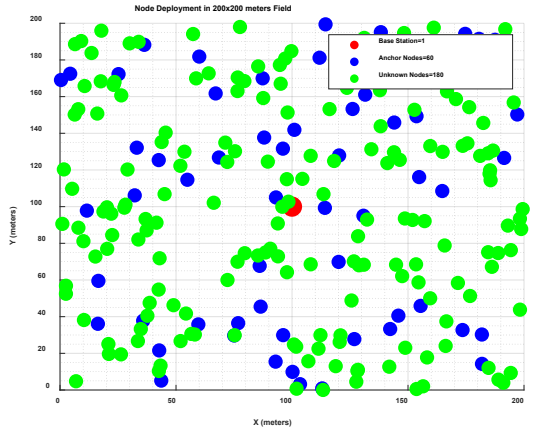

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 07:02:02 UTC from IEEE Xplore. Restrictions apply.




(d)




(e)




(f)

Fig. 4: Six scenarios of sensor nodes (AN, TN, and BS) deployment

The average NL computation time (ANLt): Fig. 7 represent that the value of ANLt increases as the number of anchor, unknown, and malicious nodes increases. The ANLt value of PSO, BOA, and COA increases from 0.355to 1.190, 0.295 to 0.900, and 0.220 to 0.695 seconds respectively. The results demonstrate that the localization computation time of COA is always less than the BOA, and PSO.

This section presents simulation results using the fitness function mentioned in the equation-(12). The MHOAs run independently using the suitable objective function. The mean values of localization, localization error and computation time of localization are presented in Table-3.

***The average node localization (ATNL):*** Fig.5 shows that the value of ATNL improving as the number of anchor, unknown, and malicious nodes increases as tabulated in the Table 3. The ATNL value of PSO, BOA, and COA_Proposed increases from 65.25% to 82.10%, 71.00 to 83.85%, and 78.00% to 96.74% respectively. The localization accuracy of COA_Proposed is always better than the PSO, and BOA.

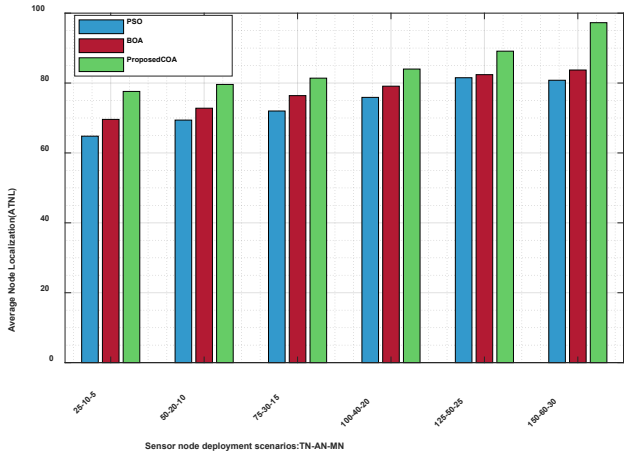



Fig. 5: The average node localization (ATNL) Vs TN-AN-MN

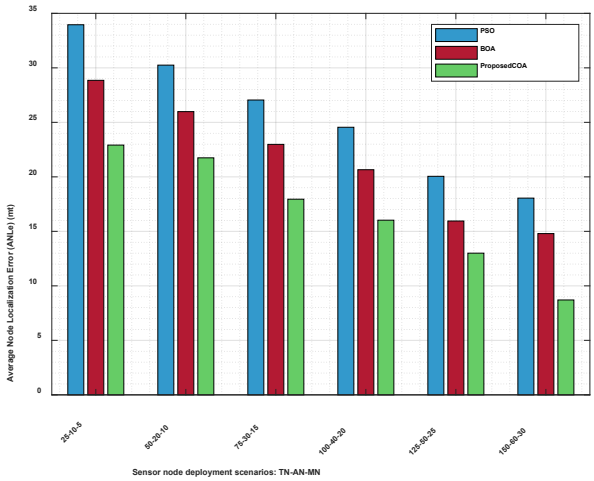



Fig. 6: The average node localization Error (ATNe) Vs TN-AN-MN

*The average node localization (ANL<sub></sub>e):* Fig.6 represent that the value of ANLe decreases as the number of anchor, unknown, and malicious nodes increases. The ANLe value of PSO, BOA, and COA_Proposed decreases from 33.955 to 18.050, 28.855 to 14.795, and 22.915 to 8.705 meters respectively. The localization error of COA_Proposed is always less than the PSO, and BOA.


Copyright © IEEE–2024 ISBN: 979-8-3503-8749-0

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 07:02:02 UTC from IEEE Xplore. Restrictions apply.

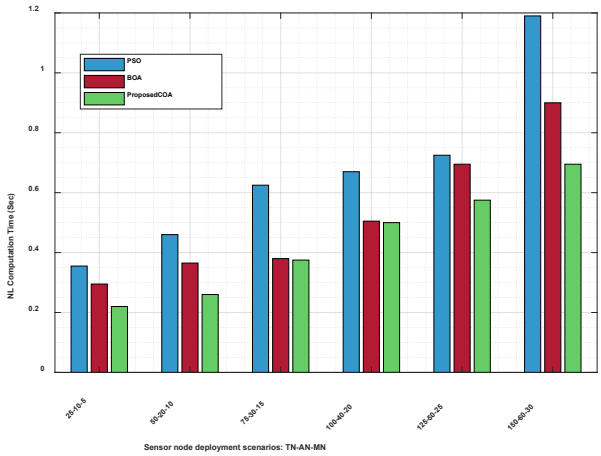



Fig. 7: The average NL computation time (ANLt): Vs TN-AN-MN

## V. Conclusion

Sybil's attack has a major impact on the normal operations of the distributed WSNs. The COA_Proposed scheme perform better than the PSO, and BOA in terms of localization error, localization accuracy, and localization estimation time. To further improve the localization performance, the proposed scheme can be hybridized with another suitable optimization algorithm. It is vital to emphasize that RSSI-based localization is cost-effective but faces many challenges. The challenges are restrictions, including signal jitter, interference, and environmental influences.

## Acknowledgement

The "North Eastern Regional Institute of Science and Technology", Deemed-To-Be-University under the Ministry of Education, Government of India, Nirjuli - 791109, Arunachal Pradesh, is sincerely acknowledged by the first author (corresponding author).

## References

[1] Misra, Sudip, Tamoghna Ojha, and Madhusoodhanan P. "SecRET: Secure range-based localization with evidence theory for underwater sensor networks." ACM Transactions on Autonomous and Adaptive Systems (TAAS) 15, no. 1 (2021): 1-26.

[2] Kumar, Gulshan, Mritunjay Kumar Rai, and Rahul Saha. "Securing range free localization against wormhole attack using distance estimation and maximum likelihood estimation in wireless sensor networks." Journal of Network and Computer Applications 99 (2017): 10-16.

[3] Goyat, Rekha, Gulshan Kumar, Rahul Saha, Mauro Conti, Reji Thomas, and Tai-hoon Kim. "RAPCOL: a range-free power efficient cooperative localization with heterogeneous devices for industrial internet-of-things." Cluster Computing 27, no. 3 (2024): 2691-2713.

[4] Kumar, Anand, Karan Singh, Tayyab Khan, Ali Ahmadian, Mohamad Hanif Md Saad, and Manisha Manjul. "ETAS: an efficient trust assessment scheme for BANs." IEEE Access 9 (2021): 83214-83233.

[5] Chen, Honglong, Wei Lou, Zhi Wang, Junfeng Wu, Zhibo Wang, and Aihua Xia. "Securing DV-Hop localization against wormhole attacks in wireless sensor networks." Pervasive and Mobile Computing 16 (2015): 22-35.

[6] Ahmed, Ayam Tawfeek, Ahmed Noori Rashid, and Khalid Shaker. "Localization in wireless sensor network." Webology 19, no. 1 (2022): 692-704.

[7] Han, Guangjie, Jinfang Jiang, Lei Shu, Mohsen Guizani, and Shojiro Nishio. "A two-step secure localization for wireless sensor networks." The Computer Journal 56, no. 10 (2013): 1154-1166.

[8] Yuan, Yali, Liuwei Huo, Zhixiao Wang, and Dieter Hogrefe. "Secure APIT localization scheme against sybil attacks in distributed wireless sensor networks." IEEE Access 6 (2018): 27629-27636.

[9] He, Jingsha, Jing Xu, Xingye Zhu, Yuqiang Zhang, Ting Zhang, and Wanqing Fu. "Reputation-Based Secure Sensor Localization in Wireless Sensor Networks." The Scientific World Journal 2014, no. 1 (2014): 308341.

[10] Dong, Shi, Xin-gang Zhang, and Wen-gang Zhou. "A security localization algorithm based on DV-hop against Sybil attack in wireless sensor networks." Journal of Electrical Engineering & Technology 15 (2020): 919-926.

[11] Nguyen, Tu N., Vinh V. Le, Shao-I. Chu, Bing-Hong Liu, and Yao-Chuan Hsu. "Secure localization algorithms against localization attacks in wireless sensor networks." Wireless Personal Communications (2022): 1-26.

[12] Han, Guangjie, Li Liu, Jinfang Jiang, Lei Shu, and Joel JPC Rodrigues. "A collaborative secure localization algorithm based on trust model in underwater wireless sensor networks." Sensors 16, no. 2 (2016): 229.

[13] Biswas, Rathindra Nath, Swarup Kumar Mitra, and Mrinal Kanti Naskar. "Wireless node localization under hostile radio environment using smart antenna." Wireless Personal Communications 116, no. 3 (2021): 1815-1836.

[14] Yu, Ning, Lirui Zhang, and Yongji Ren. "A novel D–S based secure localization algorithm for wireless sensor networks." *Security and Communication Networks* 7, no. 11 (2014): 1945-1954.

[15] Li, Xiaoyin, Lianshan Yan, Wei Pan, and Bin Luo. "Secure and robust DV-Hop localization based on the vector refinement feedback method for wireless sensor networks." (2017): 810-821.

[16] Huang, Yan, Jianying Zheng, Yang Xiao, and Miao Peng. "Robust localization algorithm based on the RSSI ranging scope." International Journal of Distributed Sensor Networks 11, no. 2 (2015): 587318.

[17] Wang, Hongbin, Liping Feng, Rong Li, and YiChi Zhang. "The secure localization algorithm of SDV-Hop in wireless sensor networks." Telkomnika 14, no. 3A (2016): 65.

[18] Mohan, Y., Yadav, R.K. & Manjul, M. Seagull optimization algorithm for node localization in wireless sensor networks. Multimed Tools Appl 83, 70793–70814 (2024). https://doi.org/10.1007/s11042-024-18331-8

[19] Gebremariam, Gebrekiros Gebreyesus, J. Panda, and S. Indu. "Blockchain-Based Secure Localization against Malicious Nodes in IoT-Based Wireless Sensor Networks Using Federated Learning." Wireless communications and mobile computing 2023, no. 1 (2023): 8068038.

[20] Gebremariam, Gebrekiros Gebreyesus, J. Panda, and S. Indu. "Secure localization techniques in wireless sensor networks against routing attacks based on hybrid machine learning models." Alexandria Engineering Journal 82 (2023): 82-100.

[21] Asif, Rameez, Muhammad Farooq-i-Azam, Muhammad Hasanain Chaudary, Arif Husen, and Syed Raheel Hassan. "A distance vector hop-based secure and robust localization algorithm for wireless sensor networks." Electronics 12, no. 10 (2023): 2237.

[22] Hai, Xiao. "Hybrid algorithm GA-PSO-BP in secure localization technology for wireless sensor networks." Journal of Cyber Security Technology (2024): 1-17.

[23] Garg, R., Gulati, T. & Kumar, S. Range free localization in WSN against wormhole attack using Farkas' Lemma. Wireless Netw 29, 2029–2043 (2023). https://doi.org/10.1007/s11276-023-03279-8

[24] Kumar, G. Hemanth, and G. P. Ramesh. "Node localization algorithm for detecting malicious nodes to prevent connection failures and improve end-to-end delay." Computer Communications 190 (2022): 37-47.

[25] Liu, Xingcheng, Shaohua Su, Feng Han, Yitong Liu, and Zhihong Pan. "A range-based secure localization algorithm for wireless sensor networks." IEEE Sensors Journal 19, no. 2 (2018): 785-796.

Copyright © IEEE–2024 ISBN: 979-8-3503-8749-0


Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 07:02:02 UTC from IEEE Xplore. Restrictions apply.

Potato Research
https://doi.org/10.1007/s11540-024-09836-8

# Performance Appraisal of a Domesticated Potato Solar Dryer: An Experimental Study

**Shimpy[1] · Mahesh Kumar[1] · Anil Kumar[2,3]**

Received: 14 October 2024 / Accepted: 26 November 2024
© The Author(s), under exclusive licence to European Association for Potato Research 2025

## Abstract

This work aims to investigate the effect of potato slice sample mass on the performance of a domesticated potato solar dryer (DPSD) under natural and forced convection modes (NC and FC modes). The drying kinetics of potato slices with varying sample masses of 68, 100 and 138 g have been evaluated. The average final moisture content under NC and FC modes was found to be 9.00 and 10.36% (wb), respectively. The drying behaviour of samples was precisely described by the Modified-Henderson and Perry model under both modes. With the increase in sample mass, the drying rate increases and effective moisture diffusivity decreases. The capital cost and embodied energy of the DPSD under NC modes were respectively 22.10% and 44.66% lower than that of the FC mode. With the increase in sample mass under both modes, payback period and energy payback time decrease, while heat transfer coefficients, thermal efficiency and exergy efficiency increase. The performance of the DPSD was observed to vary linearly with the sample mass. This work could help in reducing food waste which could directly contribute to global goals of zero hunger and food security.

**Keywords** Domestic solar dryer · Forced convection · Natural convection · Performance assessment · Potato drying

## Nomenclature

| | |
|---|---|
| $A_g$ | Area of glazing, $m^2$ |
| $C_b$ | Market cost of dried potato slices, INR 450/kg |
| $C_{cc}$ | Capital cost, INR |
| $C_{dp}$ | Cost of fresh potatoes per kg of dried potato slices, INR |

✉ Mahesh Kumar
maheshkumar@gjust.org; mkshandilya1@gmail.com

[1] Department of Mechanical Engineering, Guru Jambheshwar University of Science & Technology, Hisar, India

[2] Department of Mechanical Engineering, Delhi Technological University, Delhi, India

[3] Nodal Centre of Excellence in Energy Transition (NCEET), Delhi Technological University, Delhi, India

Published online: 06 January 2025

| | |
|---|---|
| $C_{fp}$ | Cost of fresh potato, INR 20/kg |
| $C_m$ | Annual maintenance cost, INR |
| $C_s$ | Drying cost per kg of dried potato slices, INR |
| $C_{sd}$ | Cost of drying one kg of potato slices, INR |
| $C_v$ | Specific heat, J/kg ℃ |
| $d$ | Inflation rate, 5.69% |
| $D_b$ | Drying days per batch |
| $D_{eff}$ | Effective moisture diffusivity, $m^2/s$ |
| $E_{loss}$ | Exergy loss, W |
| FCS1-3 | Forced convection samples 1 to 3 |
| $h_c$ | Convective heat transfer coefficient, $W/m^2$ ℃ |
| $h_e$ | Evaporative heat transfer coefficient, $W/m^2$ ℃ |
| $I$ | Solar radiation, $W/m^2$ |
| $I$ | Interest rate, 5.50% |
| $j$ | Life of the dryer, 30 years |
| $K_v$ | Thermal conductivity of air, W/m ℃ |
| $L$ | Half of potato slice thickness, m |
| $L_c$ | Least count |
| MC | Moisture content, % (wb) |
| $MC_i$ | Initial moisture content, % (wb) |
| $MC_t$ | Moisture content at time (t) |
| $M_d$ | Dried potato slices per batch, kg |
| $m_{ev}$ | Hourly moisture evaporation, kg |
| $m_{ev,t}$ | Total moisture evaporated, kg |
| $M_i$ | Initial mass of the potato slice sample, kg |
| $M_{ps}$ | Mass of potato sample, kg |
| $MR_{exp}$ | Experimental moisture ratio |
| NCS1-3 | Natural convection samples 1 to 3 |
| $N_D$ | Operating days per year, 300/year |
| $P$(T) | Partial pressure, $N/m^2$ |
| $R$ | Coefficient of correlation |
| RMSE | Root mean square error |
| $S_1$ | Savings at the end of one year, INR |
| $S_b$ | Saving/batch, INR |
| $S_j$ | Savings in j number of years, INR |
| $S_d$ | Saving/day, INR |
| $S_{kg}$ | Savings per kg of the dried potato slices, INR |
| $T_{amb}$ | Ambient temperature, ℃ |
| $T_c$ | Drying chamber temperature, ℃ |
| $T_e$ | Temperature just above the surface of potato sample, ℃ |
| $T_{ps}$ | Surface temperature of potato sample, ℃ |
| $T_s$ | Temperature of the sun, ℃ |
| $T_{vent}$ | Temperature at outlet air vent, ℃ |
| $U$ | Uncertainty, % |
| $V_{max}$ | Maximum value measured by an instrument |
| wb | Wet basis |

| | |
|---|---|
| $\gamma$ | Relative humidity, % |
| $\eta_{ex}$ | Exergy efficiency, % |
| $\eta_{th}$ | Thermal efficiency, % |
| $\lambda$ | Latent heat of vaporisation, J/kg |
| $\mu_v$ | Viscosity, kg/m s |
| $\rho_v$ | Density, kg/m$^3$ |
| $\tau$ | Transmitivity of glazing |
| $\chi^2$ | Chi square |

**Abbreviations**

| | |
|---|---|
| CCE | Carbon credit earned, INR |
| DPSD | Domesticated potato solar dryer |
| DR | Drying rate, g/h |
| DSD | Domestic solar dryer |
| EE | Embodied energy, kWh |
| EEC | Embodied energy coefficient, kWh/kg |
| EPBT | Energy payback time, years |
| FC | Forced convection |
| IP | Improvement potential, W |
| MAE | Mean absolute error |
| NC | Natural convection |
| OSD | Open sun drying |
| SD | Solar dryer |
| SDG | Sustainable development goal |
| SEC | Specific energy consumption, kWh/kg |

## Introduction

Food is the basic necessity for life and its scarcity is a clear threat to the existence of life. Globally, lots of efforts have been made on increasing the production of food items. But this production is not uniform throughout the globe due to geographical, climatic and many other differences. There are places where growing food is difficult and whatever the world produces is not sufficiently feeding all (Grebmer et al. 2023). Food loss and waste is one of the most prominent reasons to this food shortage and its preventions would contribute towards reducing the impact of food scarcity (FAO 2019; UNEP 2021). It is said that "food saved is as important as food produced" (Peters and Sayin 2021). Drying plays a crucial character in improving the life of perishable food items (Fernandes et al. 2022; Shimpy et al. 2022a). But the energy intensive nature of this technique makes its use economically challenging at domestic as well as industrial levels (Kadam and Samuel 2006; Kushwah et al. 2023a). Solar drying is an important advancement in this area as it is non-polluting, economic and simple (Chowdhury et al. 2011; Shimpy and Kumar 2023a).

Globally, potato is the most significant food crop after rice and wheat. There are high chances of its spoilage during storage period (Wang et al. 2023). Many investigations have been conducted on different solar dryers (SDs) for various

commodities, few of them for potato slices drying are discussed here. Yang and Sakai (2001) analysed the shrinkage and mechanical characteristics of potato slices under the effect of convective drying. Bakal et al. (2012) analysed the effect of drying temperature on cubic and cylindrical shaped potato samples with different aspect ratios during fluidized bed drying and reported increase in effective moisture diffusivity with increased temperatures. Behera et al. (2022) tested an indirect forced convection SD for potato slices drying and reported coefficient of performance and heat utilisation ratio as 0.22 and 0.7948, respectively. Pandiaraj et al. (2022) developed a laboratory scale-mixed mode SD working under forced mode (FCMMSD) and having a novel double pass solar air heater with halved trash aluminium cans attached with absorber plate. The thermal efficiency ($\eta_{th}$) and specific energy consumption (SEC) of the dryer were 23.1% and 1.61 kWh/kg, respectively. Dewangan et al. (2023) tested a tent type FCMMSD for the drying of 2.5 mm thick pre-treated potato slices which took 8 hrs at 0.053 kg/s and 14 hrs at 0.093 kg/s. The tent type SD was also tested for potato drying under free and forced modes (Kumar et al. 2023a, b, c). Forced convection mode was reported to reduce the time by 42.8 and 60% compared with free convection and open sun drying (OSD) conditions, respectively. Singh et al. (2024) reported the potato drying efficiency of an active indirect SD as 22.6%. Panchal et al. (2023) modified an indirect SD with phase change materials cascaded solar collector whose $\eta_{th}$ was reported as 9.13%. Lamrani et al. (2022) tested a FCMMSD having a sensible heat storage incorporated V-corrugated absorber for 3 mm thick potato slices drying and reported a 20% reduced drying time compared to OSD with 0.179–1.745 years of payback period. Golpour et al. (2020) reported that the SEC increased from $1.94 \times 10^5$ to $3.14 \times 10^5$ kJ/kg with increased temperature and velocity, whereas exergy efficiency increased from 57.1 to 85.8% with increased temperature and reduced velocity for potato slices drying in a hot air dryer. Karacabey et al. (2023) analysed the effect of blanching on the functional properties of purple-fleshed potato samples during ultrasound and ohmic heating. Gargari et al. (2024) dried potato slices by using infrared-hot air drying and reported that the convective heat transfer coefficient and effective moisture diffusivity varied from 5.32 to 9.98 W/m$^2$ K and $5.18 \times 10^{-8}$ to $9.93 \times 10^{-8}$ m$^2$/s, respectively.

There are severe economic and environmental drawbacks of using electricity powered dryers. Therefore, solar dryers (SDs) are recommended by many researchers around the globe for preserving different edible commodities in an easy, sustainable and economic way. The literature shows numerous design modifications for SDs such as, double glazing (Terres et al. 2015), reflector (Chaudhari and Bhavsar 2017; Sethi and Arora 2009), solar tracking (Jamroen et al. 2022; Sundari et al. 2013), phase change material (Akmak and Yildiz 2011; Azaizia et al. 2019; Poonia et al. 2018) and different cover designs (Islam et al. 2019). Each modification is a step towards designing a simple, efficient and sustainable SD. Recently, Shimpy and Kumar (2024a) designed a simple domestic solar dryer (DSD) and optimised its inlet and outlet air vent heights as 0.5 cm and 2.5 cm, respectively. Shimpy and Kumar (2024b) tested the DSD and reported its no load efficiency as 22.68%. Considering various design and performance perspectives of SDs, this work attempts to develop and test a simple and sustainable domesticated potato solar dryer (DPSD)

under natural convection and forced convection modes. No investigation has been observed in the literature analysing the sequel of varying mass of the sample on the output behaviour of such a simple, easy to use, light-weight, portable and sustainable DPSD. The drying kinetics of sliced potatoes; and econo-environmental and thermo-exergetic performances of the DPSD under both modes were investigated. The effect of sample mass (68, 100 and 138 g) on the drying kinetics, different performance indicators and output behaviour of the DPSD under both modes is also presented. In this novel study, drying behaviour of sliced potatoes and the performance of a DPSD under both modes is assessed and compared. This work will be an addition to the least explored field of domestic solar dryers. It will directly contribute towards reducing domestic food waste and loss and improving food security at household levels thereby controlling global hunger. It will be a step towards strengthening circular economy by the implementation of renewable energy systems in food processing and preservation. This research will directly support the sustainable development goals (SDG) 2, 7 and 12 and will also make an indirect contribution to SDG-3 and 13.

## Materials and Methods

### Experimental Setup

This work aims to develop and test a simple DPSD for potato slices drying under NC and FC modes. Considering the literature on DPSDs, domestic requirements and locally available materials a DPSD was developed. The geometric dimensions of the developed DPSD under FC mode are shown in Fig. 1. An aluminium sheet of 1.2 mm thickness was used to fabricate the drying chamber of 2.5 cm front height, 17 cm back height and 50 × 25 $cm^2$ base area. Inlet air vent (area: 50 × 0.5 $cm^2$) and outlet air vent (area: 2.5 × 2.5 $cm^2$) were created on the front side and back sides, respectively. Under the NC mode the outlet vent was left open (i.e., without fan) for the ventilation of humid air to the surrounding. In the case of the FC mode, a 5 V DC fan of 2.5 × 2.5 $cm^2$ was fitted to the outlet air vent for creating the condition of forced convection. The DC fan was powered through photovoltaic panel (area: 24 × 4 $cm^2$). A window glass (4 mm thick) was used as glazing at the top side of developed DPSD. The drying chamber's interior was painted mat black for maximum solar energy absorption. A composite insulating wall of glass wool, thermocol sheet and plywood of respective thicknesses of 1, 2 and 1 cm surrounded the drying chamber for preventing heat loss to open atmosphere. A drying tray (area: 50 × 25 $cm^2$) made of wooden frame and wire mesh was fabricated to carry the potato slice sample.

### Sampling, Instrumentation and Methodology

Good quality potatoes were purchased locally to prepare samples for the present study. Potatoes were washed, peeled and sliced with a hand slicer into 4-mm thickness. These slices were cut into square shapes (size: 30 × 30 mm) to reduce shape

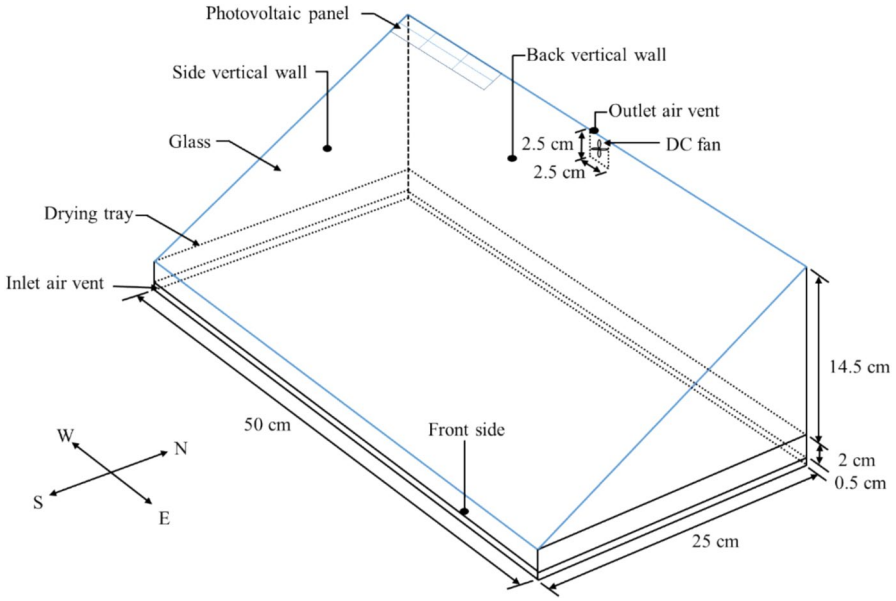



**Fig. 1** Geometric dimensions of the DPSD under forced mode

variations during different experiments. These potato slices were then washed with normal tap water, placed on cotton cloth to remove excess surface water and accommodated on wire mesh tray. Samples with initial mass of 68, 100 and 138 g were prepared and dried in DPSD under both modes. The samples having 68, 100 and 138 g initial mass dried under NC and FC modes were termed as NCS1, NCS2 and NCS3; and FCS1, FCS2 and FCS3, respectively.

Experiments were performed in the month of December, 2023 and the hourly data was collected from 8:30 A.M. onwards. Temperatures at different locations such as surface of potato sample ($T_{ps}$), drying chamber ($T_c$), ambient ($T_{amb}$) and outlet air vent ($T_{vent}$) were recorded through PT-100 temperature indicator and thermocouples (least count, $L_c$ 0.1). Just 2 mm above the potato sample, temperature ($T_e$) and relative humidity ($\gamma$) were observed through a HT-305 hygrometer ($L_c$ 0.1 ℃ and 0.1 %). The MECO-936 solarimeter ($L_c$ 0.1 W/m$^2$) and TJ-6000 digital weighing balance ($L_c$ 0.1 g) were employed to measure solar radiation ($I$) and sample mass ($M_{ps}$), respectively. The values of $M_{ps}$ were used to evaluate the hourly moisture evaporation from the potato sample. Each experiment was continued until moisture evaporation occurred. The data ($T_{ps}$, $T_c$, $T_{vent}$, $T_{amb}$, $T_e$, $\gamma$ and $I$) collection points are shown in the schematics of the DPSD under NC mode (Fig. 2).

## Modelling

Moisture content (MC) of potato slices was evaluated as follows (Ambawat et al. 2022; AOAC 1975).

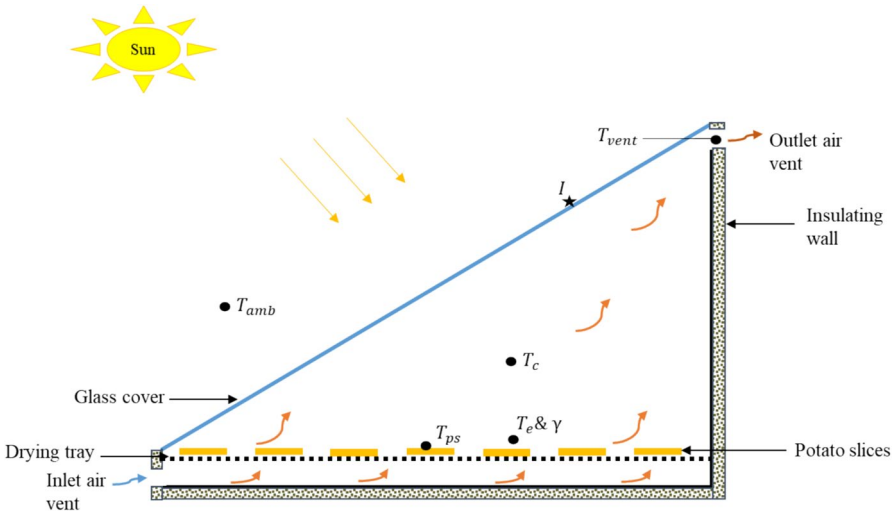



**Fig. 2** Schematics of DPSD under NC mode showing data measuring points

$$MC = \frac{\text{initial mass} - \text{final mass}}{\text{initial mass}} \times 100 = \frac{m_i - m_f}{m_i} \times 100 \tag{1}$$

The ratio of moisture evaporated ($m_{ev}$) to the time (hours) i.e. drying rate (DR) was calculated as follows (Malakar et al. 2022; Singh and Gaur 2021a).

$$DR = \frac{m_{ev}}{\text{time}} \tag{2}$$

The MC values were employed to calculate experimental moisture ratio ($MR_{exp}$) as follows (Sahdev et al. 2017a; Singh et al. 2022).

$$MR_{exp} = \frac{\text{MC at time t}}{\text{initial MC}} = \frac{MC_t}{MC_i} \tag{3}$$

The values of $MR_{exp}$ were employed to Lewis: $MR = \exp(-kt)$, Handerson and Pabis: $MR = a \times \exp(-kt)$, Midilli-Kucuk: $MR = a \times \exp(-kt^n) + bt$ and Silva et al: $MR = \exp(-at - b\sqrt{t})$ (Lewis 1921; Midilli et al. 2002; Pavis and Henderson 1961; Silva et al. 2014) models to get predicted moisture ratio ($MR_{pre}$). The $MR_{pre}$ for different modes were compared for the best suited model by evaluating the statistical parameters given in Appendix A.

The effective moisture diffusivity ($D_{eff}$) as half of potato slice thickness ($L$) and drying time ($t$) was evaluated as follows (Rani and Tripathy 2022).

$$MR = \frac{8}{\pi^2} \times \exp\left(-\frac{\pi^2 D_{eff}}{4L^2} t\right) \tag{4}$$

Rewriting Equation (4) as a linear equation by taking its natural logarithm and the slope which can be used to evaluate the value of $D_{\text{eff}}$ as follows.

$$ln(MR) = ln\left(\frac{8}{\pi^2}\right) - \left(\frac{\pi^2 D_{eff}}{4L^2}t\right) \tag{5}$$

$$\text{Slope} = \frac{\pi^2 D_{eff}}{4L^2} \tag{6}$$

The various expressions given in this section are used for performing economic assessment (ELkhadraoui et al. 2015; Mani and Natesan 2021; Mohammed et al. 2020).

$$\text{Annualised cost, } C_a = C_{ac} + C_m - S_a \tag{7}$$

$C_{ac}$ and $S_a$ are given in Appendix B. $C_m$ is annual maintenance cost ($0.3\times C_{ac}$).

From Appendix B, considering cost of fresh potatoes per kg of dried potato slices ($C_{dp}$), drying cost per kg of dried potato slices ($C_s$), cost of fresh potato ($C_{fp}$) INR 20/kg, initial mass of the potato slice sample ($M_i$), dried potato slices per batch ($M_d$), market cost of dried potato slices ($C_b$) INR 450/kg, the cost of drying one kg of potato slices ($C_{sd}$) and savings were evaluated as,

$$C_{sd} = C_{dp} + C_s \tag{8}$$

$$S_{kg} = C_b - C_{sd} \tag{9}$$

$$S_b = S_{kg} \times M_d \tag{10}$$

$$S_d = \frac{S_b}{D_b} \tag{11}$$

Here, $S_{kg}$: savings per kg of the dried potato slices, $S_b$: saving/batch, $S_d$: saving/day and $D_b$: drying days per batch.

Total savings in the life (j) of DPSD (30 years) was calculated as follows (Shimpy and Kumar 2024c).

$$S_j = S_d \times N_D(1+i)^{j-1} \tag{12}$$

$$\text{Payback period, } PBP = \frac{\ln\left[1 - \frac{C_{cc}}{S_1}(i-d)\right]}{\ln\left(\frac{1+d}{1+i}\right)} \tag{13}$$

$S_1$: savings at the end of one year and d: inflation rate (5.69%).

The amount of energy utilised in the development of the DPSD was calculated as the product of quantity of the material and the associated embodied energy coefficient (EEC) as follows (Nawaz and Tiwari 2006; Singh and Gaur 2021b).

$$\text{EE} = \text{Quantity of material used} \times EEC \tag{14}$$

The energy payback time (EPBT) of DPSD for potato slice sample drying was evaluated as (Chel and Tiwari 2009; Kushwah et al. 2023b),

$$EPBT = \frac{EE}{\text{Annual energy output}} = \frac{EE}{\text{Daily energy output} \times N_D} \tag{15}$$

$$\text{Daily energy output} = \frac{m_{ev,t} \times \lambda}{3.6 \times 10^6} \tag{16}$$

where $m_{\text{ev,t}}$ is total moisture evaporated and $N_{\text{D}}$ is the number of operating days taken as 300/year.

The annual value of $CO_2$ emission, $CO_2$ mitigation and carbon credit earned (CCE) were evaluated as (Chel and Tiwari 2009; Mishra et al. 2021; Nawaz and Tiwari 2006);

$$\text{Annual}\, CO_2 \,\text{emission} = \frac{EE \times 0.98}{Life(j)} \tag{17}$$

$$CO_2 \text{ mitigation} = \left( \frac{m_{ev,t} \times \lambda}{3.6 \times 10^6} \times 300 \times \text{Life} - EE \right) \times 2.042 \tag{18}$$

$$CCE = CO_2 \text{ mitigation} \times \text{Cost of carbon credit} \tag{19}$$

where the life (j) of the DPSD was considered 30 years.

For potato slices drying inside the DPSD, the convective heat transfer coefficient ($h_{\text{c}}$) for NC and FC modes was evaluated as (Das and Akpinar 2018; Ekka and Palanisamy 2020),

$$\text{For NC mode} : h_c = \frac{K_v}{X_d} C(GrPr)^n \tag{20}$$

$$\text{For FC mode} : h_c = \frac{K_v}{X_d} C(RePr)^n \tag{21}$$

Heat required for moisture evaporation ($Q_e$) and the amount of moisture evaporated ($m_{ev}$) from the potato samples were given by using the following expressions Kumar et al. 2023b; Shimpy et al. 2022b),

$$Q_e = 0.016 h_c [P(T_{ps}) - \gamma P(T_e)] \tag{22}$$

$$\text{For NC mode}: m_{ev} = 0.016\frac{K_v}{\lambda X}C(GrPr)^n\left[P\left(T_{ps}\right) - \gamma P\left(T_e\right)\right]tA_t \tag{23}$$

$$\text{For FC mode}: m_{ev} = 0.016\frac{K_v}{\lambda X}C(RePr)^n\left[P\left(T_{ps}\right) - \gamma P\left(T_e\right)\right]tA_t \tag{24}$$

Equations 23 and 24 can be reduced by considering $0.016\frac{K_v}{\lambda X}\left[P\left(T_{ps}\right) - \gamma P\left(T_e\right)\right]tA_t = Z$ and written as a liner equation ($Y = mX + c$) by taking natural logarithm as below (Kushwah et al. 2020).

$$ln\left[\frac{m_{ev}}{Z}\right] = nln(GrPr) + lnC \tag{25}$$

$$ln\left[\frac{m_{ev}}{Z}\right] = nln(RePr) + lnC \tag{26}$$

Equations 25 and 26 were solved by employing simple linear regression method. The evaporative heat transfer coefficient ($h_e$) was calculated as follows (Anwar and Tiwari 2001; Shimpy and Kumar 2023b).

$$h_e = 0.016h_c\left[\frac{P\left(T_{ps}\right) - \gamma P\left(T_e\right)}{T_{ps} - T_e}\right] \tag{27}$$

The values of $T_i$ were used for obtaining thermo-physical properties of drying air (Appendix C).

$$T_i = \frac{T_{ps,avg} + T_{e,avg}}{2} \tag{28}$$

Thermal efficiency ($\eta_{th}$) of the DPSD for the drying of potato samples was evaluated as follows (Almuhanna 2011; Tiwari et al. 1998),

$$\eta_{th} = \frac{m_{ev} \times \lambda}{\left(\text{IA}_g\right) \times \tau \times t} \tag{29}$$

The following expressions were used to calculate various exergetic indicators for the drying of potato slices inside the DPSD (Atalay and Aslan 2023; Maia et al. 2017; Öztürk and Hu 2005; Vijayan et al. 2020).

$$\text{Exergy efficiency } (\eta_{ex}) = \frac{\frac{m_{ev}}{3600} \times \left[\lambda - \left(1 - \frac{T_{amb}}{T_{ps}}\right)\right]}{\sum\left(\text{IA}_g\right) \times \left[1 - \frac{4}{3} \times \left(\frac{T_{amb}}{T_S}\right) + \frac{1}{3} \times \left(\frac{T_{amb}}{T_S}\right)^4\right]} \tag{30}$$

$$\text{Exergy loss } (E_{loss}) = \text{Exergy input} - \text{ Exergy output} \tag{31}$$

$$\text{Improvement potential } (IP) = \left(1 - \eta_{ex}\right) \times E_{loss} \tag{32}$$

The experimental percent uncertainty was evaluated as follows (Sehrawat et al. 2023),

$$U = \frac{L_c}{V_{max}} \times 100 \tag{33}$$

## Experimental Data

The DPSD was tested for the drying of 4-mm thick potato slice samples having initial mass of 68, 100 and 138 g under NC and FC modes. Samples of 68, 100 and 138 g mass under NC and FC modes were termed as NCS1, NCS2 and NCS3; and FCS1, FCS2 and FCS3, respectively. The hourly experimental data of solar radiations ($I$), ambient temperature ($T_{amb}$), potato slice sample surface temperature ($T_{ps}$), just above the potato slice sample temperature ($T_e$), relative humidity ($\gamma$) and sample mass ($M_{ps}$) was observed from 8:30 A.M. onwards. The data of $I$ and $T_{amb}$ observed for potato samples are shown in Figs. 3 and 4, respectively.

The average values of $I$ and $T_{amb}$ were found to be 551.7, 476.4, 532.1, 581.9, 545.4 and 585.2 W/m$^2$; and 19.9, 17.8, 18.6, 19.3, 19.2 and 19.3 ℃ for the drying of NCS1, NCS2, NCS3, FCS1, FCS2 and FCS3, respectively. The observed data of $T_{ps}$, $T_e$, $\gamma$ and $M_{ps}$ for given samples is presented in Table 1.

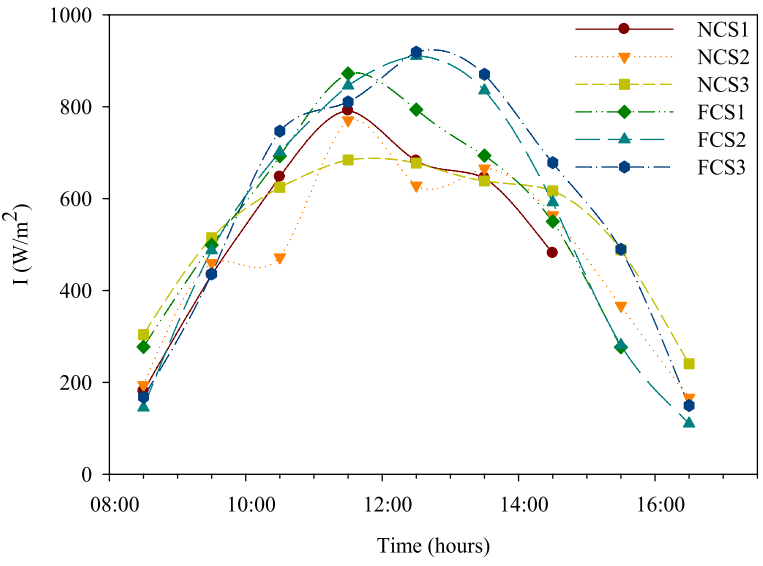



**Fig. 3** Hourly variations in $I$ for different experiments

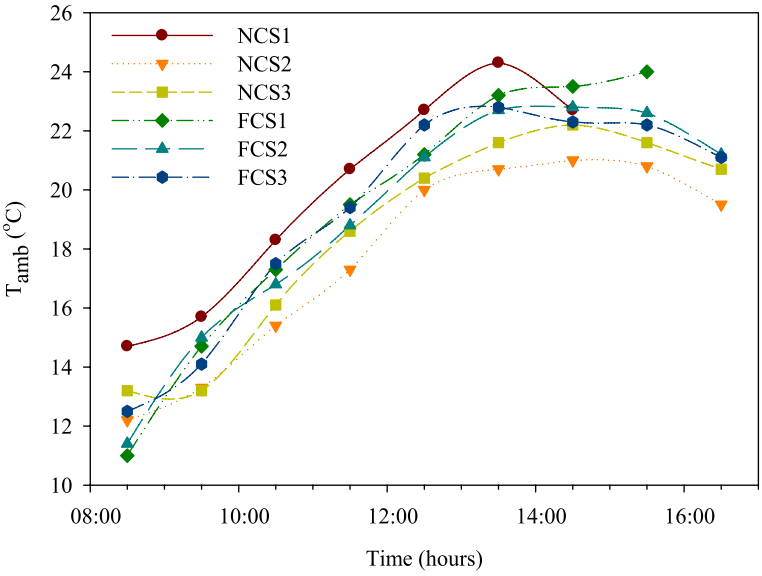



**Fig. 4** Hourly variations in $T_{amb}$ for different experiments

From Table 1, the values of $T_{ps}$, $T_e$ and $\gamma$ for the given samples were found to be in the range of 17.6–62.4 ℃, 17.9–58.9 ℃ and 11.5–79.5%, respectively. The final mass ($M_{ps}$) for NCS1, NCS2, NCS3, FCS1, FCS2 and FCS3 was 11.1, 16.4, 23.1, 11.3, 16.7 and 23.3 g, respectively.

## Results and Discussion

### Drying Kinetics

The data of $M_{ps}$ from Table 1 was used to evaluate the values of total moisture evaporated ($m_{ev,t}$) which were found as 56.9, 83.6, 114.9, 56.7, 83.3 and 114.7 g for NCS1, NCS2, NCS3, FCS1, FCS2 and FCS3, respectively. It was observed that the $m_{ev,t}$ increased with increased initial mass of the sample in both modes. The possible reason is that higher sample mass has higher amount of moisture and vice versa. The $m_{ev,t}$ was found to be slightly higher in the case of NC mode. The value of DR was calculated for given potato slice sample drying in the DPSD under both modes. The average values of DR for NCS1, NCS2, NCS3, FCS1, FCS2 and FCS3 were respectively found to be 3.49, 3.64, 5.33, 3.05, 3.44 and 4.69 g/h. It was found that the DR values under NC mode were higher than the respective values under FC mode. The value of DR were also found to increase with increased sample mass. Higher initial sample mass could be one of the reasons for higher DR values for samples NCS3 and FCS3.

The data of hourly moisture evaporation was used for the evaluation of hourly moisture content (MC) for given samples. The value of MC was reduced from 85%

**Table 1** Observed data of $T_{ps}$, $T_e$, $\gamma$ and $M_{ps}$ for given samples

| Time | $T_{ps}$ (℃) | $T_e$ (℃) | $\gamma$ (%) | $M_{ps}$ (g) | $T_{ps}$ (℃) | $T_e$ (℃) | $\gamma$ (%) | $M_{ps}$ (g) | $T_{ps}$ (℃) | $T_e$ (℃) | $\gamma$ (%) | $M_{ps}$ (g) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sample | NCS1 | | | | NCS2 | | | | NCS3 | | | |
| 08:30 | 20.4 | 18.9 | 71.0 | 68.0 | 19.7 | 17.9 | 77.0 | 100.0 | 19.6 | 18.2 | 57.8 | 138.0 |
| 09:30 | 41.8 | 37.6 | 46.9 | 63.5 | 32.4 | 34.8 | 53.4 | 93.8 | 37.8 | 30.7 | 59.8 | 126.0 |
| 10:30 | 49.1 | 42.6 | 58.5 | 52.3 | 48.1 | 46.3 | 38.9 | 80.0 | 46.5 | 44.9 | 46.3 | 105.9 |
| 11:30 | 57.4 | 53.1 | 29.6 | 38.1 | 55.8 | 51.9 | 37.0 | 58.7 | 52.3 | 50.6 | 34.4 | 80.8 |
| 12:30 | 62.4 | 55.3 | 22.5 | 22.2 | 55.3 | 52.4 | 25.2 | 40.8 | 55.1 | 51.2 | 31.7 | 57.1 |
| 13:30 | 64.3 | 57.4 | 13.1 | 13.0 | 55.0 | 53.9 | 34.8 | 26.0 | 54.6 | 52.4 | 18.9 | 35.8 |
| 14:30 | 48.2 | 44.8 | 19.8 | 11.1 | 48.6 | 48.1 | 19.8 | 19.2 | 47.5 | 49.4 | 15.9 | 26.6 |
| 15:30 | - | - | - | - | 42.2 | 43.9 | 25.7 | 16.9 | 41.2 | 43.0 | 16.8 | 23.7 |
| 16:30 | - | - | - | - | 38.0 | 28.1 | 38.7 | 16.4 | 36.9 | 33.8 | 22.3 | 23.1 |
| Time | $T_{ps}$ (℃) | $T_e$ (℃) | $\gamma$ (%) | $M_{ps}$ (g) | $T_{ps}$ (℃) | $T_e$ (℃) | $\gamma$ (%) | $M_{ps}$ (g) | $T_{ps}$ (℃) | $T_e$ (℃) | $\gamma$ (%) | $M_{ps}$ (g) |
| Sample | FCS1 | | | | FCS2 | | | | FCS3 | | | |
| 08:30 | 21.1 | 21.4 | 73.1 | 68.0 | 17.6 | 18.4 | 78.0 | 100.0 | 19.4 | 17.9 | 79.5 | 138.0 |
| 09:30 | 41.9 | 46.7 | 44.1 | 63.0 | 34.1 | 34.3 | 62.3 | 94.9 | 30.6 | 33.6 | 80.1 | 130.0 |
| 10:30 | 48.8 | 56.2 | 27.3 | 50.8 | 44.5 | 42.3 | 42.1 | 82.1 | 43.6 | 41.3 | 36.0 | 111.2 |
| 11:30 | 53.1 | 58.9 | 23.6 | 36.4 | 48.0 | 47.3 | 32.1 | 63.6 | 44.8 | 45.3 | 38.5 | 91.5 |
| 12:30 | 57.4 | 58.4 | 18.6 | 23.7 | 51.9 | 50.0 | 26.5 | 43.8 | 47.1 | 46.4 | 36.5 | 71.0 |
| 13:30 | 53.9 | 54.1 | 14.6 | 15.7 | 53.2 | 54.3 | 20.1 | 29.6 | 47.8 | 46.8 | 29.4 | 49.0 |
| 14:30 | 50.1 | 50.5 | 11.5 | 12.6 | 50.3 | 48.4 | 19.3 | 19.8 | 47.6 | 45.6 | 24.0 | 33.6 |
| 15:30 | 49.2 | 47.7 | 12.5 | 11.3 | 46.7 | 47.1 | 18.2 | 17.7 | 41.6 | 41.4 | 22.1 | 26.0 |
| 16:30 | - | - | - | - | 33.1 | 30.5 | 30.4 | 16.7 | 32.7 | 31.6 | 29.2 | 23.3 |

Abbreviations: $T_{ps}$, temperature at the surface of potato sample; $T_e$, temperature just 2 mm above the surface of potato sample; $\gamma$, relative humidity; $M_{ps}$, sample mass; *NCS1-3*, natural convection samples 1 to 3; *FCS1-3*, forced convection samples 1 to 3

(wb) to 8.10, 8.50, 10.40, 9.70, 10.18 and 11.20% (wb) for NCS1, NCS2, NCS3, FCS1, FCS2 and FCS3, respectively. The average final moisture content under NC and FC modes was 9.00 and 10.36% (wb), respectively, which is desirable for the preservation of dried food commodities. The NCS1 took 6 h of drying time followed by FCS1 that took 7 h. The remaining samples, namely, NCS2, NCS3, FCS2 and FCS3 took 8 h. The possible reason for shorter drying time of NCS1 and FCS1 could be the lowest moisture present in the samples due to low initial sample mass.

The $MR_{exp}$ values were evaluated by using hourly MC and then fitted to the given drying models to get $MR_{pre}$ values. The $MR_{pre}$ values were compared by evaluating different statistical indicators given in Appendix A and it was found that all the given models fitted very well to the values of $MR_{exp}$ for potato slice samples ( since $R > 0.95$). However, the most desirable values of $R$, RMSE, $\chi^2$ and MAE were shown by Midilli-Kucuk model. The $R$ values of Midilli-Kucuk model for given samples were found to be in the range of 0.9930 to 0.9999 and the range of RMSE, $\chi^2$ and MAE were found to be 0.0050 to 0.0149, 0.00003 to 0.00030 and 0.0040 to 0.0128, respectively. The values of constants ($k_1$, $n_1$, $a_1$ and $b_1$) of the Midilli-Kucuk model were found to be in the range of 0.051–0.073, 2.028–2.459, 0.985–0.994 and 0.0–0.0, respectively, for NC mode and 0.048–0.085, 2.043–2.235, 0.986–0.998 and 0.0–0.0, respectively, for FC mode. It is observed that the values of $b_1$ were 0.0 for all the samples for which the Midilli-Kucuk model reduces to the Modified-Henderson and Perry model (Ertekin and Firat 2017). The average of model constants ($k_1$, $n_1$ and $a_1$) were calculated to give a single model for potato slices drying under both NC and FC modes. The obtained average models ($MR_{pre,a}$) along with $MR_{exp}$ for given samples are shown in Fig. 5.

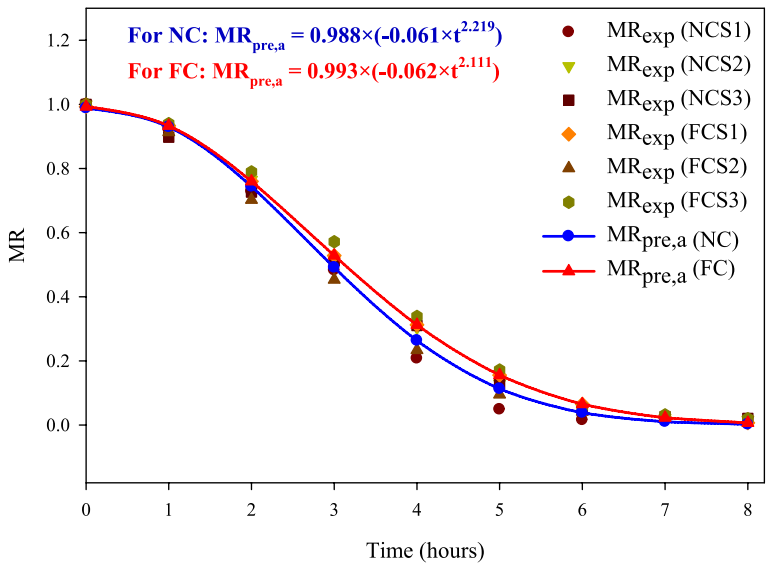



**Fig. 5** $MR_{pre,a}$ along with $MR_{exp}$ for given samples

Figure 5 shows that the obtained average models fitted very well to the values of $MR_{exp}$ for potato slice samples. For NC and FC modes, the average values of $R$, RMSE, $\chi^2$ and MAE shown by $MR_{pre,a}$ were 0.999, 0.025, 0.001 and 0.020; and 0.997, 0.042, 0.002 and 0.032, respectively. The $R$ values showed an excellent fit between $MR_{pre,a}$ and $MR_{exp}$ data for both modes. It is inferred that the reduced Midilli-Kucuk model or the Modified-Henderson and Perry model exhibited the drying behaviour of potato slices very precisely.

Figure 6 shows the plot of $\ln(MR_{exp})$ with time, linear fit equations and $R$ values for given samples.

From Fig. 6, it can be observed that the value of the slope of linear fit reduced with the increased sample mass of potato slices. It was maximum (0.7012) for NCS1 and minimum (0.05492) for NCS3 in case of NC mode; and maximum (0.0099) for FCS1 and minimum (0.008) for FCS3 in case of FC mode. Most of the $R$ values were greater than 0.85, showing a good fit. The values of $D_{eff}$ were found to be $1.15 \times 10^{-7}$, $2.36 \times 10^{-8}$ and $2.26 \times 10^{-8}$ $m^2/s$; and $1.63 \times 10^{-9}$, $1.48 \times 10^{-9}$ and $3.33 \times 10^{-10}$ $m^2/s$ for NCS1, NCS2, NCS3, NCS4, NCS5 and NCS6, respectively. The values of $D_{eff}$ for all the samples were as per the values reported for solar drying of pears (Guiné et al. 2007). It was also observed that the $D_{eff}$ were higher under NC mode in comparison to FC mode. It shows that higher moisture evaporation occurred under the NC mode in comparison to the FC mode. The values of $D_{eff}$ were also observed to be reducing with increased sample mass. It could be due to the higher drying chamber temperatures under NC mode. Higher temperature and lower sample mass resulted in higher values of $D_{eff}$ which was responsible for higher moisture evaporation.

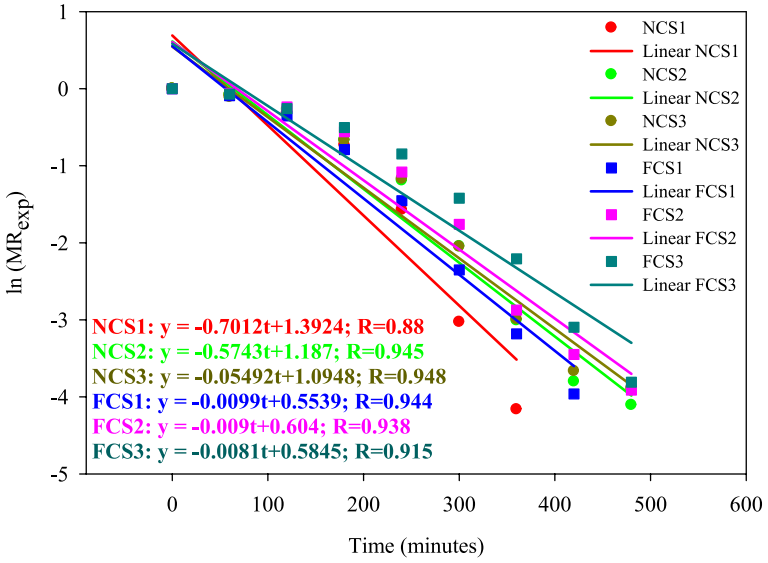



**Fig. 6** In ($MR_{exp}$) with time for given potato slice sample drying

### Econo-Environmental Assessment

The values of $C_{cc}$ and EE calculated for domesticated potato solar dryer (DPSD) under both modes are presented in Table 2.

The values of $C_{cc}$ and EE of the DPSD under NC and FC modes was calculated as INR1047.26 ($12.56) and 1278.26 ($ 15.34); and 69.48 and 100.51 kWh, respectively. The values of $C_C$ and EE of the DPSD were increased due to the addition of photovoltaic panel, DC fan and electric wires in the case of FC mode. The total cost evaluated for the DPSD under NC and FC modes was found to be lower than many of the small scale solar dryer available in the literature such as NC direct SD (Haque et al. 2018; Singh et al. 2006), FC indirect SD (Sreekumar et al. 2008), and FC direct SD (Poonia et al. 2018).

The values of $C_C$ and EE of the DPSD under both modes were used for the evaluation of various econo-environmental indicators for given potato sample drying (Table 3).

The annualised cost of the DPSD under NC and FC modes was evaluated as INR72.77 and 88.83, respectively. The drying cost per kg of potato slices ($C_{sd}$) was evaluated as INR 144.38, 136.74, 129.98, 146.56, 137.49 and 131.73 for NCS1, NCS2, NCS3, FCS1, FCS2 and FCS3, respectively. The respective values of savings/kg ($S_{kg}$) and savings at the end of the life (30 years) of the dryer ($S_j$) were evaluated as INR 305.62, 313.26, 320.02, 303.44, 312.51 and 318.27; and INR 4807.87, 7280.95, 10476.82, 4859.59, 7396.44 and 10464.76. The payback period (PBP) was evaluated as 1.03, 0.68, 0.47, 1.24, 0.82 and 0.58 years for NCS1, NCS2, NCS3, FCS1, FCS2 and FCS3, respectively. From Table 3, it can be observed that the values of $C_{sd}$ for the NC mode were comparatively lower than their respective values for the FC mode. It can also be observed that the drying cost of one kg of potato slices decreased with the increasing sample mass under both NC and FC modes. The values of $S_{kg}$ and $S_j$ were found to be increasing with the increased sample mass. The PBP of the DPSD was also found to be reducing with the increase in the sample mass. It could be due to the reduced drying cost and improved savings because of higher mass of dried product obtained for sample NCS3 and FCS3. The lowest PBP under both NC and FC was 0.47 year and 0.58 year, respectively, which are very low in comparison to the considered life of the DPSD and many SDs reported in the literature, such as FC indirect SD (Sreekumar et al. 2008) and FC direct SD (Poonia et al. 2018).

The $CO_2$ emissions for DPSD were found to be 2.27 kg/year and 3.28 kg/year under NC and FC modes, respectively. It is observed that $CO_2$ emissions by the DPSD under both NC and FC modes are very low compared to the indirect active SD (Shrivastava and Kumar 2017), passive greenhouse SD (Ahmad and Prakash 2020), indirect passive SD (Gilago et al. 2023a), direct type forced convection SD (Tiwari et al. 2021), multi-shelf NCDSD (Jain et al. 2019), and wind energy-assisted hybrid SD (Atalay 2022). From Table 3, it can be observed that all the environmental indicators for DPSD under NC mode were comparatively better than their corresponding indicators under FC mode. The possible reason for this could be the higher moisture evaporation under NC compared to FC mode. Table 3 also shows that as the sample mass increased, the EPBT of both the DPSD reduced; and the

**Table 2** $C_{cc}$ and EE of DPSD under NC and FC modes

| Material with quantity | Cost per unit (INR) | $C_C$ (INR) | | EE coefficient (kWh/kg) | EE (kWh) | |
|---|---|---|---|---|---|---|
| | | NC mode | FC mode | | NC mode | FC mode |
| Aluminium sheet (0.754 kg) | 430.00 | 324.22 | 324.22 | 55.28 | 41.68 | 41.68 |
| Glass (1.95 kg) | 77.00 | 149.84 | 149.84 | 4.417 | 8.60 | 8.60 |
| Glass wool (0.75 kg) | 24.00 | 18.00 | 18.00 | 4.06 | 3.05 | 3.05 |
| Hinges (0.03 kg, 2 nos.) | 15.00 | 30.00 | 30.00 | 55.28 | 1.66 | 1.66 |
| Nails and screws (0.19 kg) | 20.00 | 3.80 | 3.80 | 8.89 | 1.69 | 1.69 |
| Paint (0.07 kg) | 120.00 | 8.40 | 8.40 | 25.11 | 1.76 | 1.76 |
| Plywood (2.10 kg) | 210.00 | 441.00 | 441.00 | 2.88 | 6.05 | 6.05 |
| Photovoltaic panel (0.03 kg, 3 nos.) | 40.00 | - | 120.00 | 980.00 | - | 29.40 |
| DC fan (0.07 kg, 1 nos.) | 99.00 | - | 99.00 | 19.44 | - | 1.40 |
| Electric wires (0.01 kg,1.2 m) | 10.00 | - | 12.00 | 19.61 | - | 0.24 |
| Thermocol sheet (0.16 kg) | 300.00 | 48.00 | 48.00 | 24.61 | 3.94 | 3.94 |
| Wire mesh (0.12 kg) | 200.00 | 24.00 | 24.00 | 8.89 | 1.07 | 1.07 |
| Total | | 1047.26 ($ 12.56) | 1278.26 ($ 15.34) | | 69.48 | 100.51 |

Abbreviations: $C_{cc}$, capital cost; *EE*, embodied energy; *DPSD*, domesticated potato solar dryer; *NC*, natural convection; *FC*, forced convection; *INR*, Indian rupee; $1 = INR83.35 (21-11-2023).

**Table 3** Evaluated values of econo-environmental indicators for given samples

| Sample | NCS1 | NCS2 | NCS3 | FCS1 | FCS2 | FCS3 |
|---|---|---|---|---|---|---|
| Economic indicators | | | | | | |
| $C_{sd}$ (INR) | 144.38 | 136.74 | 129.98 | 146.56 | 137.49 | 131.73 |
| $S_{kg}$ (INR/kg) | 305.62 | 313.26 | 320.02 | 303.44 | 312.51 | 318.27 |
| $S_j$ (INR) | 4807.87 | 7280.95 | 10476.80 | 4859.59 | 7396.44 | 10464.80 |
| PBP (years) | 1.03 | 0.68 | 0.47 | 1.24 | 0.82 | 0.58 |
| Environmental indicators | | | | | | |
| EPBT (years) | 6.03 | 4.10 | 2.99 | 8.75 | 5.96 | 4.33 |
| $CO_2$ emissions (kg/year) | 2.27 | 2.27 | 2.27 | 3.28 | 3.28 | 3.28 |
| $CO_2$ mitigation (tons) | 0.56 | 0.90 | 1.28 | 0.50 | 0.83 | 1.22 |
| CCE (INR) | 587.59 | 932.68 | 1337.22 | 518.98 | 862.78 | 1268.61 |

*NCS1-3*, natural convection samples 1 to 3; *FCS1-3*, forced convection samples 1 to 3; $C_{sd}$, drying cost per kg of potato slice; $S_{kg}$, savings per kg of dried potato slices; $S_j$, savings in $j$ number of years; $j$, life of the dryer; *PBP*, payback period; *EPBT*, energy payback time; *CCE*, carbon credit earned; *INR*, Indian rupee

$CO_2$ mitigation and CCE increased. This could be due to higher total moisture evaporation from the drying commodity obtained from higher initial sample mass.

## Thermo-Exergetic Indicators

The data recorded during experimentation was employed to evaluate the values of thermo-exergetic indicators for given sample drying (Table 4).

Table 4 shows that all the values of $h_c$ and $h_e$ under FC mode were higher than the corresponding values under NC mode. It was also found that the values of $h_c$ and $h_e$ increased with the increased sample mass under both modes. The maximum values of $h_c$ and $h_e$ were observed for NCS3 and FCS3. The constants $C$ and $n$ varied in the

**Table 4** Calculated values of thermo-exergetic indicators

| Sample | NCS1 | NCS2 | NCS3 | FCS1 | FCS2 | FCS3 |
|---|---|---|---|---|---|---|
| Thermal indicators | | | | | | |
| $h_c$ (W/m$^2$ ℃) | 0.40 | 0.52 | 0.66 | 0.41 | 0.54 | 0.94 |
| $h_e$ (W/m$^2$ ℃) | 12.12 | 63.14 | 135.93 | 66.78 | 105.35 | 161.91 |
| $\eta_{th}$ (%) | 6.95 | 8.55 | 11.05 | 5.61 | 7.37 | 10.37 |
| Exergetic indicators | | | | | | |
| $\eta_{ex}$ (%) | 4.43 | 5.51 | 6.96 | 3.50 | 4.24 | 5.49 |
| $E_{loss}$ (W) | 79.22 | 64.96 | 70.09 | 81.46 | 76.68 | 81.26 |
| IP (W) | 75.51 | 60.87 | 64.74 | 78.31 | 72.86 | 76.44 |

Abbreviations: *NCS1-3*, natural convection samples 1 to 3; *FCS1-3*, forced convection samples 1 to 3; $h_c$, convective heat transfer coefficient; $h_e$, evaporative heat transfer coefficient; $\eta_{th}$, thermal efficiency; $\eta_{th}$, exergy efficiency; $E_{loss}$, exergy loss; *IP*, improvement potential

range of 0.55–1.27 and 0.12–0.2 for potato slices drying in DPSD. Higher values of $h_c$ and $h_e$ with increasing sample mass have also been reported by Sahdev et al. (2017a, b) for open sun groundnuts drying and Kumar and Tiwari (2007) for NC and FC greenhouse onion drying. Higher values of heat transfer coefficients under FC compared to NC have also been reported by Elangovan and Natarajan (2023) for ivy gourd drying inside a direct type SD and (Gilago et al. 2023b) for carrot drying in an indirect SD. It was observed that the $\eta_{th}$ for the given sample dried under the NC mode was higher than the corresponding values under the FC mode. The values of $\eta_{th}$ were observed to increase with increased sample mass under both modes. The maximum $\eta_{th}$ under NC mode was observed for NCS3 (11.05%) which was 37.1% and 22.62% higher than for NCS1 and NCS2, respectively. Under the FC mode, the maximum $\eta_{th}$ was observed for sample FCS3 (10.37%) which was 45.9% and 28.93% higher than for FCS1 and FCS2, respectively. The maximum $\eta_{th}$ under NC was also observed to be 6.15% higher than that of corresponding efficiency under the FC mode. The possible reason for the higher value of $\eta_{th}$ under NC mode could be the lower outlet air vent heat loss.

From Table 4, it was found that the $\eta_{ex}$ values for potato sample drying under NC modes were higher than the corresponding values under FC modes. The $\eta_{ex}$ values were found to increase with increased sample mass under both modes. The NCS3 showed the maximum $\eta_{ex}$ value (6.96%) under NC mode that was 36.4% and 20.8% higher than the NCS1 and NCS2, respectively. The maximum $\eta_{ex}$ value under FC mode was obtained for FCS3 (5.49%) that was 36.3% and 22.8% higher than the FCS1 and FCS2, respectively. The maximum $\eta_{ex}$ value under NC was also found to be 21.1% higher than the corresponding value under the FC mode.

### Comparison of Performance Indicators and Effect of Sample Mass on the Drying Behaviour of DPSD under NC and FC Modes

This section presents the effect of NC and FC modes on the performance of the DPSD for potato slices drying. Average of each performance indicator under both mode was calculated. The average values of DR, $D_{eff}$, PBP, EPBT, $\eta_{th}$ and $\eta_{ex}$ under both modes with percent difference and desirability are shown in Table 5.

Table 5 shows that there was a prominent difference in the performance indicators of DPSD under NC and FC modes. The percentage difference between the DR, $D_{eff}$, PBP, EPBT, $\eta_{th}$ and $\eta_{ex}$ under NC and FC modes was 10.27, 97.86, 21.10, 45.12, 12.05 and 21.72%, respectively. Table 5 also shows that NC mode is desirable over the FC mode for the given performance indicators. It can be stated on the basis of Table 5 that the DPSD performed better under the NC mode in comparison to the FC mode for potato slices drying.

The effect of samples mass on different performance indicators, namely, DR, $D_{eff}$, PBP, EPBT, $h_c$, $h_e$, $\eta_{th}$ and $\eta_{ex}$ was evaluated by using method of linear fitting ($y = bx + a$). The values of performance indicators were plotted against the sample

**Table 5** Comparison of performance indicators of DPSD

| Performance indicator | DPSD | | Difference (%) | Desirability |
|---|---|---|---|---|
| | NC mode | FC mode | | |
| DR (g/h) | 4.15 | 3.73 | 10.27 | NC mode |
| $D_{\text{eff}}$ ($m^2$/s) | $5.37 \times 10^{-8}$ | $1.15 \times 10^{-9}$ | 97.86 | NC mode |
| PBP (years) | 0.73 | 0.88 | 21.10 | NC mode |
| EPBT (years) | 4.37 | 6.35 | 45.12 | NC mode |
| $\eta_{\text{th}}$ (%) | 8.85 | 7.78 | 12.05 | NC mode |
| $\eta_{\text{ex}}$ (%) | 5.63 | 4.41 | 21.72 | NC mode |

Abbreviations: *NC*, natural convection; *FC*, forced convection; *DR*, drying rate; $D_{eff}$, effective moisture diffusivity; *PBP*, payback period; *EPBT*, energy payback period; $\eta_{th}$, thermal efficiency; $\eta_{th}$, exergy efficiency

**Table 6** Effect of sample mass on performance indicators of DPSD

| Indicator | *R* | *b* | *a* | *R* | *b* | *a* |
|---|---|---|---|---|---|---|
| | NC mode | | | FC mode | | |
| $M_{\text{ps}}$ (g) | 0.9976 | 35.0000 | 32.0000 | 0.9976 | 35.0000 | 32.0000 |
| DR (g/h) | 0.8483 | 0.0268 | 1.4264 | 0.9406 | 0.0236 | 1.4264 |
| $D_{\text{eff}}$ ($m^2$/s) | 0.7146 | $-1 \times 10^{-9}$ | $2 \times 10^{-7}$ | 0.8704 | $-2 \times 10^{-11}$ | $3 \times 10^{-9}$ |
| PBP (years) | 0.9633 | −0.0079 | 1.5349 | 0.9583 | −0.0093 | 1.8319 |
| EPBT (years) | 0.9590 | −0.0430 | 8.7582 | 0.9606 | −0.0625 | 12.7230 |
| $h_{\text{c}}$ (W/$m^2$ ℃) | 1.0000 | 0.0037 | 0.1479 | 0.9451 | 0.0077 | −0.1479 |
| $h_{\text{e}}$ (W/$m^2$ ℃) | 0.9973 | 1.7733 | −110.4800 | 0.9965 | 1.3630 | −27.6800 |
| $\eta_{\text{th}}$ (%) | 0.9941 | 0.0588 | 2.8529 | 0.9901 | 0.0683 | 0.0813 |
| $\eta_{\text{ex}}$ (%) | 0.9988 | 0.0362 | 1.9404 | 0.9905 | 0.0286 | 1.4962 |

Abbreviations: *R*, coefficient of correlation; *a*, constant; *b*, slope; *NC*, natural convection; *FC*, forced convection; $M_{ps}$, mass of potato samples; *DR*, drying rate; $D_{eff}$, effective moisture diffusivity; *PBP*, payback period; *EPBT*, energy payback period; $h_c$, convective heat transfer coefficient; $h_e$, evaporative heat transfer coefficient; $\eta_{th}$, thermal efficiency; $\eta_{th}$, exergy efficiency

mass for drawing linear fits for each indicators under both modes. Table 6 shows the values of *R*, slope (*b*) and constant (*a*) obtained for different indicators of DPSD.

It can be observed from Table 6 that the sample mass increased almost linearly, which is also reflected in the *R* value (0.9976) of the linear fit for sample mass. Negative or positive *b* values reflect the increasing or decreasing behaviour of an indicator with increasing sample mass. The *b* values also show the degree of impact of sample mass on a particular indicator. It was found that the DR, $h_{\text{c}}$, $h_e$, $\eta_{\text{th}}$ and $\eta_{\text{ex}}$ increase and $D_{\text{eff}}$, PBP and EPBT decreased with the increase in sample mass. It can be stated from Table 6 that all the performance indicators for the DPSD varied linearly with linearly increasing sample mass under both modes.

Hence, it can be concluded that the performance of DPSD is a linear function of sample mass under both modes.

The values of percent experimental uncertainty for given potato sample drying inside the DPSD varied in the range of 0.608–0.632% and 0.611–0.639% under NC and FC modes, respectively.

## Conclusions

A simple, small scale novel domesticated potato solar dryer (DPSD) was tested for potato slices drying under NC and FC modes. Effect of sample mass on the drying behaviour of potato slices and performance of the DPSD was investigated under NC and FC modes. The average final moisture content of dried potato slices in the DPSD under NC mode was 9% (wb) which was 15.11% lower than under the FC mode. The drying behaviour of potato slices was well exhibited by Modified-Henderson and Perry model under both modes. The values of DR increased and $D_{\text{eff}}$ decreased with increased sample mass. The $C_{\text{cc}}$ and EE of the DPSD under NC modes were 22.1% and 44.66% lower than that of FC mode, respectively. The PBP and EPBT of the DPSD under both modes was found to be decreasing with increasing sample mass. The FC mode resulted in higher heat transfer coefficients (HTCs) compared to NC mode. HTCs, $\eta_{\text{th}}$ and $\eta_{\text{ex}}$ were found to be increasing with increased sample mass for both modes. The performance of the DPSD varied linearly with sample mass under both modes.

## Limitations, Future Scope and Implications

It can be concluded from this investigation that the performance of a solar dryer (SD) varies with the sample mass. So, the most suitable sample mass should be used to get the maximum output from any SD. In this work, the effect of 68, 100 and 138 g of sample mass on the performance of the DPSD has been investigated. The designed DPSD can be further tested for different and wider range of sample mass of potato and other food and agricultural commodities. The DPSD could also be modified with energy storage elements to prevent intermittent operations. Drying is a highly practised food preservation technique and has a significant contribution in reducing global food wastage. This study presents a very simple, small scale domesticated potato solar dryer (DPSD) with a thought of empowering common households and ensure their participation in combating global food loss and wastage (SDG-12.3). The DPSDs would help in improving the food security of the users through wastage reduction.

## Appendix A

The statistical parameters used (Essalhi et al. 2018; Kumar et al. 2024; Silva et al. 2014; Singh and Kumar 2012).

$$R = \frac{N\sum_{i=1}^{N} MR_{exp,i}MR_{pre,i} - \left(\sum_{i=1}^{N} MR_{exp,i}\right)\left(\sum_{i=1}^{N} MR_{pre,i}\right)}{\sqrt{N\sum_{i=1}^{N} MR_{exp,i}{}^{2} - \left(\sum_{i=1}^{N} MR_{exp,i}\right)^{2}}\sqrt{N\sum_{i=1}^{N} MR_{pre,i}{}^{2} - \left(\sum_{i=1}^{N} MR_{pre,i}\right)^{2}}} \tag{34}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left(MR_{exp,i} - MR_{pre,i}\right)^{2}}{N}} \tag{35}$$

$$\chi^{2} = \left[\frac{\sum_{i=1}^{N}\left(MR_{exp,i} - MR_{pre,i}\right)^{2}}{N - S}\right] \tag{36}$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^{N}\left|MR_{exp,i} - MR_{pre,i}\right|}{N} \tag{37}$$

## Appendix B

$$C_{ac} = C_{cc} \times F_{c} = C_{cc} \times \frac{i(1+i)^{j}}{(1+i)^{j} - 1} \tag{38}$$

where i and j are the interest rate (5.50%) and life of the DPSD, respectively.

The value of $S_a$ can be given as follows (M. Kumar et al. 2023a; Prasanna and Deshmukh 2022).

$$S_{a} = \text{salvage value} \times \text{salvage fund factor} = \left(0.1 \times C_{cc}\right) \times F_{s} \tag{39}$$

$$F_{s} = \frac{i}{(1+i)^{j} - 1} \tag{40}$$

$$C_{dp} = C_{fp}\frac{M_{i}}{M_{d}} \tag{41}$$

$$C_{s} = \frac{C_{a}}{\left(\frac{M_{d}N_{D}}{D_{b}}\right)} \tag{42}$$

## Appendix C

Thermo-physical properties of the drying air (Henderson-Sellers 1984; Jain 2006; M. Kumar et al. 2023a).

$$C_v = 999.2 + 0.1434 \times T_i + 1.101 \times 10^{-4} T_i^2 - 6.7581 \times 10^{-8} T_i^3 \quad (43)$$

$$K_v = 0.0244 + 0.7673 \times 10^{-4} T_i \quad (44)$$

$$\mu_v = 1.718 \times 10^{-5} + 4.620 \times 10^{-8} T_i \quad (45)$$

$$\rho_v = \frac{353.44}{T_i + 273.15} \quad (46)$$

$$P(T) = \exp\left(25.317 - \frac{5144}{T + 273.15}\right) \quad (47)$$

$$\lambda = \left(2500.82 - 2.358 \times T_i\right) 10^3 \quad (48)$$

**Acknowledgements** The authors are thankful to Guru Jambheshwar University of Sciences and Technology, Hisar, and Delhi Technological University, Delhi, for providing all kind of facilities in the completion of this manuscript.

**Author Contribution** Shimpy: writing—original draft, formal analysis, visualisation, methodology and validation. Mahesh Kumar: conceptualization, supervision, writing—review and editing. Anil Kumar: conceptualization, supervision, writing—review and editing.

**Data Availability** Data available on request from the authors.

### Declarations

**Conflict of Interest** The authors declare no competing interests.

**Declaration of Originality** This article has not been published elsewhere and it has not been simultaneously submitted for publication elsewhere.

## References

Ahmad A, Prakash O (2020) Performance evaluation of a solar greenhouse dryer at different bed conditions under passive mode. J Sol Energy Eng Trans ASME 142:1–10. https://doi.org/10.1115/1.4044194

Akmak G, Yildiz C (2011) The drying kinetics of seeded grape in solar dryer with PCM-based solar integrated collector. Food Bioprod Process 89:103–108. https://doi.org/10.1016/j.fbp.2010.04.001

Almuhanna EA (2011) Utilization of a solar greenhouse as a solar dryer for drying dates under the climatic conditions of the eastern province of Saudi Arabia. J Agric Sci 4:237–246. https://doi.org/10.5539/jas.v4n3p237

Ambawat S, Sharma A, Saini RK (2022) Mathematical modeling of thin layer drying kinetics and moisture diffusivity study of pretreated moringa oleifera leaves using fluidized bed dryer. Processes 10:2464. https://doi.org/10.3390/PR10112464

Anwar SI, Tiwari GN (2001) Heat and mass transfer coefficients in four-tray solar crop drying system. Int J Ambient Energy 22:48–56. https://doi.org/10.1080/01430750.2001.9675386

AOAC (1975) Association of official analytical chemists. Official Methods of analysis. Aoac Int

Atalay H (2022) Exergoeconomic and environmental impact evaluation of wind energy assisted hybrid solar dryer and conventional solar dryer. Renew Energy 200:1416–1425. https://doi.org/10.1016/j.renene.2022.10.033

Atalay H, Aslan V (2023) Advanced exergoeconomic and exergy performance assessments of a wind and solar energy powered hybrid dryer. Renew Energy 209:218–230. https://doi.org/10.1016/j.renene.2023.03.137

Azaizia Z, Kooli S, Hamdi I et al (2019) Experimental study of a new mixed mode solar greenhouse drying system with and without thermal energy storage for pepper. Renew Energy 145:1972–1984. https://doi.org/10.1016/j.renene.2019.07.055

Bakal SB, Sharma GP, Sonawane SP, Verma RC (2012) Kinetics of potato drying using fluidized bed dryer. J Food Sci Technol 49:608–613. https://doi.org/10.1007/s13197-011-0328-x

Behera DD, Mohanty RC, Mohanty AM, et al (2022) Performance evaluation of the absorbing plate of a hybrid solar dryer for potato drying. IOP Conf Ser Earth Environ Sci 1084. https://doi.org/10.1088/1755-1315/1084/1/012011

Chaudhari RH, Bhavsar S (2017) Hybrid solar box type dryer cum cooker of chilly drying for domestic usage. Int J Sci Res 6:1614–1618

Chel A, Tiwari GN (2009) Thermal performance and embodied energy analysis of a passive house - case study of vault roof mud-house in India. Appl Energy 86:1956–1969. https://doi.org/10.1016/j.apenergy.2008.12.033

Chowdhury MMI, Bala BK, Haque MA (2011) Energy and exergy analysis of the solar drying of jackfruit leather. Biosyst Eng 110:222–229. https://doi.org/10.1016/j.biosystemseng.2011.08.011

da Silva WP, e Silva CMDPS, Gama FJA, Gomes JP (2014) Mathematical models to describe thin-layer drying and to determine drying rate of whole bananas. J Saudi Soc Agric Sci 13:67–74. https://doi.org/10.1016/j.jssas.2013.01.003

Das M, Akpinar EK (2018) Investigation of pear drying performance by different methods and regression of convective heat transfer coefficient with support vector machine. Appl Sci 8:215. https://doi.org/10.3390/app8020215

Dewangan N, Verma G, Ghritlahre HK et al (2023) An experimental investigation of mixed-mode tent house solar dryer using ultraviolet sheet for drying potato slices. Energy Sources, Part A Recover Util Environ Eff 45:11446–11466. https://doi.org/10.1080/15567036.2023.2259852

Ekka JP, Palanisamy M (2020) Determination of heat transfer coefficients and drying kinetics of red chilli dried in a forced convection mixed mode solar dryer. Therm Sci Eng Prog 19:100607. https://doi.org/10.1016/j.tsep.2020.100607

Elangovan E, Natarajan SK (2023) Convective and evaporative heat transfer coefficients during drying of ivy gourd under natural and forced convection solar dryer. Environ Sci Pollut Res 30:10469–10483. https://doi.org/10.1007/s11356-022-22865-5

ELkhadraoui A, Kooli S, Hamdi I, Farhat A (2015) Experimental investigation and economic evaluation of a new mixed- mode solar greenhouse dryer for drying of red pepper and grape. Renew Energy 77:1–8. https://doi.org/10.1016/j.renene.2014.11.090

Ertekin C, Firat MZ (2017) A comprehensive review of thin-layer drying models used in agricultural products. Crit Rev Food Sci Nutr 57:701–717. https://doi.org/10.1080/10408398.2014.910493

Essalhi H, Benchrifa M, Tadili R, Bargach MN (2018) Experimental and theoretical analysis of drying grapes under an indirect solar dryer and in open sun. Innov Food Sci Emerg Technol 49:58–64. https://doi.org/10.1016/j.ifset.2018.08.002

FAO (2019) The State of Food and Agriculture 2019. Moving Forward on Food Loss and Waste Reduction, Rome

Fernandes L, Fernandes JR, Tavares PB (2022) Design of a friendly solar food dryer for domestic overproduction. Solar 2:495–508. https://doi.org/10.3390/solar2040029

Gargari AN, Asefi N, Roufegarinejad L, Khodaei M (2024) Investigation of heat transfer in combined infrared-hot air drying: a strategy for evaluation in potato food model. Potato Res. https://doi.org/10.1007/s11540-024-09730-3

Gilago MC, Mugi VR, Chandramohan VP (2023) Evaluating the environ-economic and exergy-energy impacts of drying carrots in passive and active mode solar dryers. Therm Sci Eng Prog 43:101956. https://doi.org/10.1016/j.tsep.2023.101956

Gilago MC, Mugi VR, Chandramohan VP (2023) Evaluation of drying kinetics of carrot and thermal characteristics of natural and forced convection indirect solar dryer. Results Eng 18:101196. https://doi.org/10.1016/j.rineng.2023.101196

Golpour I, Kaveh M, Chayjan RA, Guiné RPF (2020) Energetic and exergetic analysis of a convective drier: a case study of potato drying process. Open Agric 5:563–572. https://doi.org/10.1515/opag-2020-0058

GrebmerBernstein KVJ, Wiemers M et al (2023) 2023 Global hunger index: the power of youth in shaping food systems. Glob Hunger Index Chapter 01:1–60

Guiné RPF, Ferreira DMS, Barroca MJ, Gonçalves FM (2007) Study of the drying kinetics of solar-dried pears. Biosyst Eng 98:422–429. https://doi.org/10.1016/j.biosystemseng.2007.09.010

Haque T, Tiwari M, Bose M, Kedare SB (2018) Drying kinetics, quality and economic analysis of a domestic solar dryer for agricultural products. Ina Lett 4:147–160. https://doi.org/10.1007/s41403-018-0052-1

Henderson-Sellers B (1984) A new formula for latent heat of vaporization of water as a function of temperature. Q J R Meteorol Soc 110:1186–1190. https://doi.org/10.1002/qj.49711046626

Islam MMI, Islam MMI, Tusar M, Limon AH (2019) Effect of cover design on moisture removal rate of a cabinet type solar dryer for food drying application. Energy Procedia 160:769–776. https://doi.org/10.1016/j.egypro.2019.02.181

Jain D (2006) Determination of convective heat and mass transfer coefficients for solar drying of fish. Biosyst Eng 94:429–435. https://doi.org/10.1016/j.biosystemseng.2006.04.006

Jain A, Sharma M, Kumar A et al (2019) Computational fluid dynamics simulation and energy analysis of domestic direct-type multi-shelf solar dryer. J Therm Anal Calorim 136:173–184. https://doi.org/10.1007/s10973-018-7973-5

Jamroen C, Komkum P, Yoopum P et al (2022) Improvement of an open sun drying system for dried banana product using solar tracking system: a case study in Thailand. Int J Green Energy 19:1085–1097. https://doi.org/10.1080/15435075.2021.1982718

Kadam DM, Samuel DVK (2006) Convective flat-plate solar heat collector for cauliflower drying. Biosyst Eng 93:189–198. https://doi.org/10.1016/j.biosystemseng.2005.11.012

Karacabey E, Bardakçı MS, Baltacıoğlu H (2023) Physical pretreatments to enhance purple-fleshed potatoes drying: effects of blanching, ohmic heating and ultrasound pretreatments on quality attributes. Potato Res. https://doi.org/10.1007/s11540-023-09618-8

Kumar A, Tiwari GN (2007) Effect of mass on convective mass transfer coefficient during open sun and greenhouse drying of onion flakes. J Food Eng 79:1337–1350. https://doi.org/10.1016/j.jfoodeng.2006.04.026

Kumar S, Dewangan N, Verma G et al (2023) Performance analysis of natural and forced convection mixed mode UV tent house solar dryer for potato drying. Energy Sources, Part A Recover Util Environ Eff 45:11482–11504. https://doi.org/10.1080/15567036.2023.2259839

Kumar P, Rani P, Tripathy PP (2024) Indirect mode solar drying of stevia (Stevia rebaudiana) leaves: experimental and numerical investigation of fluid flow pattern, moisture, and temperature distribution profile. Biomass Convers Biorefinery. https://doi.org/10.1007/s13399-024-05725-9

Kumar M, Shimpy, Sahdev RK et al (2023a) Natural convective greenhouse vermicelli drying: thermo-environ-econo-kinetic analyses. Sustain Energy Technol Assessments 55:103002. https://doi.org/10.1016/j.seta.2022.103002

Kumar M, Shimpy, Sahdev RK, et al (2023b) Experimental forced convection greenhouse and indirect cabinet drying of date fruits: a comparative study. J Therm Anal Calorim 148:5437–5454. https://doi.org/10.1007/s10973-023-12057-9

Kushwah A, Gaur MK, Singh P, Thakur V (2020) Experimental study on natural and force convection hybrid active greenhouse solar drying of mushroom. Glob J Energy Environ 2:1–12

Kushwah A, Kumar A, Gaur MK, Pal A (2023) Heat and mass transfer, quality, performance analysis, and modeling of thin layer drying kinetics of banana slices. ASME J Sol Energy Eng 145:1–14. https://doi.org/10.1115/1.4062447

Kushwah A, Kumar A, Kumar Gaur M, Shrivastava P (2023) Environmental sustainability and exergetic based sustainability indicators for heat exchanger-evacuated tube assisted drying system (HE-ETADS). Sustain Energy Technol Assessments 57:103277. https://doi.org/10.1016/j.seta.2023.103277

Lamrani B, Elmrabet Y, Mathew I et al (2022) Energy, economic analysis and mathematical modelling of mixed-mode solar drying of potato slices with thermal storage loaded V-groove collector: application to Maghreb region. Renew Energy 200:48–58. https://doi.org/10.1016/j.renene.2022.09.119

Lewis WK (1921) The rate of drying of solid materials. In: The Journal of Industrial and Engineering Chemistry. pp 427–432

Maia CB, Ferreira AG, Cabezas-Gómez L et al (2017) Thermodynamic analysis of the drying process of bananas in a small-scale solar updraft tower in Brazil. Renew Energy 114:1005–1012. https://doi.org/10.1016/j.renene.2017.07.102

Malakar S, Dhurve P, Arora VK (2022) Modeling and optimization of osmo-sonicated dehydration of garlic slices in a novel infrared dryer using artificial neural network and response surface methodology. J Food Process Eng. https://doi.org/10.1111/jfpe.14261

Mani P, Natesan VT (2021) Experimental investigation of drying characteristics of lima beans with passive and active mode greenhouse solar dryers. J Food Process Eng 44:1–12. https://doi.org/10.1111/jfpe.13667

Midilli A, Kucuk H, Yapar Z (2002) A new model for single-layer drying. Dry Technol 20:1503–1513. https://doi.org/10.1081/DRT-120005864

Mishra L, Sinha A, Gupta R (2021) Energy, exergy, economic and environmental (4E) analysis of greenhouse dryer in no-load condition. Sustain Energy Technol Assessments 45:101186. https://doi.org/10.1016/j.seta.2021.101186

Mohammed S, Fatumah N, Shadia N (2020) Drying performance and economic analysis of novel hybrid passive-mode and active-mode solar dryers for drying fruits in East Africa. J Stored Prod Res 88:101634. https://doi.org/10.1016/j.jspr.2020.101634

Nawaz I, Tiwari GN (2006) Embodied energy analysis of photovoltaic (PV) system based on macro- and micro-level. Energy Policy 34:3144–3152. https://doi.org/10.1016/j.enpol.2005.06.018

Öztürk HH, Hu H (2005) Experimental evaluation of energy and exergy efficiency of a seasonal latent heat storage system for greenhouse heating. Energy Convers Manag 46:1523–1542. https://doi.org/10.1016/j.enconman.2004.07.001

Panchal JM, Modi KV, Chaudhary SB et al (2023) Comparative experimental analysis of potato slice drying using a multiple phase change material cascaded modified indirect solar dryer and conventional indirect solar dryer. J Energy Storage 67:107644. https://doi.org/10.1016/j.est.2023.107644

Pandiaraj S, Ayyasamy T, Hasanuzzaman M et al (2022) An experimental investigation on a locally fabricated dryer integrated with a novel solar air heater for the drying of potato slices. Energy Sources, Part A Recover Util Environ Eff 44:9811–9826. https://doi.org/10.1080/15567036.2022.2143967

Pavis S, Henderson SM (1961) Grain drying. Temperature effect on drying coefficients. Agricul Eng Res 6:169–174

Peters T, Sayin L (2021) Status of the global food cold-chain: summary briefing

Poonia S, Singh AK, Jain D (2018) Design development and performance evaluation of photovoltaic/thermal (PV/T) hybrid solar dryer for drying of ber (Zizyphus mauritiana) fruit. Cogent Eng 5:1–18. https://doi.org/10.1080/23311916.2018.1507084

Prasanna YS, Deshmukh SS (2022) Energy, exergy and economic analysis of an air cavity appended passive solar still of different basin material at varying depth. Energy Sustain Dev 71:13–26. https://doi.org/10.1016/j.esd.2022.09.008

Rani P, Tripathy PP (2022) Investigating shrinkage corrected drying characteristics, rehydration, color profile and microstructural evolution during solar drying of pineapple slices. J Food Process Eng 45:1–13. https://doi.org/10.1111/jfpe.14036

Sahdev RK, Kumar M, Dhingra AK (2017) Development of empirical expression for thin layer groundnut drying under open sun and forced convection modes. Agric Eng Int CIGR J 19:152–158

Sahdev RK, Kumar M, Dhingra AK (2017) Effect of mass on convective heat transfer coefficient during open sun drying of groundnut. J Food Sci Technol 54:4510–4516. https://doi.org/10.1007/s13197-017-2858-3

Sehrawat R, Sahdev RK, Tiwari S, Kumar S (2023) Performance analysis and environmental feasibility of bifacial photovoltaic thermal dryer with heat storage. Energy Convers Manag 288:117150. https://doi.org/10.1016/j.enconman.2023.117150

Sethi VP, Arora S (2009) Improvement in greenhouse solar drying using inclined north wall reflection. Sol Energy 83:1472–1484. https://doi.org/10.1016/j.solener.2009.04.001

Shimpy, Kumar M, Kumar A (2023a) Designs, performance and economic feasibility of domestic solar dryers. Food Eng Rev 15:156–186. https://doi.org/10.1007/s12393-022-09323-1

Shimpy, Kumar M, Kumar A (2023b) Performance assessment and modeling techniques for domestic solar dryers. Food Eng Rev 15:525–547. https://doi.org/10.1007/s12393-023-09335-5

Shimpy, Kumar M, Kumar A (2024a) Design and optimization of a domestic solar dryer: an analytical approach. Eng Comput 41:947–965. https://doi.org/10.1108/EC-12-2023-0916

Shimpy, Kumar M, Kumar A (2024b) Experimental and simulation studies on the no-load performance assessment of a domestic solar dryer. Heat Transf 53:1312–1335. https://doi.org/10.1002/htj.22994

Shimpy, Kumar M, Kumar A (2024c) Assessment of a beeswax-packed domestic solar dryer for sustainable bitter gourd drying : an experimental study. Food Bioprod Process 148:72–87. https://doi.org/10.1016/j.fbp.2024.08.007

Shimpy, Kumar M, Kumar A et al (2022a) Comparison of groundnut drying in simple and modified natural convection greenhouse dryers: thermal, environmental and kinetic analyses. J Stored Prod Res 98:101990. https://doi.org/10.1016/j.jspr.2022.101990

Shimpy, Kumar M, Sahdev RK, et al (2022b) Experimental investigations on latent heat storage based modified mixed-mode greenhouse groundnuts drying. J Food Process Preserv 46:1–15. https://doi.org/10.1111/jfpp.16725

Shrivastava V, Kumar A (2017) Embodied energy analysis of the indirect solar drying unit. Int J Ambient Energy 38:280–285. https://doi.org/10.1080/01430750.2015.1092471

Singh P, Gaur MK (2021) Heat transfer analysis of hybrid active greenhouse solar dryer attached with evacuated tube solar collector. Sol Energy 224:1178–1192. https://doi.org/10.1016/j.solener.2021.06.050

Singh P, Gaur MK (2021) Environmental and economic analysis of novel hybrid active greenhouse solar dryer with evacuated tube solar collector. Sustain Energy Technol Assessments 47:101428. https://doi.org/10.1016/j.seta.2021.101428

Singh S, Kumar S (2012) Testing method for thermal performance based rating of various solar dryer designs. Sol Energy 86:87–98. https://doi.org/10.1016/j.solener.2011.09.009

Singh PP, Singh S, Dhaliwal SS (2006) Multi-shelf domestic solar dryer. Energy Convers Manag 47:1799–1815. https://doi.org/10.1016/j.enconman.2005.10.002

Singh P, Gaur MK, Malvi CS (2022) Effect of drying area on heat transfer coefficient and drying kinetics of high-moisture crop dried in a hybrid active greenhouse solar dryer. Heat Transf Res 53:79–97. https://doi.org/10.1615/HEATTRANSRES.2022040797

Singh DD, Sengar M, Patel SK et al (2024) Drying characteristics of thin layer of potato (Solanum tuberosum): experimental and computational studies. Environ Sci Pollut Res 31:38658–38675. https://doi.org/10.1007/s11356-022-25121-y

Sreekumar A, Manikantan PE, Vijayakumar KP (2008) Performance of indirect solar cabinet dryer. Energy Convers Manag 49:1388–1395. https://doi.org/10.1016/j.enconman.2008.01.005

Sundari UA, Neelamegam P, Subramanian CV (2013) An experimental study and analysis on solar drying of bitter gourd using an evacuated tube air collector in Thanjavur, Tamil Nadu, India. Conf Pap Energy 2013:1–4. https://doi.org/10.1155/2013/125628

Terres H, Chavez S, Lopez R, et al (2015) Study of the lemon drying process using a solar dryer. In: ASME 2015 9th International Conference on Energy Sustainability. San Diego, California, pp 1–6

Tiwari GN, Sharma PK, Goyal RK, Sutar RF (1998) Estimation of an efficiency factor for a greenhouse: a numerical and experimental study. Energy Build 28:241–250. https://doi.org/10.1016/s0378-7788(97)00062-5

Tiwari S, Sahdev RK, Kumar M et al (2021) Environmental and economic sustainability of PVT drying system: a heat transfer approach. Environ Prog Sustain Energy 40:1–11. https://doi.org/10.1002/ep.13535

UNEP (2021) Food Waste Index Report 2021. Nairobi

Vijayan S, Arjunan TV, Kumar A (2020) Exergo-environmental analysis of an indirect forced convection solar dryer for drying bitter gourd slices. Renew Energy 146:2210–2223. https://doi.org/10.1016/j.renene.2019.08.066

Wang ZJ, Liu H, Zeng FK et al (2023) Potato processing industry in China: current scenario, future trends and global impact. Potato Res 66:543–562. https://doi.org/10.1007/s11540-022-09588-3

Yang H, Sakai N (2001) Shrinkage and mechanical characteristics of potato undergoing air convection drying. Japan J Food Eng 2:67–72. https://doi.org/10.11301/jsfe2000.2.67

**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.

AIP Conference Proceedings

RESEARCH ARTICLE | JANUARY 16 2025

# Performance of aircraft with varying cant angles and sweep angles

Aman Jha ✉; Aman Ayush; Arun Kumar; M. Zunaid

*AIP Conf. Proc.* 3233, 020011 (2025)
https://doi.org/10.1063/5.0235198

View Online Export Citation

Articles You May Be Interested In

A variable cant angle and winglet span analysis to optimize winglet performance

*AIP Conf. Proc.* (July 2023)

Performance comparison of different winglets by CFD

*AIP Conf. Proc.* (July 2019)

Drag reduction analysis of wing airfoil E562 with forward wingtip fence at cant angle variations of 75° and 90°

*AIP Conference Proceedings* (August 2018)

21 January 2025 07:00:23

AIP Publishing

# Performance of Aircraft with varying cant angles and sweep angles

Aman Jha[1, a)] , Aman Ayush[1] , Arun Kumar[1], M. Zunaid[1]

[1]Department of Mechanical Engineering, Delhi Technological University, Delhi, India

[a)]Corresponding Author Email: amanjha7003@gmail.com

**Abstract:** This study's primary goal is to evaluate how well Winglets function aerodynamically at various sweep and can't angles. By simulating the wing at different sweep angles, can't angles, and angles of attack we can determine the most suitable angles for different flight phases. The purpose of this study is to analyze the properties of a Boeing 737 winglet during cruise flight operations. Different cant angles for the blended winglet attached to the wing were considered. In order to study the winglet aerodynamic characteristics, these models have been theoretically tested at subsonic flow. These characteristics include the coefficients of drag, lift, lift-to-drag ratio, and vorticity and the selection of best sweep and cant angles and then to study the variation of CL/CD with an angle of attack.

Keywords: Vorticity, lift induced drag, coefficient of lift (CL), coefficient of drag (CD), Ansys, Computational Fluid Dynamics (CFD), Angle of attack (AOA), Cant Angles (CA), Sweep Angles (SA), Angle of Sideslip (AOS), Mach Number (Ma).

21 January 2025 07:00:23

## INTRODUCTION

Aviation agencies worldwide are working towards achieving higher aircraft efficiency, recognizing that aviation is a significant contributor to global carbon emissions. To achieve this goal, they have set strict standards and regulations for aircraft manufacturers, pushing them towards developing engines that have better fuel efficiency and aircraft structures that are lighter. One of the myriad ways to achieve better efficiency in airplanes is by minimizing the lift id (LID). Reducing LID is a crucial step in improving overall aircraft efficiency [1-4].

Winglets are one of the most effective ways to reduce wingtip vortices, and as a result, wingtip vortices would be less intense, and the resulting LID would be lessened. However, adding winglets also increases bending moments and may require reinforcing the wings. Winglets are a widely used technology for improving fuel efficiency and performance in aviation, and their drag reduction is equivalent to carrying less payload over a given distance. Adding winglets can result in reduced fuel emissions through cutting down factors such as fuel burn up to 6% [5-7], shorter take-off distance (TOD),) [8-9] and faster climbing rates. They can also improve the aircraft's aesthetic design and aftermarket value.

The designer's primary objective is to reduce id with the least possible increase in profile drag, which increases with higher airspeeds. By studying the airflow around a wing, one can conclude that a longer wingspan results in a reduced tip effect on the inner sections of the wing. As the span becomes infinite, the id and downwash approach zero. Other factors affecting id are the wing planform and its efficiency in lift generation compared to id [10-14].

Large endplates can increase skin friction drag more than they decrease id. In contrast, a winglet actively reduces spanwise flow by carrying an aerodynamic load that interacts with the main wing's flow field. This is accomplished through the downwash produced by the winglet, which opposes the spanwise flow on the primary wing and results in a considerable reduction in spanwise flow [15].

In summary, reducing LID is crucial for improving aircraft efficiency, and winglets are an effective way to achieve this [16-19]. By reducing spanwise flow, winglets can minimize LID without increasing profile drag. As a result, they can help reduce fuel burn, shorten take-off distance, and improve aircraft performance and safety. While adding winglets may increase bending moments and require wing reinforcement, their benefits outweigh the drawbacks making them a widely used technology in aviation today [20], as shown in Fig. 1.

*Recent Advancements in Artificial Intelligence and its Applications in Engineering*
AIP Conf. Proc. 3233, 020011-1–020011-10; https://doi.org/10.1063/5.0235198
Published under an exclusive license by AIP Publishing. 978-0-7354-5098-1/$30.00

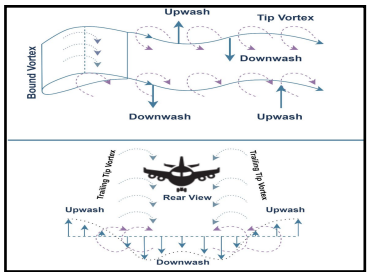


Figure 1: Downwash caused by the formation of wingtip vortices.

## OVERVIEW OF RELATED WORK

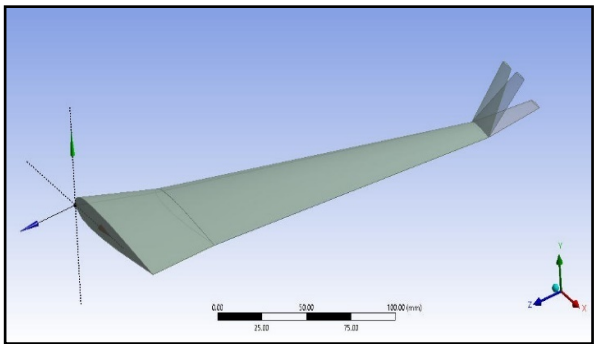
The current winglet design is optimized for a single configuration and may not be the most effective for different flight phases. To address this, a study is conducted to analyze the impact of different winglet CAs and SAs on the performance of a Boeing 737 wing during cruising flight operations. The study uses various parameters such as coefficients of drag, lift, lift-to-drag ratio, and vorticity to determine the optimal angles for different flight phases.[20] This study aims to provide an analysis of the characteristics of the winglet of a Boeing 737 wing at cruising flight operations. By simulating the wing at different sweep angles (SA), cant angles (CA), and angles of attack (AOA) we can determine the most suitable angles for different flight phases. We determine the optimal angles with the help of parameters like the coefficients of drag, lift, lift-to-drag ratio, and vorticity for varying AOA. Despite project reports being done on similar fields of study, studying the effects of different CA as well as SA at varying AOA on a Boeing 737 wing allows for a comprehensive understanding of the aerodynamics and its effect.

## METHODOLOGY

The Boeing 737 wing model was employed in this work. In addition, the standard Boeing 737 wing was made to represent the variable CA winglet (Fig. 2). Additionally, with the SA being defined, we looked into how the winglet's performance in terms of aerodynamics was altered, as shown in Table 1. Table 2 shows the wetted area (W.A) in subsequent analysis, with a Neumann-type boundary condition applied for outflow.

Table 1: Design variables investigated in the project.

| Design Variable | Value |
|---|---|
| Winglet *CA* | 15, 45 ,60 |
| Winglet *SA* | 30 ,45, 60 |
| AOA at Ma = 0.58309 | 0,2,4,6,8,10,12,14,16,18,20 |


Figure 2. Geometry with different *CA*.

21 January 2025 07:00:23

Table 2: The wetted area (W.A) in subsequent analysis, with a Neumann-type boundary condition applied for outflow.

| Wing | W.A ($mm^2$) |
|---|---|
| Original Boeing 737 wing without wingtip | 20460 mm² |
| Winglet *SA* = 30◦ | 21243 mm² |
| Winglet *SA* = 45◦ | 21105 mm² |
| Winglet *SA* = 60◦ | 20920 mm² |

## VALIDATION AND VERIFICATION

The section discusses the use of CFD analysis to calculate pressure difference, LID, and analyze vortex formation on an airfoil and wing. The 3D model of the airfoil and winglet was designed using Solid Works, and the model was imported into ANSYS workbench fluid flow (fluent) and Autodesk CFD module for analysis, as shown in Table 3, 4.

Table 3: Flow test conditions

| ***AOA (deg)*** | ***AOS (deg)*** | ***Ma*** | ***R. N*** |
|---|---|---|---|
| 3.06 | 0.0 | 0.83295 | 11.72E+06 |

Table 4: Design Properties of Onera M6 wing

| | |
|---|---|
| ***Leading edge SA*** | 30 ° |
| ***Trailing edge SA*** | 15◦ |
| ***Span, b*** | 1.1963 meters |
| ***Mean aerodynamic chord, c*** | 0.64607 meters |
| ***Aspect ratio*** | 3.8 |
| ***Taper ratio*** | 0.562 |

Using experimental data that is available for the wing's location at the 1/5th span, the resulting CFD findings will next be verified. The outcomes of NASA's WIND® simulation can also be contrasted with the CFD results. Solidworks was used for creating the required geometry. The ONERA M6 designed in Solid Works is presented in Fig. 3.

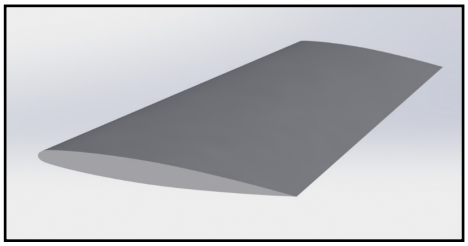


Figure 3. Onera M6 wing model in Solidworks



21 January 2025 07:00:23

*Mesh:*
The mesh used is tetrahedral mesh and the number of nodes present in it is 148544 and the number of meshing elements is 503315, as shown in Fig. 4.

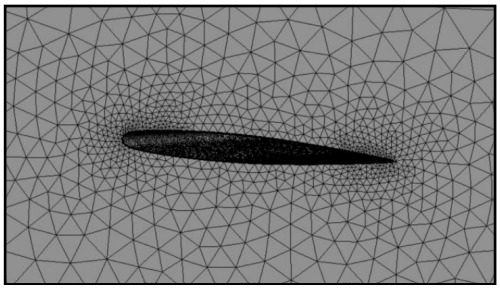


Figure 4. Mesh generated in Ansys

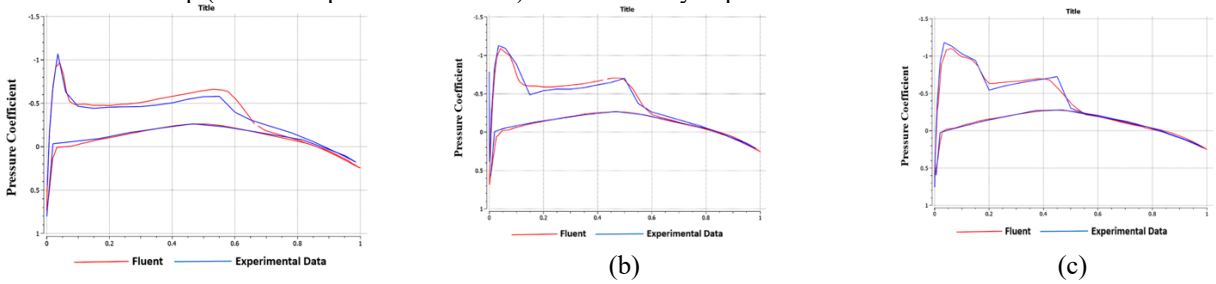
Boundary Conditions:
The study implemented a Pressure far-field boundary condition at the Inlet, Outlet, and Far side of the flow domain while applying a symmetrical boundary condition on the near side. The Wing surface and wing tip were subjected to Wall/Stationary boundary conditions (No-slip condition).
Air was directed through the inlet at an AOA of 3.06° towards the wing, with a pressure of 45.892 psi and a temperature of 460k. The air had a velocity of Ma= 0.8395. The same conditions were applied to the Outlet and Far side.

Verification:
The study involved building a polyline at different locations on a wing and comparing the pressure coefficient (Cp) values obtained from experimental data and CFD simulations. The results were presented in a figure to validate the CFD results with the experimental data. Different polylines obtained from NASA data, reference tests, and Ansys Fluent simulations were compared to validate the Cp CFD results as shown in Fig. 5 and Table 5.

Table 5: Span-wise location

| Section Number | η= Y/b |
|---|---|
| 1 | 0.20 |
| 2 | 0.44 |
| 3 | 0.65 |

1. Validation of Cp (at 1/5th span-wise location) CFD results by experimental data:
2. Validation of Cp (at 0.44ft span-wise location) CFD results by experimental data:
3 Validation of Cp (at 0.65 ft span-wise location) CFD results by experimental data:




(a) (b) (c)

Figure 5 (a+b+c): Variation of Cp at different sections and validating Fluent and experimental results.

The CFD findings showed excellent agreement with the available experimental data.

21 January 2025 07:00:23

## NUMERICAL METHOD & ANALYSIS

The compressible Reynolds-averaged Navier-Stokes (RANS) equations are solved using finite volume Ansys Fluent [16]. The Spalart-Allmaras Turbulence model is used to model turbulence [21, 22].
Before beginning the parametric analysis, we assessed the numerical scheme's accuracy and the mesh's resolution. In this validation study, numerical results were compared with physical test data under similar operating conditions as described in the report [19]. So, the R.N is 11.72x106, the Ma is 0.8395, and the AOA = 3.06°.

## ANALYSIS

*WINGLET GEOMETRY: -*
A blended winglet, similar to the one used by the Boeing 737, was used in this analysis.

*DESIGNING :*

1. Wing model:

The measurements are from the Boeing 737 handbook. The dimensions of the Boeing 737 wings employed in this project's wind design are depicted in the drawings below in the figure. The dimensions were taken from 737 Airplane Characteristics for Airport Planning (2013)

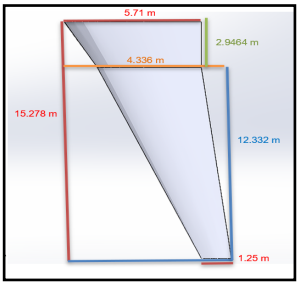


Figure 6: 737 wings with dimensions.

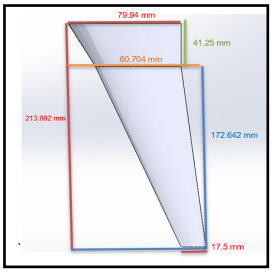


Figure 7: Wing design with dimensions.

Solid work is used for designing wings. The scale has been reduced by considering the dimension at 1 meter = 14mm as the original dimension of the wing is 15.278 meters long and studying it would take a lot of time.
Below is an illustration of the wing design at the new scale:

2. Winglet design:

A blended winglet, similar to those used on Boeing 737 aircraft, is the winglet used in this analysis. The provided figure defines the SA of the winglet. Table 1 lists the values of CA, SA, and AOA used in this study, as shown in Fig. 6-8.

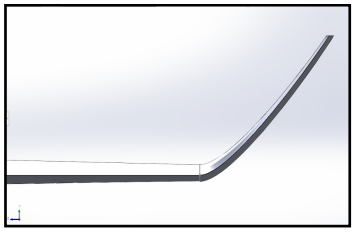


Figure 8: Blended winglet added to the wingtip.

*Boundary Conditions: -*
The computational domain and the arrangement of the boundary conditions are shown in the figure. A Dirichlet-type boundary condition is used to represent the outflow boundary, while a Neumann-type boundary condition is used to



21 January 2025 07:00:23

represent the far-field boundary. In order to model the wing, we used no-slip boundary conditions and continuous wall function boundary conditions for the turbulence variables [23-35], as shown in Fig. 9.

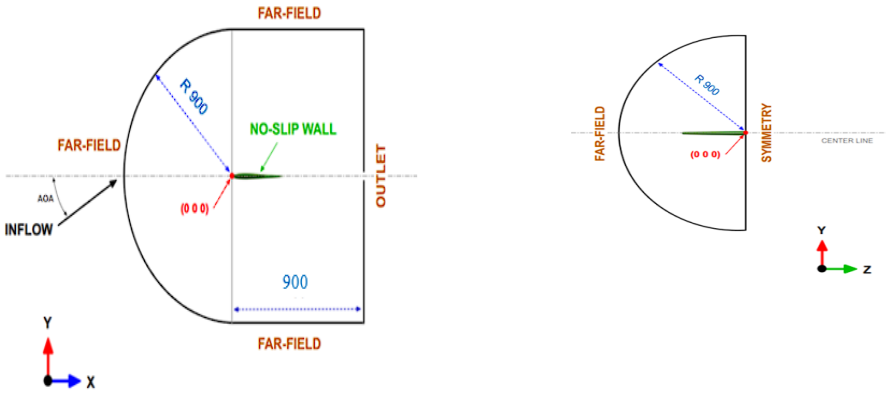



Figure 9: Size of the domain for flow simulation.

*Analysis setup:* -
Velocity = 200 m/s
Flow liquid = Air (Ideal gas)
Density p =1.2256 kg/m^3
Turbulent viscosity ratio = 10 (constant)
Pressure = 0.74 atm = 11.123 psi

*Mesh setup:* -

1. Body sizing
2. Face sizing
3. Inflation layer

Fig. 10 shows the mesh statistics.

| Statistics | |
|---|---|
| Nodes | 335223 |
| Elements | 1160288 |

Figure 10: Mesh statistics.

*Reports:* -

1. Goals: - The aim of this analysis is to calculate the lift and drag coefficient.

- Lift (Newtons)
- Drag (Newtons)
- CL
- CD

2. The equations involved in this analysis are: -

$$Cl = \frac{2 * Lift\ (L)}{(Density * Velocity^2) * Area\ Of\ Wing}$$

$$Cd = \frac{2 * Drag\ (D)}{(Density * Velocity^2) * Area\ Of\ Wing}$$

Where:
Density p = 1.2256 kg/m^3
Velocity = 200 m/s

21 January 2025 07:00:23

Area of wing (wing without winglet) = 20460 mm²

*Result:* -
1. Lift, drag, and CL/CD results.
The following conditions and properties given to fluids are: -
Velocity = 200m/s
AOA = 3.06 º
Density = 1.2256 kg/m^3
Pressure = 0.74 atm = 11.123 psi
Temperature = -55*C or 218.15 K
1.1 Wing without winglet, details shown in Table 6.

Table 6: Analysis Result for wing without winglet

| | |
|---|---|
| ***Lift*** | 74.9773 N |
| ***Drag*** | 4.2714 N |
| ***Cl*** | 0.4110 |
| ***Cd*** | 0.0234 |
| ***Cl/Cd*** | 17.5530 |
| ***Trailing edge velocity*** | 270.562 m/s |

1.2 Winglet with different *CA*s at a fixed *Sa*s, details shown in Table 7-9.

Table 7: Winglet with **30 *SA***

| ***CA*s** | **Trailing Edge Velocities (m/s)** | **Cl/Cd** | **Lift (N)** | **Drag (N)** |
|---|---|---|---|---|
| 15 | 273.493 | 18.2309 | 80.8315 | 4.4337 |
| 45 | 273.363 | 18.0413 | 79.4244 | 4.4023 |
| 60 | 273.975 | 17.8934 | 79.2112 | 4.4268 |

Table 8: Winglet with **45 *SA***

| ***CA*s** | **Trailing Edge Velocities (m/s)** | **Cl/Cd** | **Lift (N)** | **Drag (N)** |
|---|---|---|---|---|
| 15 | 274.144 | 18.0732 | 79.4612 | 4.3966 |
| 45 | 274.647 | 18.0455 | 79.3100 | 4.3950 |
| 60 | 273.402 | 17.7555 | 78.2211 | 4.4054 |

21 January 2025 07:00:23

Table 9: Winglet with **60 *SA***

| ***CA*s** | **Trailing Edge Velocities (m/s)** | **Cl/Cd** | **Lift (N)** | **Drag (N)** |
|---|---|---|---|---|
| 15 | 274.202 | 17.9101 | 77.4532 | 4.3524 |
| 45 | 274.711 | 18.1908 | 78.0916 | 4.3707 |
| 60 | 274.166 | 17.8791 | 78.1943 | 4.3734 |

The result of lift (in N), drag (in N), cl/cd, and trailing edge velocity is calculated with different CAs and a fixed SA.

The study focused on how the addition of wings to a wing design can increase speed compared to a wing without wings. The presence of additional fluid particles deflected from the wing surface leads to regions of higher velocity on the upper as well as lower surfaces of the wing and aircraft. Due to the shape of the airfoil of the wing, air moves faster over the upper surface, creating a high-velocity region on the backplane. The wing divides airflow between both surfaces, resulting in higher velocities in the low-pressure region based on the Bernoulli force. The interface of different pressures behind the wing produces vortices that cause pressure drag, which can be reduced to increase speed and improve performance. The design of the wing with wings increased the speed value by 1.51% compared to the wingless wing by reducing pressure drag. The best result was achieved at 60° SA and 45°CA, with CL/CD of 18.1908 and trailing edge velocity of 274.71 m/s, as per the calculations.

*Grid independent Test: -*
To ensure that the results were not dependent on the mesh used, a thorough grid independence study was carried out. The grid independence test had two primary objectives:

- To obtain a converged solution that was independent of the mesh used and solely based on the boundary conditions and physics involved.
- Optimize the number of cells used in the computational domain to ensure the best possible results.

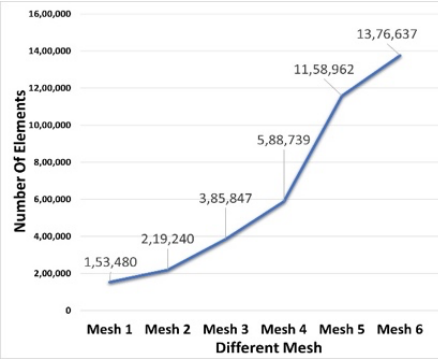



Figure 11: Variation in the number of elements in different types of mesh

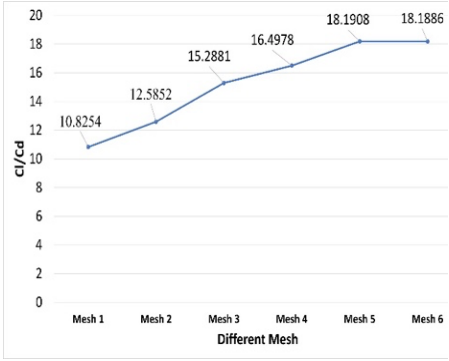



Figure 12: Variation of Cl/Cd with different mesh types

After finding optimal CL/CD ratio and trailing edge velocity i.e., SA of 60° and a CA of 45° CL/CD ratio was calculated in different types of mesh, as shown in Figs 11-12.
For the above analysis, we kept the AOA constant to find the optimal value of the CL/CD ratio. The result was then calculated keeping Pressure= 11.123 Psi, Temperature =218.15 K, and Mach no= 0.58309 as shown in the figure 13.

21 January 2025 07:00:23

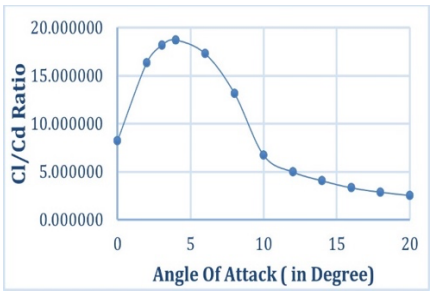


Figure 13: Variation of Cl/Cd with changing AOA

After finding the best optimal cant and SA, AOA has also been varied in this case to find the best CL/CD ratio. The angle has been varied from 0∘ to 20 ∘ (spaced at intervals of 2).
By observing the graph above, we can conclude that the angle of attack that yields the highest CL/CD ratio of 18.7085 is at 4º.

## CONCLUSION

Winglets are designed to reduce LID and enhance aircraft performance and fuel economy. This study focuses on the cruising flight operation parameters to determine the best sweep and can't angles for the winglet to minimize lift-induced drag. The study uses simulations at different AOA, SA, and CA to determine the most suitable angles for different flight phases. The results show that winglets at the wingtip improved aerodynamic efficiency when measured in terms of Cl/Cd, and reduced air vortex intensity behind the wing.
The study finds that significant drag reductions for the same lift value are achieved by cautiously regulating the winglet can't and sweep angles. The aerodynamic performance gains made possible by adopting the suggested winglet are positively impacted by large winglet SA values. The analysis results show that the airplane performed best during cruising conditions with winglets having a CA of 45 degrees and a SA of 65 at the AOA.
This study can serve as a reference for future research on winglets, particularly if a mechanical device is developed that can adjust the winglets' cant and sweep angles during different phases of flight. There is still scope for further work in this area, such as exploring additional cant and sweep angles and determining the optimal cant and sweep angles for different flight operating conditions, such as take-off, climb, and landing. Ultimately, improving the efficiency of winglets during different flight phases will have a significant impact on aircraft performance and fuel economy.

## REFERENCES

1. Board, Air Force Studies, and National Research Council. Assessment of wingtip modifications to increase the fuel efficiency of air force aircraft. National Academies Press, 2007..
2. Yin, Kwong-sang. "A study of carbon dioxide emissions reduction opportunities for airlines on Australian international routes." (2017).
3. Guerrero JE, Maestro D, Bottaro A (2012) Biomimetic spiroid winglets for lift and drag control. Comptes Rendus Mec 340:67-80
4. Myilsamy, Dinesh, Yokesh Thirumalai, and P. S. Premkumar. "Performance Investigation of an Aircraft Wing at Various Cant Angle of Winglet Using CFD Simulation." In Altair Technology Conference India. 2015.
5. Azlin, M. An, CF Mat Taib, Salmiah Kasolang, and F. H. Muhammad. "CFD analysis of winglets at low subsonic flow." In Proceedings of the World Congress on Engineering, vol. 1, pp. 6-8. 2011.
6. Anderson, John. EBOOK: Fundamentals of Aerodynamics (SI units). McGraw hill, 2011..
7. Jupp, Jeff. "Wing aerodynamics and the science of compromise." The Aeronautical Journal 105, no. 1053 (2001): 633-641.
8. Jupp, Jeff. "Wing aerodynamics and the science of compromise." The Aeronautical Journal 105, no. 1053 (2001): 633-641.
9. McLean, Doug. "Wingtip devices: what they do and how they do it." In Boeing performance and flight operations engineering conference. 2005.
10. Chambers, Joseph R. Concept to reality: contributions of the Langley Research Center to US Civil Aircraft of the 1990s. No. NASA/SP-2003-4529. 2003.

21 January 2025 07:00:23

11. Whitcomb, Richard T. A design approach and selected wind tunnel results at high subsonic speeds for wing-tip mounted winglets. No. L-10908. 1976.
12. Gomez-Tamm, A. E., P. Ramon-Soria, B. C. Arrue, and A. Ollero. "Current state and trends on bioinspired actuators for aerial manipulation." In 2019 Workshop on Research, Education and Development of Unmanned Aerial Systems (RED UAS), pp. 352-361. IEEE, 2019.
13. Prasad, Kondapalli Siva, Vommi Krishna, and BB Ashok Kumar. "Aerofoil profile analysis and design optimisation." J Aerosp Eng Technol 3 (2013): 1-10.
14. Ansys Academic Research, Release 19, Help System, Ansys Fluent Theory Guide; ANSYS, Inc.: Canonsburg, PA, USA, 2018.
15. Wilcox, David C. Turbulence modeling for CFD. Vol. 2. La Canada, CA: DCW industries, 1998.
16. Huang, Junji, Jorge-Valentino Bretzke, and Lian Duan. "Assessment of turbulence models in a hypersonic cold-wall turbulent boundary layer." Fluids 4, no. 1 (2019): 37.
17. Schmitt, V. "Pressure distributions on the ONERA M6-wing at transonic mach numbers, experimental data base for computer program assessment." AGARD AR-138 (1979).
18. Junk, Michael, and Zhaoxia Yang. "Outflow boundary conditions for the lattice Boltzmann method." Progress in Computational Fluid Dynamics, an International Journal 8, no. 1-4 (2008): 38-48.
19. C√©ron-Mu√±oz, Hern√°n Dario, Renato Cosin, Rog√©rio Frauendorf de Faria Coimbra, Luiz Guilherme Noujain Correa, and Fernando Martini Catalano. "Experimental investigation of wing-tip devices on the reduction of induced drag." Journal of Aircraft 50, no. 2 (2013): 441-449.
20. Guerrero, Joel E., Marco Sanguineti, and Kevin Wittkowski. "Variable cant angle winglets for improvement of aircraft flight performance." Meccanica 55 (2020): 1917-1947.
21. Panagiotou, Periklis, S. Antoniou, and K. Yakinthos. "Cant angle morphing winglets investigation for the enhancement of the aerodynamic, stability and performance characteristics of a tactical Blended-Wing-Body UAV." Aerospace Science and Technology 123 (2022): 107467.
22. Kazim, Ali Hussain, Abdullah Hamid Malik, Hammad Ali, Muhammad Usman Raza, Awais Ahmad Khan, Tauseef Aized, and Aqsa Shabbir. "CFD analysis of variable geometric angle winglets." Aircraft Engineering and Aerospace Technology 94, no. 2 (2022): 289-301.
23. Raja, Bino Prince D., G. Ramanan, and Diju G. Samuel. "Computational analysis of blended winglet model performance by varying cant angle." Journal of Computational and Theoretical Nanoscience 16, no. 2 (2019): 467-471.
24. Smith, Marilyn, Narayanan Komerath, Re Ames, O. Wong, and J. Pearson. "Performance analysis of a wing with multiple winglets." In 19th AIAA Applied Aerodynamics Conference, p. 2407. 2001.
25. Seshaiah, T., B. Vasu, K. Vijaya Kumar Reddy, and P. Bridjesh. "Analysis on air craft winglet at different angles by using CFD simulation." Materials today: Proceedings 49 (2022): 275-283.

21 January 2025 07:00:23

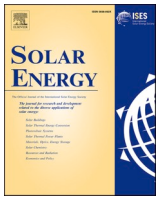
Solar Energy 287 (2025) 113253



Contents lists available at ScienceDirect

Solar Energy

journal homepage: www.elsevier.com/locate/solener





# Performance optimization of lead-free $KGeCl_3$ based perovskite solar cells using SCAPS-1D

Rahul Kundara *, Sarita Baghel

*Department of Applied Physics, Delhi Technological University, Delhi, India*

ARTICLE INFO

Handling editor: 'Ranga Pitchumani'

*Keywords:*
$KGeCl_3$-based PSC
SCAPS-1D
ETLs and HTLs

ABSTRACT

The commercial application of lead-based Perovskite Solar Cells (PSCs) is restrained by their toxicity. In this work, we have investigated $KGeCl_3$ which is a potential candidate for non-lead PSC. SCAPS-1D software is employed to examine the impact of distinct Hole Transport Layers (HTLs), absorber layer and Electron Transport Layers (ETLs) to maximize the device performance. The effect of absorber layer thickness, its acceptor and donor concentration ($N_A$ and $N_D$) have been analyzed. We have obtained a maximum efficiency of 29.82 % with open circuit voltage ($V_{OC}$) = 1.19 V, short circuit current density ($J_{SC}$) = 30.08 mA/cm$^2$, and fill factor (FF) = 82.74 % for the configuration of FTO/$WS_2$/$KGeCl_3$/PEDOT: PSS/Au at radiative recombination coefficient of $10^{-12}$ cm$^3$/sec, operating temperature of 300 K and defect density of $10^{15}$ cm$^{-3}$. The optimized thickness of the absorber layer is 600 nm. Thus, the investigation suggests that $WS_2$ as ETL and PEDOT: PSS as HTL is most suitable for the Ge-based PSC. Hence, this configuration is most suitable to manufacture highly efficient PSC devices.

## 1. Introduction

Global warming, rapid industrialization and growing population will drive the demand for renewable sources of energy which should also be environment friendly. In this scenario solar energy devices offer a clean and sustainable solution for the increase in energy demand. Currently Silicon based Photovoltaic (PV) modules dominate the solar cell market due to their stable and efficient performance. But the Silicon technology is not completely environment friendly when analyzed through its complete life cycle assessment. They are difficult to recycle and they have high carbon emission during the production stage. However, perovskite solar cells (PSCs) are the emerging class of photovoltaics due to its promising properties such as tunable bandgap ($E_g$), high absorption coefficient, low manufacturing cost and low temperature fabrication. Perovskite material has an $ABX_3$ structure (where A is an organic or inorganic cation, B is an inorganic metal cation and X is halogen anion) [1]. In the typical structure of a PSC, the absorber layer is positioned between Electron Transport Layer (ETL) and Hole Transport Layer (HTL). The generation of photocurrent occurs through injection of photoelectrons into the ETL and the transfer holes through the HTL from the absorber layer. Therefore, ensuring appropriate alignment of the energy band levels across the layers is essential for achieving optimal device performance.

In 2009, PSC was reported with an efficiency of 3.8 % and has achieved maximum power conversion efficiency (PCE) approximately 26 % till now for $CH(NH_2)_2PbI_3$ and $CH_3NH_3PbI_3$ perovskite materials [2,3,4,5]. Recently various hybrid organic–inorganic perovskite materials have been reported which possess high charge carrier mobilities and also offer low temperature solution-based fabrication techniques [6,7,8]. They possess some challenges related to stability, toxicity and scalability [9]. The inorganic PSCs, particularly those employing materials like cesium lead halides have advantages over hybrid PSCs as follows: improved stability, scalability and non-toxic alternatives [10,11] but they have phase stability issues and complexity in fabrication [12]. It is mostly lead (Pb)-based PSCs [3,4,5] that have demonstrated experimental PCE of 20.17 % for $CsPbI_3$ [13]. But, despite the high performance of Pb based PSC, its stability is a major challenge. There are many degradation reactions which impact the PV performance and device stability. There is instability due to rapid oxidation of Pb cation, interfacial reactions of Pb with other materials in the device [14]. Apart from this toxicity concerns also need to be addressed. Though Pb is a toxic substance, there are recent studies which indicate that its use in PSC is limited to less than 1 gm/m$^2$ which can be easily managed by proper encapsulation and appropriate recycling systems [15]. On the other hand, the effect of toxicity of lead versus non-lead PSCs have also been investigated [16] and the studies show that, due

* Corresponding author.
*E-mail address:* rahul_2k20phdap10@dtu.ac.in (R. Kundara).

https://doi.org/10.1016/j.solener.2025.113253
Received 13 March 2024; Received in revised form 19 December 2024; Accepted 3 January 2025
Available online 7 January 2025
0038-092X/© 2025 International Solar Energy Society. Published by Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

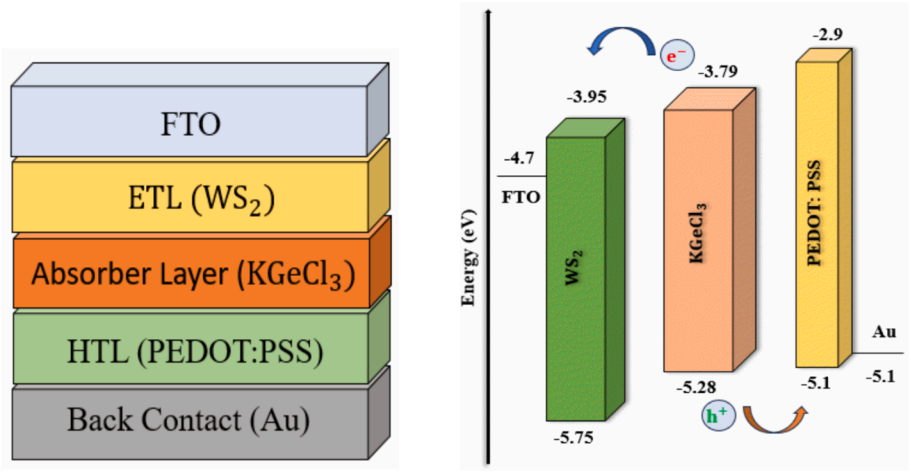



**Fig. 1.** Architecture and band alignment diagram for $KGeCl_3$ PSC.

**Table 1**
Input parameters were used in simulation of distinct ETLs.

| Parameter | CdS [27] | PCBM [28,29] | $C_{60}$ [30] | $WS_2$ [31] | $KGeCl_3$ [32] |
|---|---|---|---|---|---|
| Thickness (μm) | 0.050 | 0.500 | 0.050 | 0.100* | 0.300 |
| Eg (eV) | 2.42 | 2.1 | 1.7 | 1.8 | 1.49 |
| χ (eV) | 4.3 | 3.9 | 3.9 | 3.95 | 3.52 |
| $\varepsilon_r$ | 9.35 | 3.9 | 4.2 | 13.6 | 3.52 |
| $N_C$ ($cm^{-3}$) | $2.2 \times 10^{18}$ | $2.2 \times 10^{19}$ | $8.0 \times 10^{19}$ | $1 \times 10^{18}$ | $1 \times 10^{18}$ |
| $N_V$ ($cm^{-3}$) | $1.8 \times 10^{19}$ | $2.2 \times 10^{19}$ | $8.0 \times 10^{19}$ | $2.4 \times 10^{19}$ | $1 \times 10^{19}$ |
| $\mu_n$ ($cm^2$/Vs) | 100 | 0.001 | $8.0 \times 10^{-2}$ | 100 | 20 |
| $\mu_p$ ($cm^2$/Vs) | 25 | 0.002 | $3.5 \times 10^{-3}$ | 100 | 20 |
| $N_D$ ($cm^{-3}$) | $1.15 \times 10^{17}$ | $1 \times 10^{19}$ | $2.6 \times 10^{17}$ | $1 \times 10^{18}$ | $1 \times 10^{15}$ |
| $N_A$ ($cm^{-3}$) | 0 | – | – | 0 | $1 \times 10^{15}$ |
| $N_t$ ($cm^{-3}$) | $1 \times 10^{15*}$ | $1 \times 10^{9*}$ | $1 \times 10^{14}$ | $1 \times 10^{13*}$ | $1 \times 10^{13}$ |

*In this work.

to the toxic nature of lead, there is a need to find environmentally friendly substitutes for it. However, before doing so, one has to consider the results of life cycle assessment studies of available alternative materials. There have been various investigations to substitute hazardous Pb ions with other alternatives such as Sn [17], Ge [18], Bi [19], Sb [20] and Ag [21]. But most of the alternatives do not offer substantial advantages in terms of cost, toxicity or environmental safety. Recently reported DFT study shows that $KGeCl_3$ is a promising perovskite material for PV application [Namisi et al. 2023]. There are very few experimental or theoretical investigations on $KGeCl_3$ for solar cell application so far. The study comprises the application of $KGeCl_3$ as a perovskite material and simulated the PSC using SCAPS-1D. The recently reported studies demonstrated that $KGeCl_3$-based PSC achieved a PCE of 15.83 % [22] and 19.62 % [23] using SCAPS-1D. Previously germanium-based double PSC has been simulated and obtained a maximum PCE of 22.5 % [24] whereas $MAGeI_3$-based PSC has achieved a theoretical PCE of 25.34 % [25]. SnGe-based PSC has been theoretically obtained at an efficiency of 24.20 % [26]. We have simulated the various ETLs (CdS, PCBM, $C_{60}$ and $WS_2$) and HTLs (CuI, Spiro-OMeTAD, $Cu_2O$, PEDOT: PSS and $MoS_2$) and analyzed the effects on the PV parameters of PSC. The optimization of thickness of absorber layer, operating temperature, energy band gap ($E_g$) and carrier concentration ($N_D$, $N_A$) have been done to achieve the maximum PCE. The simulated $KGeCl_3$ device achieved a highest PCE of 29.82 % (with $V_{OC} = 1.19$ V, $J_{SC} = 30.08$ mA/$cm^2$, and FF = 82.74 %) at 300 K with $WS_2$ as ETL and PEDOT: PSS as HTL. These photovoltaic outcomes showed that $KGeCl_3$ is promising material to fabricate lead-free low cost and efficient PSCs.

## 2. Material and Methods

### 2.1. Device architecture

The Fig. 1 depicts the device architecture and band alignment of the $KGeCl_3$-based PSC. The PSC structure typically includes distinct layers, an absorber layer situated between ETL and HTL. The gold (Au) is chosen as a back contact electrode with a work function of 5.1 eV. During the simulation of the Ge-based PSC, different ETLs and HTLs are simulated to determine the optimal combination that yields the maximum PCE while also optimizing the thickness of the absorber layer. The thickness of various ETLs (CdS, PCBM, $C_{60}$ and $WS_2$) varies from 50 nm to 500 nm shown below in Table 1. The employed $KGeCl_3$ absorber

**Table 2**
Input parameters used in simulation of various HTLs.

| Parameter | FTO [33,34] | CuI [35] | Spiro-OMeTAD [36,37] | $Cu_2O$ [33,35] | PEDOT: PSS [28,30] | $MoS_2$ [38] |
|---|---|---|---|---|---|---|
| Thickness (μm) | 0.400 | 0.100 | 0.213 | 0.250 0.080 | | 0.200 |
| Eg (eV) | 3.5 | 3.1 | 3 | 2.17 2.2 | | 1.29 |
| χ (eV) | 4.0 | 2.1 | 2.2 | 3.2 2.9 | | 4.2 |
| $\varepsilon_r$ | 9.0 | 6.5 | 3 | 7.11 3.0 | | 3.0 |
| $N_C$ ($cm^{-3}$) | $2.02 \times 10^{18}$ | $2.8 \times 10^{19}$ | $1 \times 10^{19}$ | $2.02 \times 10^{17}$ $2.2 \times 10^{15}$ | | $2.2 \times 10^{18}$ |
| $N_V$ ($cm^{-3}$) | $1.8 \times 10^{19}$ | $1 \times 10^{19}$ | $1 \times 10^{19}$ | $1.1 \times 10^{19}$ $1.8 \times 10^{18}$ | | $1.8 \times 10^{19}$ |
| $\mu_n$ ($cm^2$/Vs) | 20 | 100 | $10^{-4}$ | 200 $1 \times 10^{-2}$ | | 100 |
| $\mu_p$ ($cm^2$/Vs) | 10 | 43.9 | $10^{-4}$ | 80 $2 \times 10^{-4}$ | | 150 |
| $N_D$ ($cm^{-3}$) | $2 \times 10^{19}$ | – | – | – – | | $1 \times 10^{15}$ |
| $N_A$ ($cm^{-3}$) | 0 | $1 \times 10^{18}$ | $2 \times 10^{19}$ | $1 \times 10^{18}$ $10^{19}$ | | $1 \times 10^{17}$ |
| $N_t$ ($cm^{-3}$) | $1 \times 10^{15}$ | $1 \times 10^{14}$ | $1 \times 10^{14}$ | $1 \times 10^{14}$ $1 \times 10^{14}$ | | $1 \times 10^{14}$ |

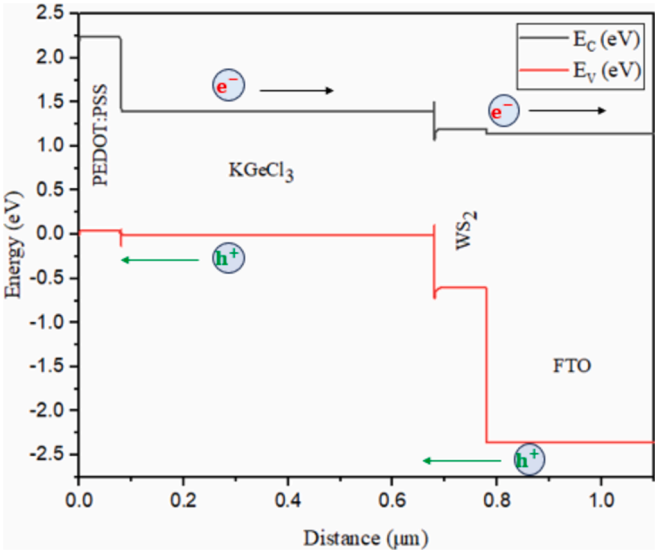



**Fig. 2.** Energy band diagram of FTO/$WS_2$ /$KGeCl_3$/PEDOT: PSS/Au PSC.

**Table 3**
$KGeCl_3$ PSC performance before optimization (at thickness of absorber layer = 300 nm, $E_g$ = 1.49 eV, $N_D = N_A = 1 \times 10^{14}$ $cm^{-3}$ and $N_t = 10^{13}$ $cm^{-3}$) with different ETLs and HTLs.

| Device architecture | $V_{OC}$ (V) | $J_{SC}$ (mA/$cm^2$) | FF (%) | PCE (%) |
|---|---|---|---|---|
| FTO/$C_{60}$/$KGeCl_3$/ Spiro-OMeTAD/Au | 1.30 | 22.60 | 80.39 | 23.63 |
| FTO/$C_{60}$/$KGeCl_3$/ PEDOT: PSS/Au | 1.30 | 22.67 | 80.49 | 23.74 |
| FTO/$C_{60}$/$KGeCl_3$/ $Cu_2O$/Au | 1.30 | 22.80 | 80.56 | 23.90 |
| FTO/$WS_2$/$KGeCl_3$/$MoS_2$/Au | 1.02 | 23.04 | 81.11 | 19.09 |
| FTO/$WS_2$/$KGeCl_3$/CuI/Au | 1.31 | 23.10 | 84.41 | 25.57 |
| FTO/$WS_2$/$KGeCl_3$/Spiro-OMeTAD/Au | 1.31 | 23.10 | 84.22 | 25.51 |
| **FTO/$WS_2$/$KGeCl_3$/PEDOT: PSS/Au** | 1.31 | 23.15 | 84.36 | **25.61** |
| FTO/$WS_2$/$KGeCl_3$/CuI/Au | 1.31 | 23.10 | 84.41 | 25.57 |
| FTO/CdS/$KGeCl_3$/ PEDOT: PSS/Au | 1.30 | 22.50 | 74.98 | 22.05 |
| FTO/$WS_2$/$KGeCl_3$/CBTS/Au | 1.31 | 23.20 | 83.92 | 25.52 |
| FTO/$WS_2$/$KGeCl_3$/CFTS/Au | 1.12 | 25.34 | 87.86 | 25.01 |
| FTO/PCBM/$KGeCl_3$/$Cu_2O$/Au | 1.28 | 13.61 | 84.24 | 14.73 |

layer with thickness of 300 nm shown below in Table 1. Various HTLs are used to enhance the PCE of the PSC, including CuI, Spiro-OMeTAD, $Cu_2O$, PEDOT: PSS and $MoS_2$ with thickness varying from 80 nm to 250 nm shown below in Table 2. The band alignment diagram of the optimized device shown below in Fig. 2 illustrates the transport of charge carriers, specifically the movement of photo-generated electrons ($e^-$s) and holes ($h^+$s) across the device. Electrons have propensity to migrate from higher to lower energy levels, transitioning from PEDOT: PSS ($E_C \sim$ 2.2 eV) to $KGeCl_3$ (with $E_C$ ranging from 0.05 to 1.50 eV), then $WS_2$ ($E_C \sim$ 1.2 eV), and ultimately to FTO ($E_C \sim$ 1.1 eV). Likewise, hole transport is facilitated by the configuration of valence levels. Holes have ability to migrate freely from FTO ($E_V \sim -2.36$ eV) to $WS_2$ ($E_V \sim -0.60$ eV), then to $KGeCl_3$ ($E_V \sim 0$ eV), and finally to PEDOT: PSS ($E_V \sim 0.04$ eV). At the interface of $WS_2$/$KGeCl_3$, a moderate upward cleavage is observed. Instead of being detrimental, this phenomenon acts as a barrier, preventing excessive photo-generated $e^-$s and $h^+$s from flowing through the conduction and valence bands, respectively. Such regulation is crucial as it helps prevent rapid current saturation. The simulated performance data of $KGeCl_3$ PSC before optimization with distinct ETLs and HTLs demonstrated in Table 3.

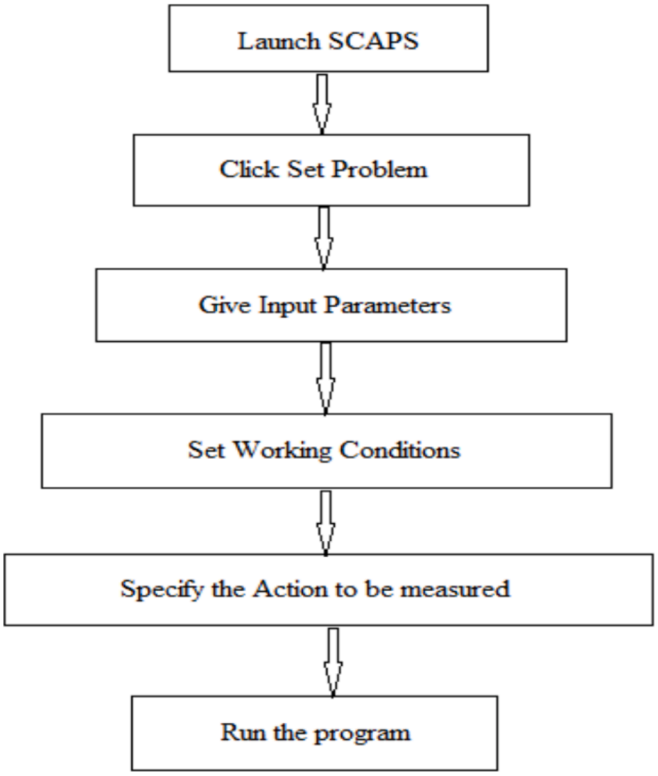



**Fig. 3.** Flow chart of simulation procedure.

### 2.2. Simulation Method

The data of different parameters are chosen from previous studies. The SCAPS-1D (version 3.3.10) has been employed for simulation. This software can be used for simulations at different input physical conditions. This tool calculation is dependent on differential equations, including Poisson's and continuity equations. It can calculate Quantum efficiency (QE), J-V curve and carrier lifetime for both $e^-$s and $h^+$s.

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{q}{\epsilon}[p(x) - n(x) + N_D - N_A + \rho_p - \rho_n = 0 \quad (1)$$

$$\frac{1}{q}\frac{dJ_p}{dx} = G_{op}(x) - R(x) \quad (2)$$

$$\frac{1}{q}\frac{dJ_n}{dx} = -G_{op}(x) + R(x) \quad (3)$$

The simulations were run at 300 K using a "air mass 1.5 global" spectrum and a light power of 1000 W/$m^2$. The work is dependent on change in thickness, $E_g$, $N_D$, and $N_A$ at which simulation of the proposed PSC has been done. The outcome of the device has been reflected by the I–V characteristics which represent the following PV parameters: $J_{SC}$, FF, $V_{OC}$, and PCE. The simulation procedure is depicted in Fig. 3.

## 3. Results and discussion

### 3.1. Impact of distinct parameters on $KGeCl_3$-based PSC

#### 3.1.1. Effect of bandgap with thickness

This section includes effect of the absorber layer thickness, $E_g$, donor and acceptor carrier concentration ($N_D$ and $N_A$) on PV parameters ($V_{OC}$, $J_{SC}$, FF and PCE) of $KGeCl_3$-based PSC where thickness of absorber layer, bandgap, $N_D$ and $N_A$ are fixed at 0.3 μm, 1.49 eV and $10^{15}$ $cm^{-3}$ respectively. The contour plot shown below in Fig. 4 demonstrates that change in PV parameters of simulated PSC as a function of thickness of

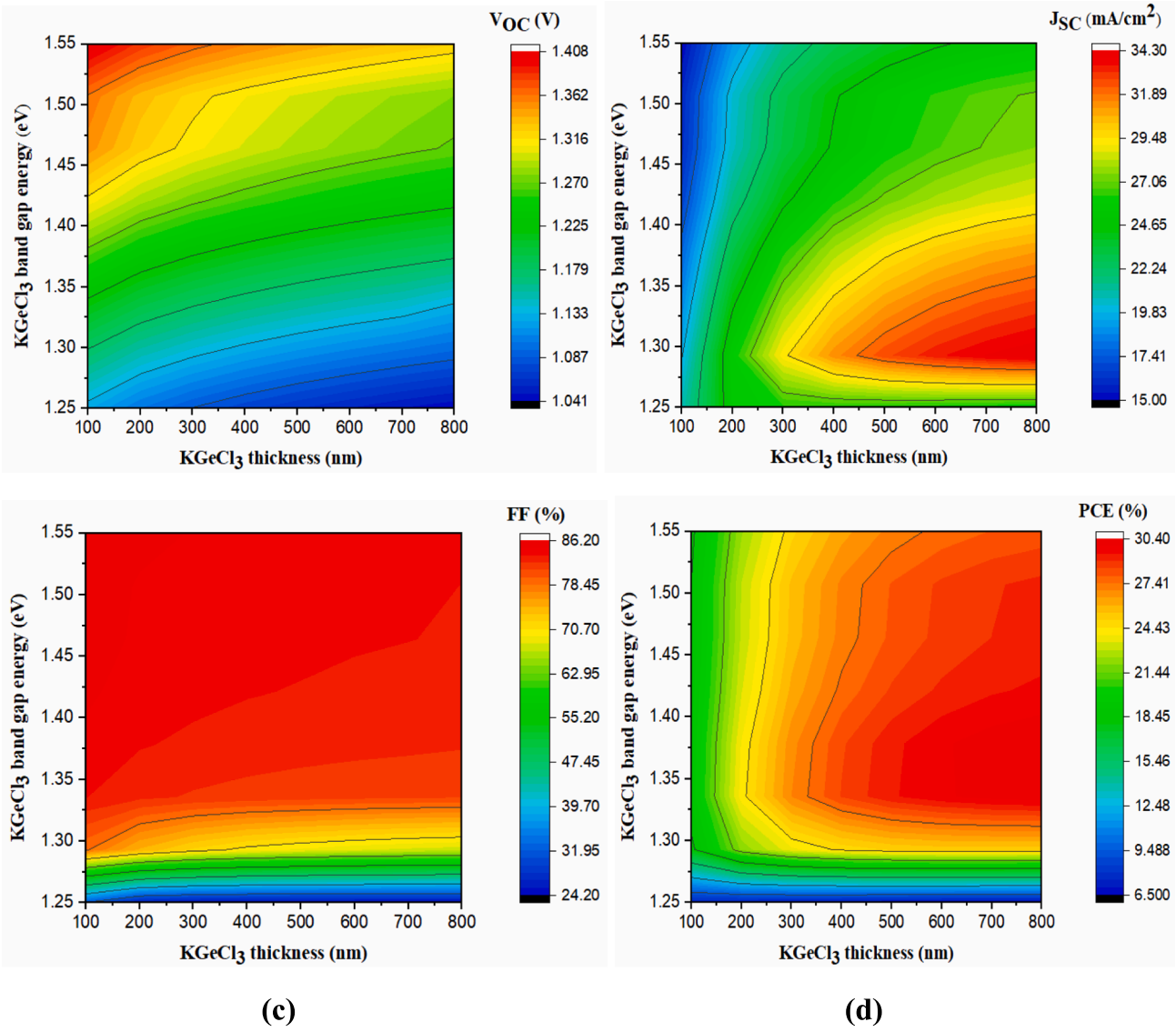



**Fig. 4.** Change in PV parameters of simulated PSC such as (a) $V_{OC}$ (b) $J_{SC}$ (c) FF and (d) PCE with thickness of absorber layer and bandgap at doping concentration of $10^{15}$ $cm^{-3}$.

$KGeCl_3$ and its $E_g$ change from 100 to 800 nm and from 1.25 to 1.55 eV respectively. Fig. 4(a) shows that $V_{OC}$ increases from 1.041 V at $E_g$ of 1.25 eV to about 1.408 V at $E_g$ of 1.55 eV while the change in $V_{OC}$ with thickness rises from 100 to 800 nm. For Fig. 4 (a), a higher $E_g$ leads to a higher $V_{OC}$, it allows a larger separation between the quasi-fermi levels of electrons and holes. It reduces thermal carrier generation and intrinsic recombination rates. The decrease in $V_{OC}$ with increasing temperature can be attributed to the creation of more interfacial defects, an increase in series resistance, and a reduction in carrier diffusion length [39,40]. The value of $J_{SC}$ falls from 34.30 to 15.09 mA/$cm^2$ with an increment in value of 1.25 to 1.55 eV depicted in Fig. 4(b). For Fig. 4 (b), $J_{SC}$ has low values as $E_g$ increases for the thickness between 100 to 300 nm. This is perhaps due to low charge carrier generation due to very small value of thickness. High value of $J_{SC}$ obtained for thickness between 600 to 700 nm and at $E_g$ of 1.30 eV. For the as $E_g$ greater than 1.30 eV causes decreasing the absorption within the layer and hence $J_{SC}$ decreases [41]. The value of $J_{SC}$ increases for each thickness of $KGeCl_3$, ranging from 100 to 800 nm.

The Fig. 4(c) indicate that the FF was increased from 31.69 to 86.02 % with increment in the value of $E_g$ from 1.25 to 1.55 eV and this with simultaneous increment in $KGeCl_3$ thickness from 100-800 nm with approximate decrement of 2.59 % in value of FF. The Fig. 4(d) depicted that the value of PCE was 6.89 % at $E_g$ of 1.25 eV with increment in thickness of $KGeCl_3$ from 100-800 nm at various regions. An average PCE of approximately 23.30 % is attained across variations of $E_g$ ranging from 1.25 to 1.55 eV. Additionally, within a thickness range of 500 to 800 nm, a higher performance level of about 28.85 % is achieved. These findings can be elucidated as follows: with an increase in the $E_g$, there is a concurrent rise in the local collection of light absorption within the $KGeCl_3$ thin film. These outcomes can be clarified by observing that as the $E_g$ increases, the local collection efficiency of light absorption within the absorber layer increases. This enhancement augments the carrier generation rate, consequently leading to a dramatic increase in the $V_{OC}$, as indicated by the equation [42].

$$V_{oc} = \frac{K_B T}{q} \ln(1 + I_{Ph} I_0) \tag{4}$$

Where $V_{OC}$ is open circuit voltage, $K_B$ is Boltzmann constant, T is temperature in K, q is electronic charge, $I_{Ph}$ is Photogenerated current and $I_0$ is reverse saturation current. The outcomes suggest that to maximize the PCE in PSC, it is optimal to have a $KGeCl_3$ $E_g$ of 1.40 eV and a thickness ranging from 500 to 800 nm. JV plots were simulated for PSCs varying

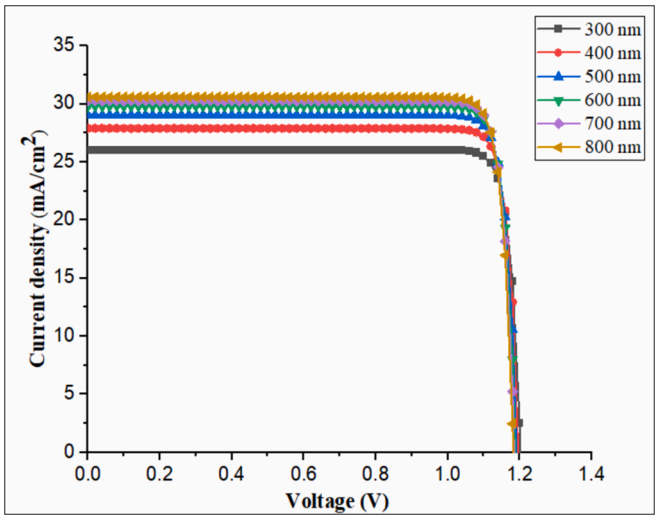



**Fig. 5.** The JV curves of the PSC with thickness of the absorber layer.

$KGeCl_3$ thickness ranging from 300 to 800 nm while maintaining a fixed $E_g$ of 1.40 eV, as depicted in Fig. 5. It is suggested that the $E_g$ should be tuned using suitable dopants to get this optimal value of band gap. These results reveal the following trends: $V_{OC}$ ranges between 1.18 V and 1.20 V, the $J_{SC}$ changes from 26.05 to 30.60 mA/cm$^2$, FF remains constant at 89 % and PCE ranges from 28.06 to 29.82 % with increase in thickness of $KGeCl_3$ thickness. Notably, the PCE trend closely follows the variations in $V_{OC}$ and FF values.

### 3.1.2. *Effect of donor carrier concentration ($N_D$) with thickness*

The Fig. 6 demonstrates the contour plot depicting the PV parameters as functions of the $KGeCl_3$ layer thickness varies 100 to 800 nm (x-axis) and $N_D$ ranging between $1 \times 10^{14}$ to $1 \times 10^{21}$ cm$^{-3}$ (y-axis) respectively. From Fig. 6(a), it is revealed that the $V_{OC}$ increases from approximately 0.96 to 1.41 V as the $N_D$ rises from $1 \times 10^{14}$ to $1 \times 10^{21}$ cm$^{-3}$, with major fluctuations in the $V_{OC}$ values observed with increasing $KGeCl_3$ thickness. According to Fig. 6(b), the $J_{SC}$ value increases approximately 17.72 to 31.06 mA/cm$^2$ as $N_D$ rises from $1 \times 10^{14}$ to $1 \times 10^{19}$ cm$^{-3}$ and $J_{SC}$ remain constant at about 28.48 mA/cm$^2$ for $KGeCl_3$ thickness exceeding 400 nm. Conversely, for $KGeCl_3$ thickness below 400 nm, $J_{SC}$ maintains around 21.90 mA/cm$^2$ below $N_D$ of $1 \times 10^{20}$ cm$^{-3}$. Above $N_D$ of $1 \times 10^{21}$ cm$^{-3}$, the value of $J_{SC}$ falls to 0.37 mA/cm$^2$ irrespective of the $KGeCl_3$ thickness. Fig. 6(c) suggests that the FF

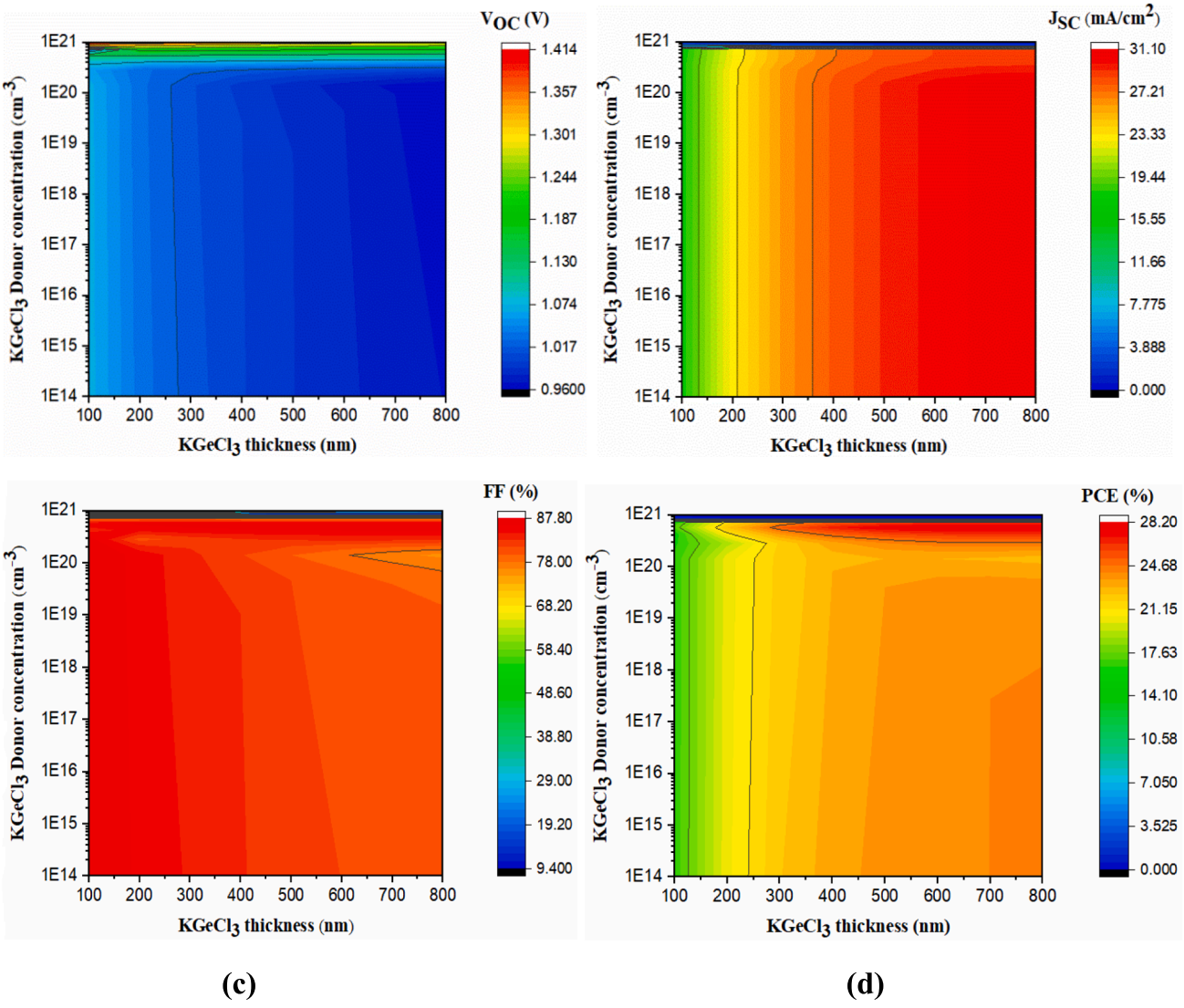



**Fig. 6.** Variation in PV parameters of simulated PSC such as (a) $V_{OC}$ (b) $J_{SC}$ (c) FF and (d) PCE with thickness of absorber layer and $N_D$ at $E_g$ = 1.40 eV.

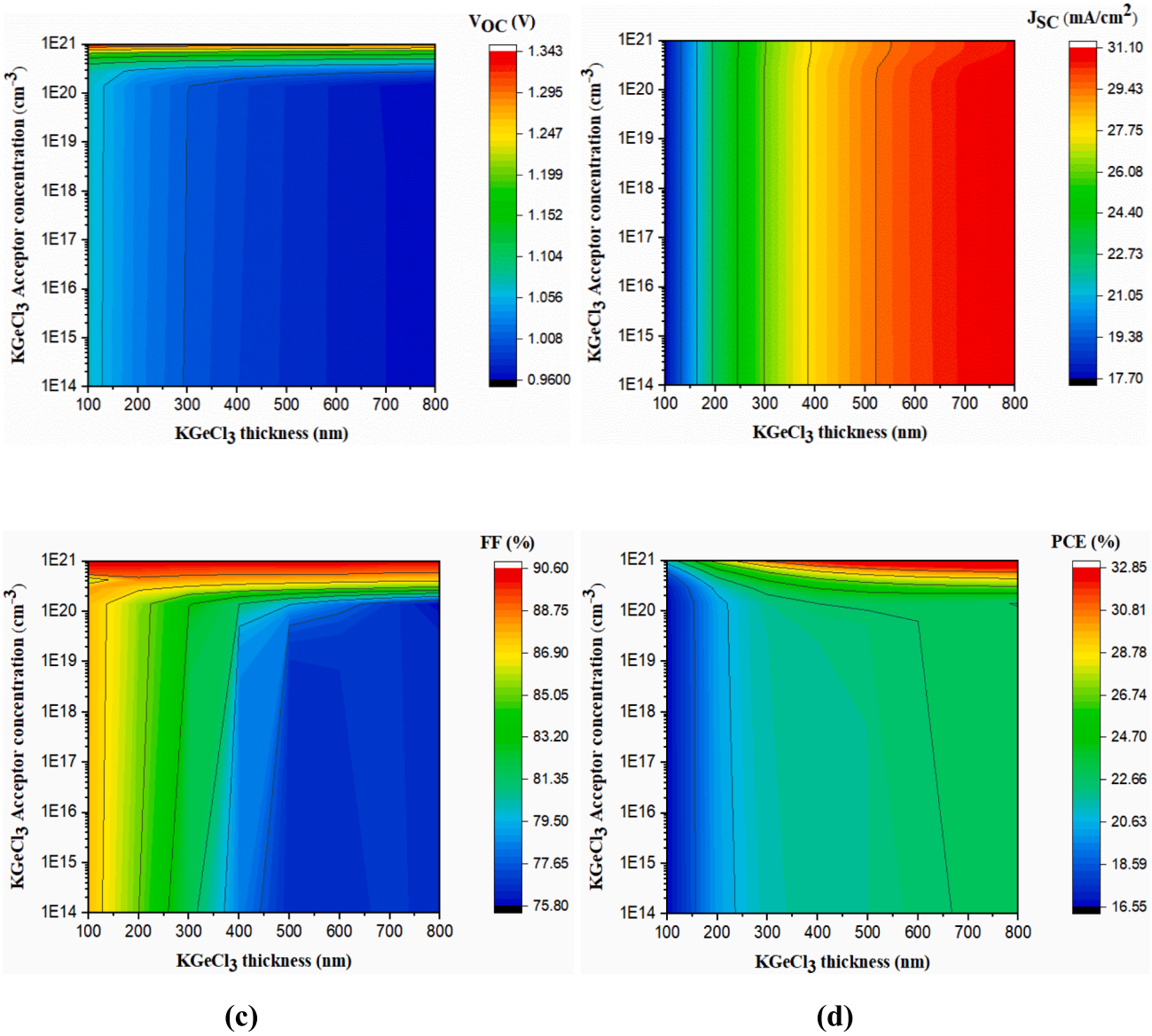



**Fig. 7.** Variation in PV parameters of simulated PSC such as (a) $V_{OC}$ (b) $J_{SC}$ (c) FF and (d) PCE with thickness of absorber layer and $N_A$.

maintains a relatively steady value of approximately 82.44 % across various $N_D$ and thickness, except for $N_D$ ranging from $1 \times 10^{19}$ to $1 \times 10^{20}$ $cm^{-3}$. Within a thickness range of 100 to 800 nm, first FF decreases approximately 13.46 % for the range of $N_D$ of $1 \times 10^{14}$ to $1 \times 10^{16}$ $cm^{-3}$. The FF shows a slight increase by around 3.27 % for the range of $N_D$ of $1 \times 10^{17}$ to $1 \times 10^{19}$ $cm^{-3}$ and further decrease for higher value of $N_D$. Fig. 6(d) depicts a correlation between $N_D$ and PCE with $KGeCl_3$ thickness. As $N_D$ increases from $1 \times 10^{14}$ to $1 \times 10^{19}$ $cm^{-3}$, PCE increases whereas for $N_D$ greater than the $1 \times 10^{19}$ $cm^{-3}$ which results in decrease in PCE. Hence, the optimized value of $N_D$ is $1 \times 10^{18}$ $cm^{-3}$ for $KGeCl_3$-based PSC. Specifically, it rises from approximately 16.07 to 28.18 % with initial increment in $N_D$ and thickness of the $KGeCl_3$ layer. The PCE decreased to 0.01 % with a high value of $N_D$. The high value of doping results in formation of coulomb traps, which decreases the mobility of charge carriers. So, we need to take the optimal value of $N_D$ for simulation [43]. This pattern suggests a direct relationship between $N_D$ and PCE, with a more pronounced effect at higher concentrations. It is evident that increasing $N_D$ of $KGeCl_3$ absorber layer has no discernible impact on the $V_{OC}$ value, particularly when the concentrations are below $1 \times 10^{20}$ $cm^{-3}$, across each individual thickness. This could be attributed to the complete formation of the depletion width, impacting the generation rate of charge carriers and the diffusion length of minority carrier's holes in ETL and electrons in HTL. Conversely, the $V_{OC}$ value increases with the thickness of $KGeCl_3$ for each value of $N_A$. This phenomenon can be ascribed to the prolonged lifetime and enhanced length of charge collection, resulting in greater photogenerated current. Consequently, leakage current and recombination rate decreases as the $KGeCl_3$ thickness grows [44]. The PCE values encapsulate the overall performance of $V_{OC}$, $J_{SC}$ and FF, with its value primarily influenced by $V_{OC}$ and FF shown below in Fig. 6.

### 3.1.3. *Effect of acceptor carrier concentration ($N_A$) with thickness*

Fig. 7 illustrates a contour plot showcasing the change in the photovoltaic parameters with respect to two key factors: thickness of $KGeCl_3$ ranging from 100 to 800 nm and $N_A$ ranging from $1 \times 10^{14}$ to $1 \times 10^{21}$ $cm^{-3}$. From Fig. 7(a), it is evident that the $V_{OC}$ shows improvement, increasing from 0.96 to 1.34 V as $N_A$ enhances from $1 \times 10^{14}$ to $1 \times 10^{21}$ $cm^{-3}$. Meanwhile, there is only a minor change in the $V_{OC}$ values with increase in $KGeCl_3$ thickness. In Fig. 7(b), it is analyzed that the value of $J_{SC}$ increases from approximately 17.71 to 30.60 $mA/cm^2$ as $N_A$ rises from $1 \times 10^{14}$ to $1 \times 10^{21}$ $cm^{-3}$ for increase in thickness of $KGeCl_3$. However, the value of $J_{SC}$ remains constant around 17.71 $mA/cm^2$ for

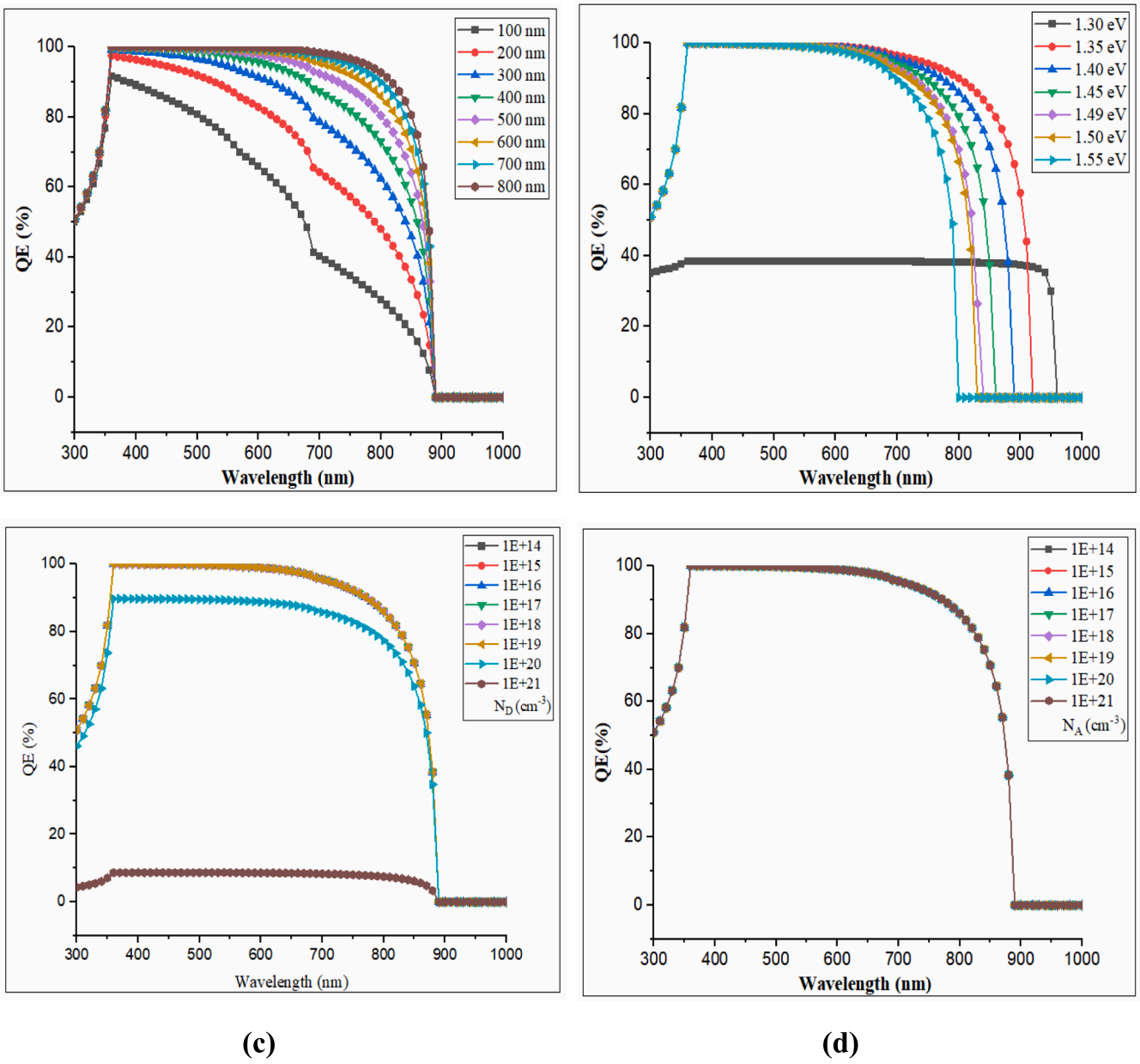



**Fig. 8.** The QE curves of the simulated PSC as a function of absorber layer (a) thickness, (b) $E_g$, (c) $N_D$ and (d) $N_A$.

$N_A$ ranging from $1 \times 10^{14}$ to $1 \times 10^{21}$ $cm^{-3}$ at constant thickness of $KGeCl_3$. From Fig. 7(c), it can be inferred that the FF values generally remain constant at around 88 % across various $N_A$ and thicknesses. However, there is an exception: for $N_A$ ranging from $1 \times 10^{14}$ to $1 \times 10^{21}$ $cm^{-3}$ and thickness of the absorber layer between 100 to 500 nm, the FF increases from approximately 76.48 to 90.58 %. Additionally, a higher FF value of about 91 % is achieved for $N_A$ of $1 \times 10^{21}$ $cm^{-3}$ and for various thickness ranging from 100 to 800 nm. In Fig. 7(d), it is evident that the PCE increases from 16.56 % to 28.63 % as $N_A$ rises from $1 \times 10^{14}$ to $1 \times 10^{17}$ $cm^{-3}$, accompanied by an increase in thickness. Subsequently, as $N_A$ further increases from $1 \times 10^{17}$ to $1 \times 10^{21}$ $cm^{-3}$, the value of PCE rises from 17.50 % to 31.10 %. These findings may be rationalized by the increase in carrier concentration leading to heightened electron diffusion from the $KGeCl_3$ layer to ETL. This elevated electron diffusion subsequently augments the device's built-in voltage ($V_{bi}$), predominantly formed by the depletion width, consequently resulting in a significant rise in $V_{OC}$ [42,45]. However, the $J_{SC}$ values improve with increasing $N_A$ due to enhancement of photogenerated charge carriers and the reduction of both leakage current and recombination current, particularly noticeable for $KGeCl_3$ thickness exceeding 300 nm. These simulation results conclude that the optimized value of $N_A$ is $1 \times 10^{19}$ $cm^{-3}$ for $KGeCl_3$-based PSC.

In this instance, the shape of PCE closely mirrors $J_{SC}$, as the $N_A$ predominantly affects the photogenerated current. Consequently, the QE is measured as a function of thickness of absorber layer, $E_g$ and carrier concentration, as illustrated in Fig. 8. It's apparent that QE exhibits a different shape for the different values of thickness, reaching 50 % at 300 nm and 100 % at 800 nm. As the $E_g$ increases from 1.30 to 1.55 eV, the region under the curve representing higher QE shifts downward from 1000 to 800 nm depicted in Fig. 8(b). The Fig. 8(c) showed that the $N_D$ is raised from $1 \times 10^{14}$ to $1 \times 10^{20}$ $cm^{-3}$, the QE value enhances from approximately 50 to 98 % across wavelengths from 300 to 640 nm. Conversely, the QE value declines from 98 to 34 % for longer wavelengths from 640 to 880 nm. The QE value falls from 99 to 4 % for all further wavelengths at $N_D$ of $1 \times 10^{21}$ $cm^{-3}$. Regardless of the $N_A$ value, the QE maintains consistent values and shapes demonstrated in Fig. 8

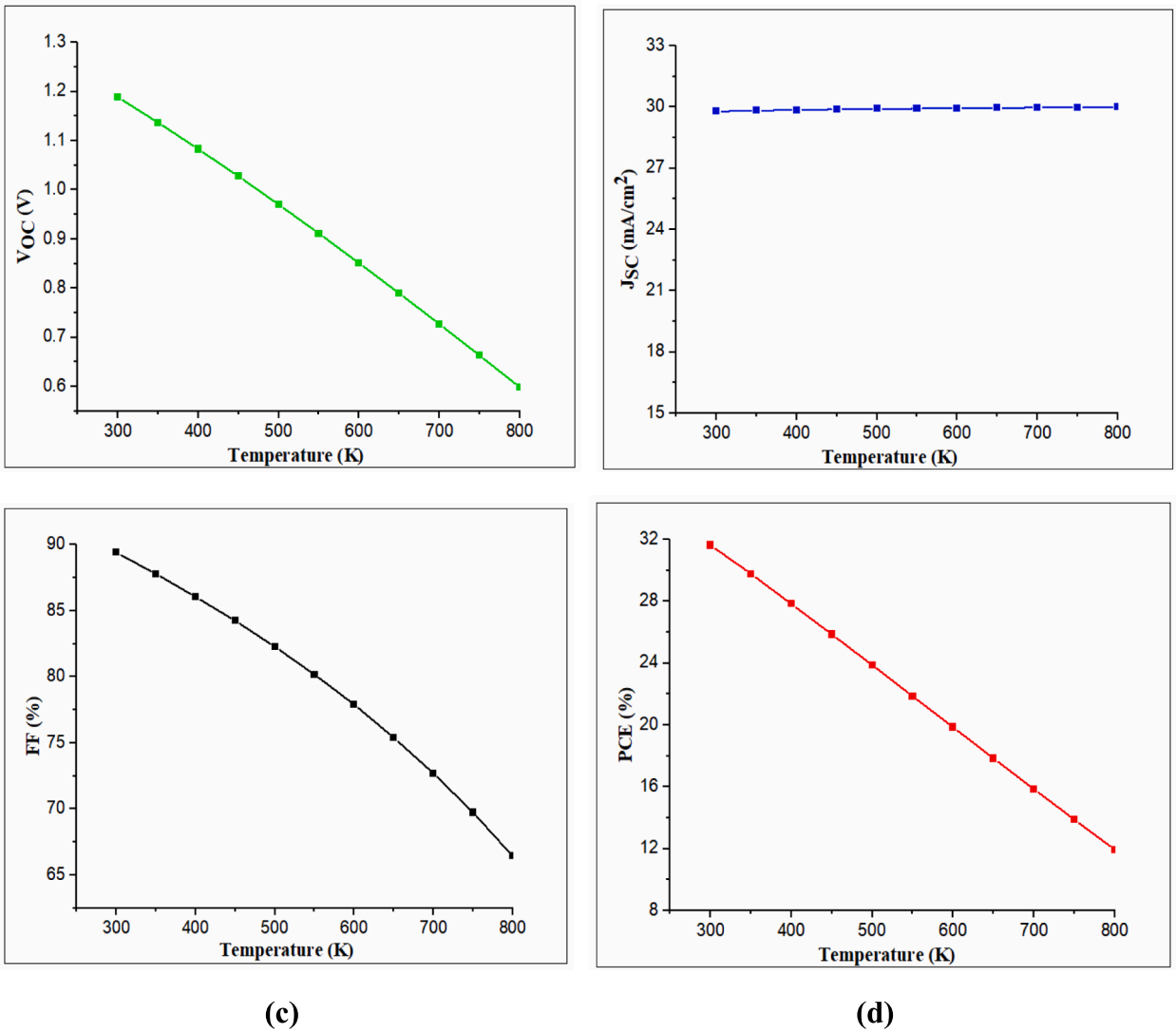



**Fig. 9.** Impact of temperature on PV parameters of $KGeCl_3$ PSC.

**Table 4**
Effect of change in temperature on PV parameters.

| Temperature | $V_{OC}$ (V) | $J_{SC}$ (mA/cm$^2$) | FF (%) | PCE (%) |
|---|---|---|---|---|
| 300 K | 1.188 | 29.78 | 89.39 | 31.64 |
| 350 K | 1.136 | 29.83 | 87.78 | 29.78 |
| 400 K | 1.083 | 29.86 | 86.05 | 27.84 |
| 450 K | 1.027 | 29.89 | 84.23 | 25.87 |
| 500 K | 0.970 | 29.91 | 82.27 | 23.88 |
| 550 K | 0.911 | 29.93 | 80.17 | 21.87 |
| 600 K | 0.851 | 29.95 | 77.91 | 19.86 |
| 650 K | 0.789 | 29.96 | 75.39 | 17.84 |
| 700 K | 0.727 | 29.98 | 72.70 | 15.85 |
| 750 K | 0.663 | 29.99 | 69.74 | 13.88 |
| 800 K | 0.598 | 30.00 | 66.44 | 11.93 |

(d). The findings suggest that the photogenerated current increases with high-energy photons but decreases for lower-energy ones, indicating differing rates of charge generation and recombination that influence $J_{SC}$. These results confirm and support the notion that PCE is primarily affected by $J_{SC}$ [46]. Based on the simulations, it can be inferred that the $KGeCl_3$ $E_g$, thickness, $N_D$ and $N_A$ should fall within the ranges of 1.40 eV, 600 nm, $1 \times 10^{18}$ cm$^{-3}$ and $1 \times 10^{19}$ cm$^{-3}$, respectively to attain higher PCE.

### 3.2. Effect of temperature on PSC

To examine the impact of operating temperature on the PV parameters, $WS_2$ is used as ETL and PEDOT: PSS as HTL as illustrated in Fig. 9. The operating temperature at which the PSC operates has been varied between 300 K and 800 K depicted in Table 4. Fall in the performance with temperature increase is related to a decrease in diffusion length of charge carriers [47]. Several solar cell architectures show the instability of their overall performance due to deformation among layers at high temperatures. In $KGeCl_3$-based PSC, fall in efficiency with temperature is about 63 %, for thickness of ETL and HTL is 100 nm and 80 nm respectively. When the enhancement of temperature occurs, the PCE drops significantly from 31.64 % at 300 K to 11.93 % at 800 K. The sharp fall in PCE shows that the optimized device is highly unstable towards higher values of temperature. The minor rise in $J_{SC}$ with increasing temperature results in reduced $E_g$ and creation of additional $e^-$-$h^+$ pairs [48]. The value of $V_{OC}$ is decreasing with an increase in the operating temperature with increasing interfacial defects and series resistance. The decrease in diffusion length with increase in temperature, leading to a rise in series resistance, which in turn impacts the FF and overall PCE of the device [49,50]. The increment in temperature results in enhancement in reverse saturation current and recombination rate of charge carriers, therefore the $V_{OC}$, FF and PCE decreased [51,52]. The simulation outcomes showed that the optimized value of temperature is

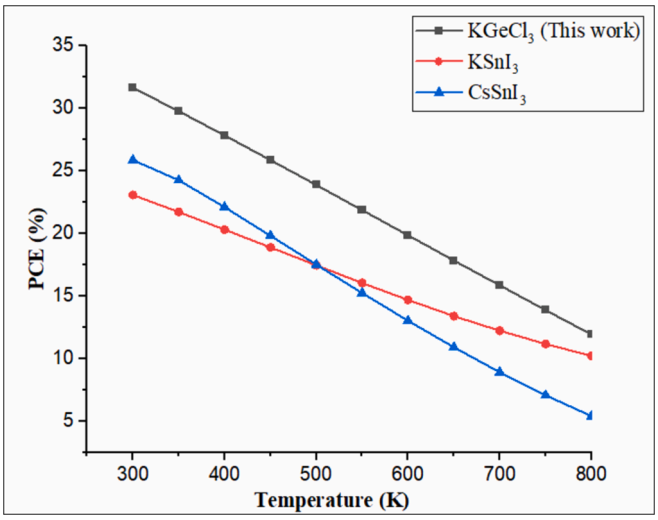



**Fig. 10.** PCE vs temperature curves of various PSC materials [53,54].

300 K. Comparison of observed temperature dependance with PCE of different PSC materials is shown below in Fig. 10.

### 3.3. *Verification with respect to the ideal cell*

The comparison of the obtained PV parameters of $KGeCl_3$-based PSC as a function of absorber layer $E_g$ with the ideal single junction solar cell parameters has been done. The maximum PCE of 29.82 % with $V_{OC}$ of 1.19 V, $J_{SC}$ of 30.08 mA/cm$^2$, and FF of 82.74 % for the optimized configuration of FTO/$WS_2$/$KGeCl_3$/PEDOT: PSS/Au has been obtained at $E_g$ of 1.49 eV and radiative recombination coefficient of $10^{-12}$ cm$^3$/sec. The ideal single junction solar cell parameters are as follows: PCE of 31.00 %, $V_{OC}$ of 1.20 V, $J_{SC}$ of 29.80 mA/cm$^2$, and FF of 89.80 % at the absorber layer $E_g$ of 1.50 eV [55]. This demonstrated that the work outcomes are within the limit of the ideal single junction solar cell parameters.

## 4. Conclusions

This work comprises simulation study of $KGeCl_3$-based PSC. The effect of distinct parameters, including thickness of absorber layer, ETLs and HTLs, $E_g$ $N_D$, $N_A$ and operating temperature have been examined using SCAPS-1D. The study demonstrated PEDOT: PSS as a promising HTL that could replace the expensive and less-conductive Spiro-OMeTAD. The maximum PCE of 29.82 % was obtained for the configuration of FTO/$WS_2$/$KGeCl_3$/PEDOT: PSS/Au. The SCAPS-1D outcomes, deduced that the optimized thickness of the absorber layer was 600 nm, with $N_D$ of $1 \times 10^{18}$ cm$^{-3}$ and a $N_A$ of $1 \times 10^{19}$ cm$^{-3}$ results in maximum PCE at $E_g$ of 1.40 eV. The optimized PSC is simulated at $N_t$ of $1 \times 10^{15}$ cm$^{-3}$. The simulation trend demonstrated that photovoltaic performance of a PSC falls with increment in operating temperature hence 300 K is optimized temperature. The device PCE improved beyond 29 % by optimizing thickness, operating temperature, $E_g$, $N_D$ and $N_A$. The proper optimization of suitable ETL and HTL has been done for $KGeCl_3$-based PSC hence, this can be an alternative to Pb-based PSC for the preparation of efficient, low-cost and eco-friendly PSCs.

## CRediT authorship contribution statement

**Rahul Kundara:** Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Sarita Baghel:** Writing – review & editing, Validation, Supervision, Resources, Formal analysis.

## 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.

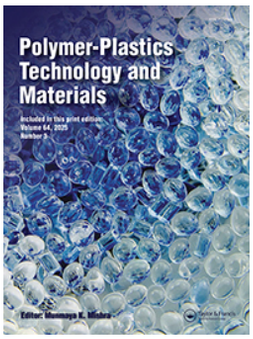
## Acknowledgment

The author, Rahul Kundara is thankful to Delhi Technological University, New Delhi for providing a senior research fellowship. We are very grateful to Marc Burgelman, the Electronics and Information Systems (ELIS), University of Gent, Belgium, for providing us with SCAPS-1D software.

## References

[1] M.S. Islam, K. Sobayel, A. Al-Kahtani, M.A. Islam, G. Muhammad, N. Amin, M. Shahiduzzaman, M. Akhtaruzzaman, Defect study and modelling of SnX3-based perovskite solar cells with SCAPS-1D, Nanomater. 11 (2021), https://doi.org/10.3390/nano11051218.
[2] A. Singha, A. Paul, S. Koul, V. Sharma, S. Mallick, K.R. Balasubramaniam, D. Kabra, Stable and Efficient Large Area 4T Si/perovskite Tandem Photovoltaics with Sputtered Transparent Contact, Sol. RRL. 7 (2023) 2300117, https://doi.org/10.1002/solr.202300117.
[3] J. Jeong, M. Kim, J. Seo, H. Lu, P. Ahlawat, A. Mishra, Y. Yang, M.A. Hope, F. T. Eickemeyer, M. Kim, Y.J. Yoon, Pseudo-halide anion engineering for α-FAPbI3 perovskite solar cells, Nature. 592 (2021) 381–385, https://doi.org/10.1038/s41586-021-03406-5.
[4] Y. Tang, Z. Gu, C. Fu, Q. Xiao, S. Zhang, Y. Zhang, Y. Song, FAPbI3 Perovskite Solar cells: From Film morphology regulation to device optimization, Sol. RRL. 6 (2022) 2200120, https://doi.org/10.1002/solr.202200120.
[5] C. Liang, Z. Huang, J. Su, L. Shi, S. Liang, Y. Dong, Study on Performance Optimization of Perovskite Solar Cells Based on MAPbI3, Adv. Theor Simul. 7 (2024) 2301015, https://doi.org/10.1002/adts.202301015.
[6] Y. Liu, Y. Liu, Y. Guo, Organic–inorganic hybrid perovskite materials and their application in transistors, Mater. Chem. Front. 7 (2023) 5215–5246, https://doi.org/10.1039/D3QM00697B.
[7] A. Kojima, K. Teshima, Y. Shirai, T. Miyasaka, Organometal halide perovskites as visible-light sensitizers for photovoltaic cells, J. Am. Chem. Soc. 131 (2009) 6050–6051, https://doi.org/10.1021/ja809598r.
[8] N.G. Park, Perovskite solar cells: an emerging photovoltaic technology, Mater. Today. 18 (2015) 65–72, https://doi.org/10.1016/j.mattod.2014.07.007.
[9] Y. Zhao, K. Zhu, Organic–inorganic hybrid lead halide perovskites for optoelectronic and electronic applications, Chem. Soc. Rev. 45 (2016) 655–689, https://doi.org/10.1039/C4CS00458B.
[10] M. Kulbak, D. Cahen, G. Hodes, How important is the organic part of lead halide perovskite photovoltaic cells? Efficient CsPbBr3 cells, J. Phys. Chem. Lett. 6 (2015) 2452–2456, https://doi.org/10.1021/acs.jpclett.5b00968.
[11] G.E. Eperon, S.D. Stranks, C. Menelaou, M.B. Johnston, L.M. Herz, H.J. Snaith, Formamidinium lead trihalide: a broadly tunable perovskite for efficient planar heterojunction solar cells, Energy Environ. Sci. 7 (2014) 982–988, https://doi.org/10.1039/C3EE43822H.
[12] N.A. Ouedraogo, Y. Chen, Y.Y. Xiao, Q. Meng, C.B. Han, H. Yan, Y. Zhang, Stability of all-inorganic perovskite solar cells, Nano Energy. 67 (2020) 104249, https://doi.org/10.1016/j.nanoen.2019.104249.
[13] N. Sun, S. Fu, Y. Li, L. Chen, J. Chung, M.M. Saeed, K. Dolia, A. Rahimi, C. Li, Z. Song, Y. Yan, Tailoring Crystallization Dynamics of CsPbI3 for Scalable Production of Efficient Inorganic Perovskite Solar Cells, Adv. Funct. Mater. 34 (2024) 2309894, https://doi.org/10.1002/adfm.202309894.
[14] J. Zhuang, J. Wang, F. Yan F, Review on chemical stability of lead halide perovskite solar cells, Nano-Micro Letters. 15 (2023) 84. doi: 10.1007/s40820-023-01046-0.
[15] S. Gressler, F. Part, S. Scherhaufer, G. Obersteiner, M. Huber-Humer, Advanced materials for emerging photovoltaic systems–Environmental hotspots in the production and end-of-life phase of organic, dye-sensitized, perovskite, and quantum dots solar cells, Sustain. Mater. Techno. 34 (2022) e00501. doi: 10.1016/j.susmat.2022.e00501.
[16] G. Schileo, G. Grancini, Lead or no lead? Availability, toxicity, sustainability and environmental impact of lead-free perovskite solar cells, J. Mater. Chem. c. 9 (2021) 67–76, https://doi.org/10.1039/D0TC04552G.
[17] N.K. Noel, S.D. Stranks, A. Abate, C. Wehrenfennig, S. Guarnera, A.A. Haghighirad, A. Sadhanala, G.E. Eperon, S.K. Pathak, M.B. Johnston, A. Petrozza, L.M. Herz, H. J. Snaith, Lead-free organic-inorganic tin halide perovskites for photovoltaic applications, Energy Environ. Sci. 7 (2014) 3061–3068, https://doi.org/10.1039/c4ee01076k.
[18] M.G. Ju, M. Chen, Y. Zhou, H.F. Garces, J. Dai, L. Ma, N.P. Padture, X.C. Zeng, Earth-Abundant Nontoxic Titanium(IV)-based Vacancy-Ordered Double Perovskite Halides with Tunable 1.0 to 1.8 eV Bandgaps for Photovoltaic Applications, ACS Energy Lett. 3 (2018) 297–304, https://doi.org/10.1021/acsenergylett.7b01167.

[19] Z. Zhang, X. Li, X. Xia, Z. Wang, Z. Huang, B. Lei, Y. Gao, High-Quality (CH3NH3) 3Bi2I9 Film-Based Solar Cells: Pushing Efficiency up to 1.64%, J. Phys. Chem. Lett. 8 (2017) 4300–4307, https://doi.org/10.1021/acs.jpclett.7b01952.
[20] M. Wang, P. Zeng, S. Bai, J. Gu, F. Li, Z. Yang, M. Liu, High-Quality Sequential-Vapor-Deposited Cs2AgBiBr6 Thin Films for Lead-Free Perovskite Solar Cells, Sol. RRL. 2 (2018) 1800217, https://doi.org/10.1002/solr.201800217.
[21] Y. Zong, Y. Zhou, Y. Zhang, Z. Li, L. Zhang, M.G. Ju, M. Chen, S. Pang, X.C. Zeng, N. P. Padture, Continuous Grain-Boundary Functionalization for High-Efficiency Perovskite Solar Cells with Exceptional Stability, Chem. 4 (2018) 1404–1415, https://doi.org/10.1016/j.chempr.2018.03.005.
[22] M.A. Siddique, A.S. Rahman, The SCAPS-1D simulation of non-toxic KGeCl3 perovskite from DFT derived properties, Mater. Sci. Eng. b. 303 (2024) 117268, https://doi.org/10.1016/j.mseb.2024.117268.
[23] N. Shrivastav, M.A. Hamid, J. Madan, R. Pandey, Exploring KGeCl3 material for perovskite solar cell absorber layer through different machine learning models, Sol. Energy. 278 (2024) 112784, https://doi.org/10.1016/j.solener.2024.112784.
[24] K. Sekar, L. Marasamy, S. Mayarambakam, H. Hawashin, M. Nour, J. Bouclé, Lead-free, formamidinium germanium-antimony halide (FA 4 GeSbCl 12) double perovskite solar cells: the effects of band offsets, RSC Adv. 13 (2023) 25483–25496, https://doi.org/10.1039/D3RA03102K.
[25] V. Kumar, R. Kumar, F. Chand, Performance analysis of ecofriendly Ge based perovskite solar cell using computational approach, Materials Letters. 361 (2024) 136145, https://doi.org/10.1016/j.matlet.2024.136145.
[26] M.S. Azizman, A.W. Azhari, D.S. Halin, N. Ibrahim, S. Sepeai, N.A. Ludin, M. N. Nor, L.N. Ho, Progress in tin-germanium perovskite solar cells: A review, Synthetic Metals. 299 (2023) 117475, https://doi.org/10.1016/j.synthmet.2023.117475.
[27] A. Kuddus, M.F. Rahman, S. Ahmmed, J. Hossain, A.B.M. Ismail, Role of facile synthesized V2O5 as hole transport layer for CdS/CdTe heterojunction solar cell: Validation of simulation using experimental data, Superlattices Microstruct. 132 (2019) 106168, https://doi.org/10.1016/j.spmi.2019.106168.
[28] U. Mandadapu, S.V. Vedanayakam, K. Thyagarajan, Simulation and analysis of lead based perovskite solar cell using SCAPS-1D, Indian, Journal of Science and Technology. 10 (2017) 65–72, https://doi.org/10.17485/ijst/2017/v11i10/110721.
[29] F. Azri, A. Meftah, N. Sengouga, A. Meftah, Electron and hole transport layers optimization by numerical simulation of a perovskite solar cell, Sol. Energy. 181 (2019) 372–378, https://doi.org/10.1016/j.solener.2019.02.017.
[30] K.D. Jayan, V. Sebastian, Comprehensive device modelling and performance analysis of MASnI3 based perovskite solar cells with diverse ETM, HTM and back metal contacts, Sol. Energy. 217 (2021) 40–48, https://doi.org/10.1016/j.solener.2021.01.058.
[31] M.K. Hossain, G.F.I. Toki, A. Kuddus, M.H.K. Rubel, M.M. Hossain, H. Bencherif, M.F. Rahman, M.R. Islam, M. Mushtaq, An extensive study on multiple ETL and HTL layers to design and simulation of high-performance lead-free CsSnCl3-based perovskite solar cells, Sci Rep. 13 (2023) 2521, https://doi.org/10.1038/s41598-023-28506-2.
[32] H. Wei, P. Qiu, Y. Li, Y. He, M. Peng, X. Zheng, X. Liu, Challenges and strategies of all-inorganic lead-free halide perovskite solar cells, Ceram. Int. 48 (2022) 5876–5891, https://doi.org/10.1016/j.ceramint.2021.11.184.
[33] S. Rai, B.K. Pandey, D.K. Dwivedi, Modeling of highly efficient and low cost CH3NH3Pb (I1-xClx) 3 based perovskite solar cell by numerical simulation, Opt. Mater. 100 (2020) 109631, https://doi.org/10.1016/j.optmat.2019.109631.
[34] S.Z. Haider, H. Anwar, M. Wang, Theoretical device engineering for high-performance perovskite solar cells using CuSCN as hole transport material boost the efficiency above 25%, Phys. Status Solidi (a). 216 (2019) 1900102 https://doi.org/10.1002/pssa.201900102.
[35] A.A. Kanoun, M.B. Kanoun, A.E. Merad, S. Goumri-Said, Toward development of high-performance perovskite solar cells based on CH3NH3GeI3 using computational approach, Sol. Energy. 182 (2019) 237–244, https://doi.org/10.1016/j.solener.2019.02.041.
[36] S. Bhattarai, T.D. Das, Optimization of carrier transport materials for the performance enhancement of the MAGeI3 based perovskite solar cell, Sol. Energy. 217 (2021) 200–207, https://doi.org/10.1016/j.solener.2021.02.002.
[37] I.M. De Los Santos, H.J. Cortina-Marrero, M.A. Ruíz-Sánchez, L. Hechavarría-Difur, F.J. Sánchez-Rodríguez, M. Courel, H. Hu, Optimization of CH3NH3PbI3 perovskite solar cells: A theoretical and experimental study, Sol. Energy. 199 (2020) 198-205. doi: 10.1016/j.solener.2020.02.026.
[38] S. Kohnehpoushi, P. Nazari, B.A. Nejand, M. Eskandari, MoS2: a two-dimensional hole-transporting material for high-efficiency, low-cost perovskite solar cells, Nanotechnol. 29 (2018) 205201, https://doi.org/10.1088/1361-6528/aab1d4.
[39] D. Ompong, J. Singh, High open-circuit voltage in perovskite solar cells: The role of hole transport layer, Org. Electron. 63 (2018) 104–108, https://doi.org/10.1016/j.orgel.2018.09.006.
[40] T. Ouslimane, L. Et-Taya, L. Elmaimouni, A. Benami, Impact of absorber layer thickness, defect density, and operating temperature on the performance of MAPbI3 solar cells based on ZnO electron transporting material, Heliyon. 7(3) (2021) e06379. doi: 10.1016/j.heliyon.2021.e06379.
[41] M. Asaduzzaman, M. Hasan, A.N. Bahar, An investigation into the effects of band gap and doping concentration on Cu (In, Ga) Se 2 solar cell efficiency, SpringerPlus. 5 (2016) 1–8, https://doi.org/10.1186/s40064-016-2256-8.
[42] M. Abdelfatah, J. Ledig, A. El-Shaer, A. Wagner, V. Marin-Borras, A. Sharafeev, P. Lemmens, M.M. Mosaad, A. Waag, A. Bakin, Fabrication and characterization of low cost Cu2O/ZnO: Al solar cells for sustainable photovoltaics with earth abundant materials, Sol. Energy Mater. Sol. Cells. 145 (2016) 454–461, https://doi.org/10.1016/j.solmat.2015.11.015.
[43] Z. Shang, T. Heumueller, R. Prasanna, G.F. Burkhard, B.D. Naab, Z. Bao, M. D. McGehee, A. Salleo, Trade-off between trap filling, trap creation, and charge recombination results in performance increase at ultralow doping levels in bulk heterojunction solar cells, Adv. Energy Mater. 6 (2016) 1601149, https://doi.org/10.1002/aenm.201601149.
[44] K.P. Musselman, A. Marin, L. Schmidt-Mende, J.L. MacManus-Driscoll, Incompatible length scales in nanostructured Cu2O solar cells, Adv. Funct. Mater. 22 (2012) 2202–2208, https://doi.org/10.1002/adfm.201102263.
[45] W.M. Lin, N. Yazdani, O. Yarema, S. Volk, M. Yarema, T. Kirchartz, V. Wood, Simulating nanocrystal-based solar cells: A lead sulfide case study, J. Chem. Phys. 151 (2019) 241104, https://doi.org/10.1063/1.5129159.
[46] M. Abdelfatah, A.M. El Sayed, W. Ismail, S. Ulrich, V. Sittinger, A. El-Shaer, SCAPS simulation of novel inorganic ZrS2/CuO heterojunction solar cells, Sci. Rep. 13 (2023) 4553, https://doi.org/10.1038/s41598-023-31553-4.
[47] G. Pindolia, S.M. Shinde, P.K. Jha, Optimization of an inorganic lead free RbGeI3 based perovskite solar cell by SCAPS-1D simulation, Sol. Energy. 236 (2022) 802–821, https://doi.org/10.1016/j.solener.2022.03.053.
[48] V. Deswal, S. Kaushik, R. Kundara, S. Baghel, Numerical simulation of highly efficient Cs2AgInBr6-based double perovskite solar cell using SCAPS 1-D, Mater. Sci. Eng. b. 299 (2024) 117041, https://doi.org/10.1016/j.mseb.2023.117041.
[49] M.K. Hossain, M.H. Rubel, G.I. Toki, I. Alam, M.F. Rahman, H. Bencherif, Effect of various electron and hole transport layers on the performance of CsPbI3-based perovskite solar cells: A numerical investigation in DFT, SCAPS-1D, and wxAMPS frameworks, ACS Omega. 7 (2022) 43210–43230, https://doi.org/10.1021/acsomega.2c05912.
[50] R. Kundara, S. Baghel, Device modelling of lead free (CH3NH3) 2CuX4 based perovskite solar cells using SCAPS simulation, Opt. Quantum Electron. 55 (2023) 968, https://doi.org/10.1007/s11082-023-05244-3.
[51] S. Ahmed, F. Jannat, M.A.K. Khan, M.A. Alim, Numerical development of eco-friendly Cs2TiBr6 based perovskite solar cell with all-inorganic charge transport materials via SCAPS-1D, Optik. 225 (2021) 165765, https://doi.org/10.1016/j.ijleo.2020.165765.
[52] A. Chen, K. Zhu, Computer simulation of a-Si/c-Si heterojunction solar cell with high conversion efficiency, Sol. Energy. 86 (2012) 393–397, https://doi.org/10.1016/j.solener.2011.10.015.
[53] G. Pindolia, S.M. Shinde, P.K. Jha, Non-leaded, KSnI3 based perovskite solar cell: A DFT study along with SCAPS simulation, Mater Chem Phys. 297 (2023) 127426, https://doi.org/10.1016/j.matchemphys.2023.127426.
[54] M. Jiang, J. Tang, Simulated development and optimized performance of narrow-bandgap CsSnI3-based all-inorganic perovskite solar cells, J. Phys. d: Appl. Phys. 54 (2021) 465104, https://doi.org/10.1088/1361-6463/ac1e4c.
[55] A. Morales-Acevedo, Fundamentals of solar cell physics revisited: Common pitfalls when reporting calculated and measured photocurrent density, open-circuit voltage, and efficiency of solar cells, Sol. Energy. 262 (2023) 111774, https://doi.org/10.1016/j.solener.2023.05.051.




Taylor & Francis
Taylor & Francis Group

Polymer-Plastics Technology and Materials

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/lpte21

# Polyvinyl Alcohol: Recent Advances and Applications in Sustainable Materials

**Teena Saini, Jagram Meena, Vratika Verma, Sakshi Saini & Rajdeep Malik**

**To cite this article:** Teena Saini, Jagram Meena, Vratika Verma, Sakshi Saini & Rajdeep Malik (30 Dec 2024): Polyvinyl Alcohol: Recent Advances and Applications in Sustainable Materials, Polymer-Plastics Technology and Materials, DOI: 10.1080/25740881.2024.2438046

**To link to this article:** https://doi.org/10.1080/25740881.2024.2438046

Published online: 30 Dec 2024.

Submit your article to this journal

Article views: 40

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/10.1080/25740881.2024.2438046

Taylor & Francis
Taylor & Francis Group

Check for updates

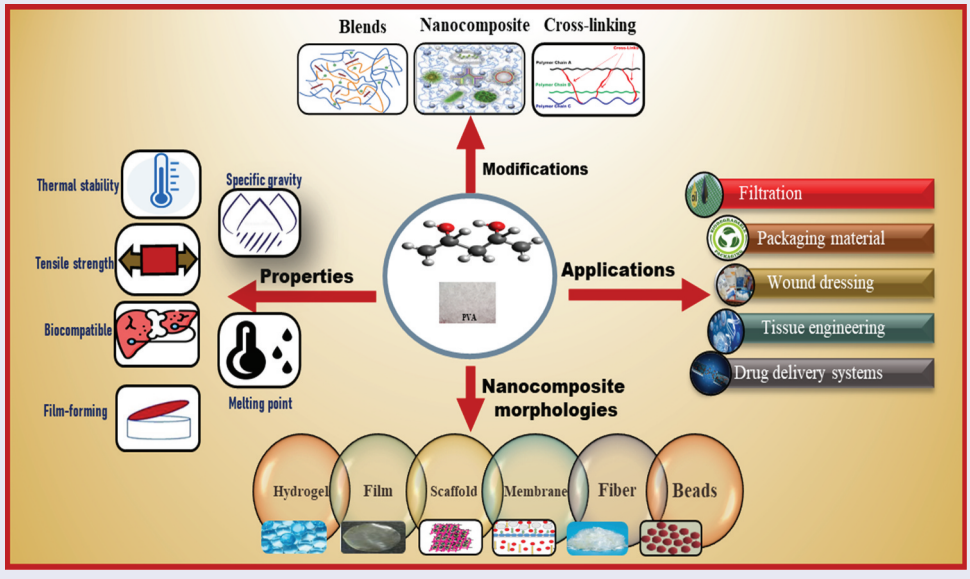
# Polyvinyl Alcohol: Recent Advances and Applications in Sustainable Materials

Teena Saini, Jagram Meena, Vratika Verma, Sakshi Saini, and Rajdeep Malik

Department of Chemistry, Gurukul Kangri Deemed to be University, Haridwar, India

**ABSTRACT**

Polyvinyl alcohol (PVA) is a synthetic biopolymer widely recognized for its biodegradability, biocompatibility, and excellent film-forming properties. Recent advancements in PVA-based materials have expanded its use in diverse fields, including packaging, biomedicine, and wastewater treatment. This review offers a comprehensive analysis of the latest innovations in PVA applications, emphasizing their role as a sustainable material, particularly through the development of PVA-based nanocomposites. The addition of nanoparticles, cross-linking agents, and biopolymer combinations has significantly enhanced the mechanical strength, antibacterial properties, and UV-blocking capacity of PVA, making it well-suited for advanced packaging and biomedical applications. Cross-linking improves water resistance and mechanical durability while combining with biopolymers such as starch and chitosan enhances biodegradability. This review highlights the structural, mechanical, optical, and electrical properties of PVA-based nanocomposites.




**SUMMARY**

- Review covers PVA's biodegradability, biocompatibility, and eco-friendly uses.
- PVA properties enhanced by nanoparticles, crosslinking, and biopolymer blending.
- The review discusses PVA applications in packaging, biomedical, and more.
  (Please include this as a summary point in the list with a bullet for consistency) The review emphasizes recent PVA advancements and potential in sustainability.

**KEYWORDS**

Polyvinyl alcohol (PVA); PVA modifications; nanocomposites; biopolymer blends; tissue engineering

## 1. Introduction

Polymers, large molecules with repeating structural units, are versatile compounds that have revolutionized various industries. Their diverse properties make them essential in industrial, biological, and synthetic applications. These macromolecules, forming the basis of countless materials, have transformed technology and daily life, showcasing their significant impact on the world.[1, 2]. The influence of these polymers on our current lifestyle is nearly immeasurable. Products crafted from polymers surround us everywhere: synthetic fiber clothing, plastic bags, nylon bearings,

**CONTACT** Jagram Meena jrm.sv3@gmail.com Department of Chemistry, Gurukul Kangri Deemed to be University, Haridwar 249404, India

© 2024 Taylor & Francis Group, LLC

fiberglass, polyethylene cups, paint with polymer bases, adhesive epoxy, cushioning made from polyurethane foam, and cookware coated with Teflon, among many others. The range of items is extensive [3]. Once their purpose is fulfilled, they are often discarded and can pose a threat to living organisms and the environment by releasing hazardous substances. This is particularly the case for synthetic polymers, as they tend to be non-biodegradable [4]. The growing concern over non-biodegradable plastic waste has spurred increasing research into biodegradable polymers to reduce environmental pollution. This has resulted in the commercialization of biopolymers and biopolymer composites in industries such as water purification, therapeutics, food, petroleum, and cosmetics.[5]

PVA is a synthetic biopolymer with a creamy or whitish appearance, devoid of taste and odor. It is nontoxic, biocompatible, and thermally stable, existing in a granular or powdered form with a semicrystalline or linear structure.[6] It was first developed in 1924 by Hermann and Haehnel through the hydrolysis of polyvinyl acetate in ethanol with potassium hydroxide. In contemporary commercial production, a continuous process is commonly employed. This process involves the hydrolysis of polyvinyl acetate through ester interchange with methanol. The reaction takes place in the presence of anhydrous sodium methylate or aqueous sodium hydroxide, resulting in the elimination of acetate groups and the formation of polyvinyl alcohol.[7] Due to the incomplete nature of this reaction, it leads to the formation of a copolymer comprising both polyvinyl alcohol (PVA) and polyvinyl acetate. The extent of hydrolysis is contingent upon the number of residual acetate groups present.[8] Polyvinyl alcohol has captured attention throughout its history due to its simple structure and its propensity for chemical reactions through its secondary alcohol functionality. The water solubility of PVA makes it particularly advantageous in environmentally conscious and green chemistry applications. Additionally, PVA serves as an embedding matrix for metallic nanoparticles such as silver nitrate, used in applications ranging from hydrogen sensors to antibacterial activities.[9–11] The favorable physical and chemical attributes of PVA make it a popular choice in polymer blends. It is widely utilized for its excellent film-forming properties and effective emulsifying capabilities.[12]

The objective of this review is to provide a focused examination of PVA applications, properties, modifications, and recent advances, particularly in packaging, biomedical applications, and environmental sustainability. It explores PVA biodegradability, biocompatibility, and key modifications such as nanoparticle incorporation, cross-linking, and biopolymer blends. The review highlights how these modifications enhance PVA mechanical strength, antibacterial activity, and UV-blocking properties, contributing to its growing role as a sustainable material across industries.

## 2. PVA structure

It exhibits a straightforward chemical arrangement, consisting of a central carbon atom backbone. This foundational structure is complemented by the incorporation of hydroxyl groups that extend like pendant branches from the main backbone, contributing to the overall molecular arrangement of PVA. Depending on the degree of hydrolysis of polyvinyl acetate, it consists of either 1,3-diol units or 1,2-diol units. Figure 1 presents the structure of PVA with different degrees of hydrolysis, including both fully hydrolyzed and partially hydrolyzed forms.[13]

## 3. PVA properties

PVA is commercially available in three common grades: Partially hydrolyzed, intermediate hydrolyzed, and fully hydrolyzed (Table 1). The degree of hydrolysis, also referred to as the degree of saponification significantly influences various mechanical properties of PVA, including its water solubility. In particular, fully hydrolyzed PVA exhibits lower water solubility when compared to partially hydrolyzed PVA.[14] Despite the hydrolysis process, PVA still contains 1–2 mol% of acetyl groups, impacting its properties. The mol% of acetate groups in PVA varies across commercial grades, with fully hydrolyzed PVA exhibiting a lower mol% of acetate groups compared to partially hydrolyzed counterparts.

The degree of hydroxylation in this synthesis process significantly influences the physical, chemical, and mechanical characteristics of the resulting PVA (Figure 2).[15]

It is categorized under nonionic water-soluble polymers containing a vinyl group. It is biocompatible, biodegradable, nontoxic, and nonhazardous, making it safe to handle and environmentally friendly. PVA is commonly available in the form of white to ivory granules or powder and is odorless, tasteless, and translucent. Its strong chemical resistance and high mechanical strength make it a versatile material, though it does have limitations, including suboptimal barrier and thermal properties and a relatively high cost.[13, 16] Despite these drawbacks, the properties of PVA, as outlined in Table 2, make it a widely used material in various applications due to its eco-friendly nature and ease of modification.

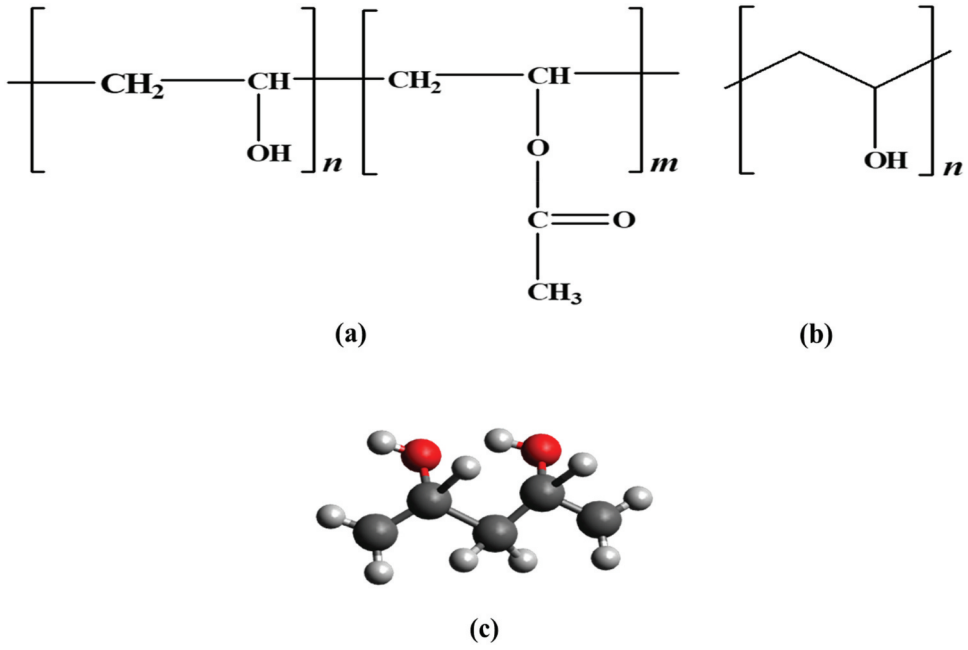



**Figure 1.** Structural representation of PVA in various forms: (a) Partially hydrolyzed PVA, (b) Fully hydrolyzed PVA, and (c) a 3D molecular model of PVA.

**Table 1.** PVA profile: Grade, hydrolysis degree, acetate group mol %.

| PVA grade | Degree of hydrolysis | Mol% of acetate groups |
|---|---|---|
| Partially hydrolyzed | 84.2–89.0% | 10–14 |
| Intermediate hydrolyzed | 92.5–96.5% | 3–7 |
| Fully hydrolyzed | 98.0–99.0% | 1–2 |

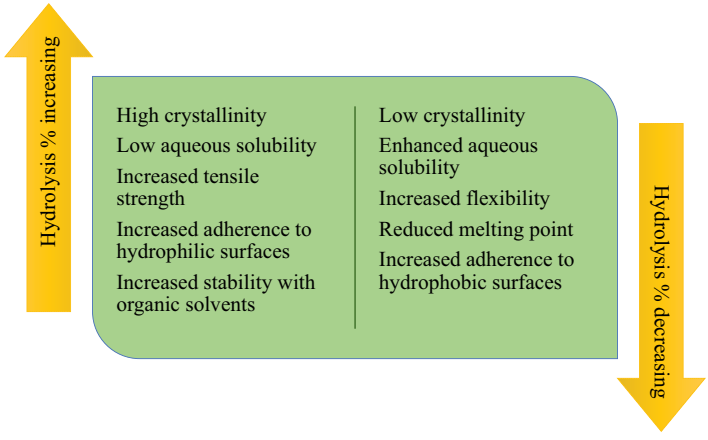



**Figure 2.** Effect of degree of hydrolysis on the properties of PVA.

### *3.1. Biodegradability and biocompatibility*

Biodegradable polymers undergo decomposition through the action of microorganisms, resulting in their conversion into water ($H_2O$) and carbon dioxide ($CO_2$). Plant systems subsequently absorb these by-products.[22] Biodegradability is the process by which organic substances are broken down by the enzymatic activity of living organisms, such as fungi, yeast, and bacteria, under appropriate environmental conditions. These conditions include factors like oxygen availability, temperature, and humidity.[23] In aerobic environments, the degradation process yields $CO_2$, biomass,

**Table 2.** Polyvinyl alcohols: Physical and Chemical properties.

| Chemical composition and physical properties | | |
|---|---|---|
| Parameter | Explanation | Reference |
| Molecular weight | 20,000–400,000 | [13] |
| Structural formula | $(\text{-CH}_2\text{CHOH})\text{-}_n\text{-}(\text{CH}_2\text{CHOCOCH}_3\text{-})\text{-}_m$ | [6] |
| Empirical formula | $(C_2H_4O)_n(C_4H_6O_2)_m$ | [6] |
| Solubility | Soluble in water, insoluble in aliphatic and aromatic hydrocarbons, as well as in oils, ketones, and esters | [17] |
| Tensile strength | 99% hydrolyzed PVA: 67–110 MPa<br>89% hydrolyzed PVA: 24–79 MPa | [13, 18] |
| Melting point | Partially hydrolyzed: 180°C-190°C<br>Fully hydrolyzed: 230°C | [6] |
| Boiling point | 228°C | [6] |
| Thermal stability | 173–473 K | [19] |
| Specific gravity | 1.19–1.31 | [17] |
| Polydispersity index for commercial grade PVA | 2–2.5 | [20] |
| Glass transition temperature | 65–85°C | [17] |
| Decomposition temperature | 220–250°C | [21] |

and water; in anaerobic conditions, it produces methane. The degradation should proceed without leaving harmful residues.[24] PVA is a water-soluble synthetic polymer characterized by its hydrophilic nature, which allows it to dissolve in water. However, dissolution in water does not necessarily equate to complete biodegradation. PVA requires specific conditions for degradation, including the presence of particular microorganisms and enzymes. In soil, PVA can be degraded by microbial activity, but the degradation rate depends on factors such as soil type and environmental conditions.[25, 26] A study examining the biodegradation and ecotoxicity of PVA in marine environments found that PVA exhibits limited biodegradability under natural marine conditions, with slight biodegradation attributed to glycerol in blended materials. It also highlighted the need for standardized testing methods to assess biodegradation and toxicity in marine ecosystems.[27] A case study on PVA with a viscosity of 18 mPa.s and a degree of hydrolysis of 88% demonstrated that the polymer undergoes substantial biodegradation during wastewater treatment, with a removal rate of 97.4%. The study also confirmed that PVA poses no significant aquatic toxicity, ensuring its environmental safety in freshwater ecosystems.[28]

PVA has been found to have excellent biocompatibility, making it well-suited for medical applications. Studies have indicated that PVA, particularly in the form of hydrogels, is accepted by biological tissues. Its natural biocompatibility with bodily organs makes it highly appropriate for various biomedical uses, such as artificial intestines, blood vessels, contact lenses, and drug delivery systems. Furthermore, PVA has demonstrated effectiveness as a material for wound dressing.[29] A study on grafted chitosan/PVA/AuNPs nanocomposites demonstrated excellent biocompatibility, as the materials exhibited strong antimicrobial activity comparable to Amoxicillin and effectively inhibited the A549 cancer cell line. Furthermore, toxicity tests confirmed that these nanocomposites are nontoxic, supporting their potential for biomedical applications.[30] Another study confirmed the biocompatibility of PVA in a bi-layer hydrogel for osteochondral defect repair, showing successful tissue regeneration and strong interfacial bonding without chemical crosslinking.[31] A recent study demonstrated the biocompatibility of boric acid (BA)-crosslinked PVA coatings used for paper packaging. In vitro tests with HEK 293 and MEF cells showed negligible cytotoxicity, while in vivo testing on mice revealed no adverse effects on vital organs after 30 days of exposure to sample-dissolved drinking water. These findings confirm the biocompatibility and environmental safety of BA-crosslinked PVA, making it an excellent choice for sustainable packaging solutions.[32]

### 3.2. Mechanical and thermal properties

PVA features a unique set of mechanical and thermal properties, making it a versatile choice for diverse applications. The tensile properties of pure PVA are largely determined by its molecular weight and degree of hydrolysis. Fully hydrolyzed PVA has been reported to possess a tensile strength of 1.6 GPa, Young's modulus of 48 GPa, and an elongation at break of 6.5%. In comparison, partially hydrolyzed PVA exhibits tensile strength, Young's modulus, and elongation values of 25.4 MPa, 27.6 MPa, and 260%, respectively.[13, 33] The addition of 6 wt% nanocrystalline cellulose (NCC) to PVA films enhances their mechanical and thermal properties. The tensile strength increases to 35.30 MPa and elongation to 65.54%. Thermal stability improves,

with the glass transition temperature rising from 109°C to 114°C and the maximum degradation temperature increasing from 275°C to 300°C.[34]

Crosslinking significantly influences the mechanical and thermal properties of PVA. Agents like glutaraldehyde (GA), boric acid, and citric acid form covalent bonds between PVA chains, enhancing their structural integrity and performance. Studies show that chemically crosslinked PVA membranes exhibit increased tensile strength and Young's modulus, with tensile strength improving as the degree of crosslinking increases. Specifically, PVA membranes crosslinked with 3 wt% GA demonstrate enhanced mechanical properties, where tensile strength increases with longer crosslinking times, while strain at break decreases due to the restricted mobility of the polymer chains caused by crosslinking.[35, 36] A study showed that boric acid greatly improves the tensile strength, reduces water absorption, and lowers the solubility of PVA films, enhancing their mechanical and thermal properties. Citric acid also improves these properties but is less effective than boric acid. Glycerol increases viscosity and water resistance but weakens the film structure and increases water absorption. Overall, boric acid and citric acid are more effective crosslinkers for PVA films than glycerol.[37]

## 4. PVA modifications

PVA has gained significant attention for its versatile nature and wide range of applications. It possesses outstanding properties such as chemical resistance, mechanical strength, film-forming abilities, and biocompatibility. To further enhance these properties, various chemical and physical modification strategies have been explored. These include the incorporation of nanofillers, blending PVA with other polymers, and fine-tuning its chemistry to meet specific application requirements.[38–40] Modifying PVA is also essential to reduce inter- and intramolecular interactions among the hydroxyl groups in its molecular chain. PVA can be modified in its bulk solution, within crosslinked structures, or on its surface, enabling its functionalization for a range of uses.[9] The modified form of PVA often appears as a terpolymer composed of vinyl acetate, vinyl alcohol, and functionalized vinyl units, which enables its adaptability for diverse applications.[41]

### 4.1. Nanocomposites

Nanocomposites are materials that contain at least one dimension on the nanometer scale. They offer enhanced performance and unique properties that result from combining different composites, which possess distinct physical and chemical characteristics.[42–44] These composites often exist in different phases separated by an interface. As a result, nanocomposites exhibit novel properties that cannot be achieved by their components alone. The larger component is referred to as the matrix, while the process of enhancing the properties of nanocomposites involves reinforcements. Reinforcement typically involves the addition of nanomaterials, such

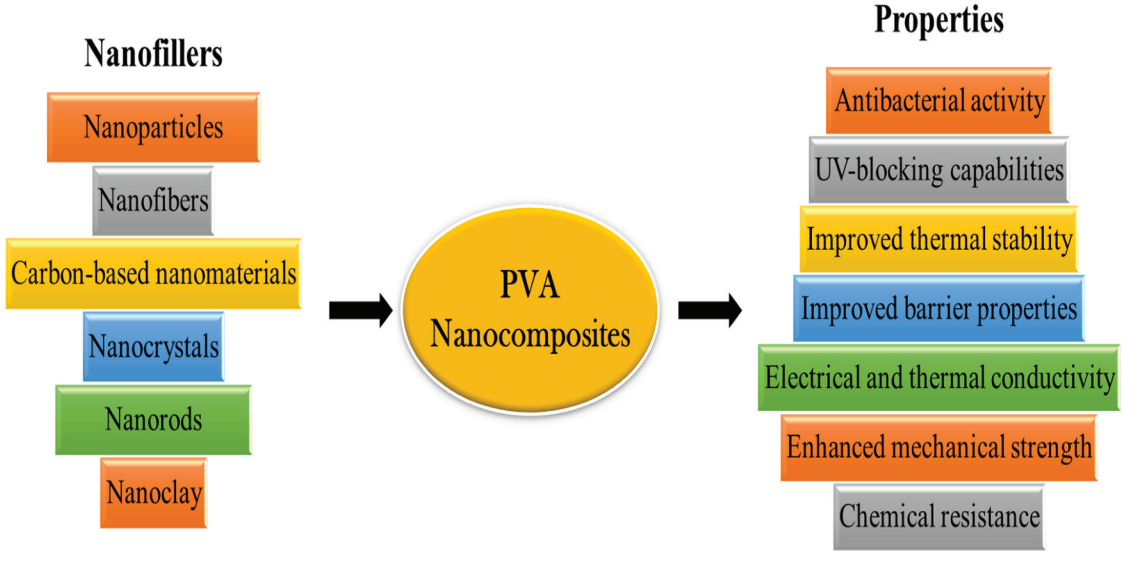



**Figure 3.** Nanofillers in PVA nanocomposites and enhanced properties.

as nano-fillers, into the matrix to improve overall performance.[45–47] Incorporating nanofillers into PVA matrices significantly enhances their mechanical properties, UV-blocking capabilities, antibacterial effects, and thermal stability (Figure 3).

#### 4.1.1. PVA/ZnO nanocomposites

PVA/ZnO nanocomposites exhibit improved optical, mechanical, and antibacterial properties, making them highly suitable for diverse applications. Incorporating ZnO nanoparticles enhances optical characteristics and crystallinity, although it slightly reduces the transparency of PVA films while significantly increasing their strength. Moreover, ZnO nanoparticles provide antibacterial properties, expanding the applications of PVA/ZnO composites in medical and packaging fields. Due to their enhanced functional properties, these nanocomposites are promising materials for advancements in optoelectronic and antibacterial applications.[48–50]

#### 4.1.2. PVA/CuO nanocomposites

PVA/CuO nanocomposites exhibit significant enhancements in several key properties due to CuO nanoparticles incorporation. The inclusion of CuO nanoparticles effectively reduces the optical band gap of PVA, leading to improved optical characteristics with altered absorption peaks, which enhances the suitability of the material for photonic applications.[51] Mechanically, the incorporation of CuO significantly increases the strength and flexibility of the PVA films, with optimal improvements observed at CuO concentrations up to 0.5 wt%. Additionally, the dielectric constant of the nanocomposites rises with increased CuO concentration, reflecting an increase in polarization and density of states. The electronic conductivity of the films also improves with higher CuO content, demonstrating greater ionic diffusivity and conductivity.[52, 53] The performance of PVA-based nanocomposites with metal oxide nanofillers is detailed in Table 3.

#### 4.1.3. PVA/graphene oxide (GO) nanocomposites

PVA/Graphene oxide (GO) nanocomposites exhibit significant improvements in mechanical and optical properties, rendering them highly suitable for advanced applications. The incorporation of GO into the PVA matrix leads to substantial enhancements, with the 2.5 wt% GO/PVA film showing a 43.65% increase in tensile strength and improved toughness. These enhancements are due to robust interactions between GO and PVA, including hydrogen bonding and the formation of intercalated nanolaminate structures. Additionally, the composites demonstrate excellent UV absorption capabilities, making them effective for UV shielding applications.[61–63]

**Table 3.** Comparative performance of PVA-based nanocomposites with metal oxide nanofillers.

| Nanocomposite type | Synthesis method | Mechanical performance | Functional properties | References |
|---|---|---|---|---|
| CuO/PVA | Solution casting | Tensile strength increased from 85.8 to 154.0 MPa | Antibacterial activity-*E. coli*: 2 wt% CuO reduced growth; 4 wt% CuO and higher inhibited growth | [54] |
| PVA-ZnO | Solvent casting | - | Transparency- Up to 65% with 3 wt% ZnO<br>Band gap- Narrows with added ZnO<br>Optical conductivity- Improved at 3 wt% ZnO | [55] |
| Acrylic acid/PVA/MgO | Gamma radiation | - | The inhibition zone increased from-<br>For *S. aureus*: 20.67 mm to 26.34 mm<br>For *E. coli*: 12 mm to 17.67 mm | [56] |
| PVA/gelatin methacrylate-$CeO_2$ | Coaxial electrospinning | Tensile strength- 8.49 MPa | Antibacterial activity- Reduced bacterial colonies by 88.5% for *S. aureus* and 93.1% for *E. coli*<br>Cell Viability- nontoxic to human dermal fibroblast cells | [57] |
| PVA-$TiO_2$ | Freezing-thawing | Compression Strength- Increased with $TiO_2$ and $Fe_2O_3$, especially at 0.1 wt%<br>Young's Modulus- Increased, peaking at 0.1 wt% $TiO_2$ | Improved porosity and uniform pore distribution, compatibility for wastewater treatment | [58] |
| PVA/CS/NPs-$SiO_2$ | Freeze-drying | Compressive Strength- Best at 1% $SiO_2$ (1.180 MPa)<br>Elastic Modulus- Best at 0.5% $SiO_2$ (0.0455 MPa)<br>Strain capacity- Best at 2% $SiO_2$ (243.4%) | Antibacterial activity-<br>*K. pneumoniae*: 25.0 ± 1.0 mg/ml<br>*E. cloacae*: 3.13 ± 0.13 mg/ml<br>*E. coli*: 3.13 ± 0.14 mg/ml<br>*S. enterica* ATCC 53648: 25.0 ± 1.0 mg/ml | [59] |
| Hyaluronic acid-PVA film loaded with $Ag_3(PO_4)$/$Sm_2O_3$/CNT | Casting method | Tensile strength- 15.19 MPa<br>Elongation at break- 98.7% | Biocompatibility- High cell viability with 125.8% at 8.8 µg/ml<br>Antibacterial activity- Shows 20.5 mm inhibition against *S. aureus* and 17.5 mm against *E. coli* | [60] |

#### 4.1.4. PVA/nanoclay nanocomposites

Recent studies on PVA/Nanoclay nanocomposites have highlighted significant enhancements in material properties with the incorporation of Nanoclay. The addition of Nanoclay to PVA hydrogels improved thermal stability and facilitated a more uniform dispersion of fertilizer, leading to optimized control release properties. Specifically, an optimal Nanoclay content of 0.5 wt% was found to enhance the water absorption levels, while higher concentrations reduced swelling and improved controlled release efficiency.[64] In food packaging applications, incorporating clay nanoparticles into PVA films improved thermal conductivity, water vapor transport rates, and antibacterial performance, despite a decrease in tensile strength and elongation.[65] These advancements indicate that PVA/Nanoclay nanocomposites offer promising benefits for applications requiring improved stability, controlled release, and enhanced mechanical and barrier properties.

#### 4.1.5. PVA/carbon nanotubes (CNTs) nanocomposites

PVA/carbon nanotubes (CNTs) nanocomposites offer significant enhancements in mechanical, thermal, and electrical properties. By incorporating multi-walled carbon nanotubes (MWCNTs) into crosslinked PVA matrices, significant improvements in mechanical strength, thermal stability, and shape recovery can be achieved. MWCNTs in PVA-based composites accelerate the breaking of C-O bonds and enhance yield strength while maintaining tunable shape recovery properties.[66] Functionalized CNTs (f-CNTs) and short carbon fiber ($C_f$) further improve mechanical behavior by enhancing crystallinity and polymer chain interactions. The addition of f-CNTs in particular results in superior strength gains due to better distribution and stronger interaction with PVA chains, compared to MWCNTs or $C_f$. Electrospun PVA/CNT composites also exhibit impressive electrical conductivity and biodegradability, with potential applications in nerve conduits, where they promote cell growth and viability. Moreover, single-walled carbon nanotubes (SWCNTs) embedded in PVA matrices enable a significant increase in anisotropic conductivity, determined by factors like SWCNTs concentration and alignment, offering broad potential in designing conductive polymer-based composites.[67–69]

#### 4.1.6. PVA/graphene nanocomposites

The integration of graphene into PVA matrices, through techniques like solution casting and electrospinning, significantly alters the crystallinity, thermal stability, and optical properties of the nanocomposites. Studies show that the incorporation of graphene reduces crystallinity and thermal stability due to restricted polymer chain mobility and synergistic instability. However, microwave irradiation for short durations improves these properties by enhancing crystallinity and thermal stability, although prolonged exposure leads to degradation. Graphene also influences the optical bandgap and reactivity of PVA.[70, 71] Additionally, Electrospun PVA/graphene nanocomposites demonstrate high porosity, improved hydrophilicity, and excellent biocompatibility, supporting cell adhesion and proliferation. This makes PVA/graphene nanocomposites suitable for applications in flexible electronics and tissue engineering, particularly in developing vascular grafts for clinical use.[72]

#### 4.1.7. PVA/nanocrystalline cellulose (NCC) nanocomposites

PVA nanocomposites reinforced with nanocrystalline cellulose (NCC) exhibit substantial improvements in mechanical, thermal, and barrier properties, making them highly suitable for packaging applications. The addition of NCC enhances the tensile strength, elongation, thermal strength, elongation, and thermal stability of PVA films, with optimized concentrations typically ranging between 6 wt% and 12 wt%. These nanocomposites show increased glass transition temperatures and maximum degradation temperatures, contributing to better performance under thermal stress. Additionally, the inclusion of NCC reduces water vapor permeability, solubility, and water absorption while improving the contact angle and UV-blocking capability. As an environmentally friendly reinforcement, NCC provides a biodegradable alternative to mineral fillers, offering enhanced mechanical strength, barrier properties, and thermal resilience without compromising transparency. This makes PVA/NCC composites an excellent choice for developing sustainable packaging materials with superior performance characteristics.[34, 73–75]

In addition to the discussed nanocomposites, several studies have explored various PVA-based nanocomposites that demonstrated significant improvements in mechanical, thermal, and functional properties. The table below summarizes some prominent examples, highlighting the types of nanofillers incorporated into PVA, the observed enhancements in properties, and their potential applications (Table 4).

### 4.2. Cross-linking

Due to its hydrophilic nature, PVA can be crosslinked through various techniques, with chemical crosslinking

**Table 4.** PVA-based nanocomposites, enhanced properties and applications.

| Matrix | Nanofiller | Enhanced properties | Applications | References |
|---|---|---|---|---|
| PVA | $CuO/SiO_2$ | Thermal and mechanical stability, antimicrobial properties, High photocatalytic efficiency | Wastewater treatment | [76] |
| PVA | Silver NPs | Improved conductivity and dielectric permittivity | Schottky diodes | [77] |
| PVA | GO, CuO | Optical and electrical properties | Optoelectronics | [78] |
| PVA | $TiO_2$, C NPs | Mechanical and thermal properties, high tensile strength | Alkaline direct methanol fuel cells (ADMFCs) | [79] |
| Modified PVA with plant extracts (Hemp seeds, St. John's wort, Rosemary, Cherry Laurel) | Silver NPs | Thermal resistance, Antibacterial properties, Mechanical properties, swelling behavior | Eco-friendly nanocomposite film | [80] |
| PVA/PEG | Aluminum Nitride NPs | Thermal and mechanical properties, optical properties, water resistance | UV-blocking films | [81] |
| PVA | Plasma functionalized-MWCNTs | Optical properties, conductivity, Thermal stability | Dielectric and conductive materials | [82] |
| Chitosan/PVA | Folic acid-functionalized graphene oxide | Mechanical properties, humidity sensing, dispersibility | Intelligent food packaging and preservation | [83] |
| PVA | $SnS_2$/SnS nanoflakes | Enhanced voltage and current output, power density | Portable self-powered devices | [84] |
| PVA/CMC | Graphene nanoplatelets, $TiO_2$ NPs | Thermal stability, optical properties, Mechanical properties, electrical conductivity and dielectric properties | Flexible micro and optoelectronic devices | [85] |
| PVA/CMC | Multi-walled CNTs, α-$Fe_2O_3$ | Optical and dielectric properties, electrical conductivity | Electrochemical and optoelectronic devices | [86] |
| PVA | PZT | Optical properties, photoluminescence | Optoelectronic applications | [87] |

and thermal methods being among the most commonly employed. Chemical crosslinking often involves the incorporation of acids, bases, or their mixtures, facilitating reactions between the free hydroxyl groups of the polymer and the functional group of the crosslinking agents. This process reduces water solubility, increases rigidity, and enhances the chemical stability of the polymer. Moreover, the addition of hydrophilic groups such as -$SO_3$-, or -COO- improves thermal stability and enhances ionic conductivity.[88, 89] The crosslinking of PVA nanofibers, particularly through a solution-based method, enhances their adsorption capacity for heavy metal ions like $Cu^{2+}$ and $pb^{2+}$, while also improving mechanical strength, thermal stability, and crystallinity.[90] A study demonstrated that heat treatment induces crosslinking in PVA, leading to enhanced crystallinity, optical, and mechanical properties. Optimal results were achieved by heating at 140°C for 40 minutes.[36]

The most commonly used crosslinkers for PVA include dialdehydes such as glutaraldehyde, [91] glyoxal, [92] and malic anhydride, [93] along with diisocyanates, [94] transglutaminase, [14] polyurethane, [95] genipin, [96] citric acid, and boric acid.[37] A study showed that crosslinking electrospun PVA fibers with citric acid enhances their water stability, mechanical properties, and thermal stability, with the melting temperature increasing from 170°C to 190°C. Crosslinked fibers remained stable in water for 72 hours and supported NIH 3T3 cell attachment and proliferation, showing potential for tissue engineering and other biomedical applications.[97] Co-crosslinking of PVA with glutaraldehyde and ammonium persulfate improved water resistance and thermal stability, reducing the swelling percentage from 590% to 129% and increasing the decomposition temperature from 242.9°C to 344.7°C.[98] Overall, crosslinking significantly enhances the properties of PVA, improving its mechanical strength, thermal stability, and water resistance. These modifications expand the material potential applications, particularly in fields requiring durability and stability, such as packaging, textiles, and biomedical engineering. The tailored crosslinked structures also offer improved performance in environments where traditional PVA films would degrade or lose functionality.

### 4.3. PVA blends with biopolymers

Blending refers to the physical mixing of polymeric materials without involving any chemical reactions.[99] It is an effective method for creating innovative materials with desirable properties. The goal of blending may include enhancing material performance, reducing water sensitivity, lowering production costs, or tailoring properties for specific applications. The primary advantage of blending lies in its ability to achieve these improvements more simply and flexibly than other modification techniques.[100, 101] PVA has a relatively

high material cost and a low biodegradation rate, making it necessary to blend with other polymers. PVA is highly compatible with several natural biopolymers, including chitosan, guar-gum, lignocellulose, and starch. Blending PVA with these biopolymers not only enhances its mechanical strength and thermal stability but also improves its biodegradability, making it more versatile for various applications.[102, 103]

### 4.3.1. PVA/Chitosan blends

The blend of PVA with chitosan offers significant improvements in both biodegradability and functionality. Chitosan, a biopolymer derived from chitin, has a linear structure composed of glucosamine and N-acetylglucosamine units (Figure 4). Its amino groups give it strong antimicrobial properties, making it effective against a range of bacteria and fungi. When combined with PVA, which is known for its excellent film-forming properties and hydrophilicity, the resulting blend shows enhanced mechanical strength and flexibility. This is due to strong hydrogen bonding between the hydroxyl groups of PVA and the amine groups of chitosan, leading to a more durable and functional material.[104–106]

Recent research on PVA/Chitosan blends reveals significant advancements and applications. The fabrication of CuO-doped PVA/Chitosan blends using the solution casting technique has shown that incorporating CuO particles enhances optical properties, including reduced transmittance and optical bandgap, making these blends suitable for optical device applications. The uniform dispersion of CuO particles within the PVA/chitosan matrix was confirmed through FT-IR and scanning electron microscopy (SEM) analysis.[107] Chitosan/PVA blend membranes with decorated graphene oxide (GO), antibacterial rosin, and Ag NPs were fabricated via solution casting. The addition of 2 wt.% decorated GO enhanced the tensile strength to □48 MPa compared to 33 MPa for the unfilled blend. The chitosan/PVA-3% NPs membrane showed higher adsorption for heavy metals (85–99.7% removal) and demonstrated good thermal stability and antibacterial properties.[108] Blending PVA and chitosan with edible fillers like Zein (ZN) and starch can improve their suitability for food packaging. This approach enhances the chemical and physical properties of the resulting edible films. This development contributes to more sustainable food packaging solutions by reducing pollution and advancing the creation of eco-friendly materials.[109]

### 4.3.2. PVA/Starch blends

Blending PVA with starch offers a promising approach to developing biodegradable materials that harness the advantages of both polymers.[110] Starch, a natural polymer sourced from materials like corn, potatoes, and cassava, consists of amylose, a linear chain of glucose units linked by α-1,4-glycosidic bonds, and amylopectin, a branched polymer with α-1,4 and α-1,6-glycosidic linkages (Figure 5).[111] This unique structure enables starch to form gels and films, making it an ideal candidate for blending with PVA. Blending PVA with starch results in several advantageous properties. The addition

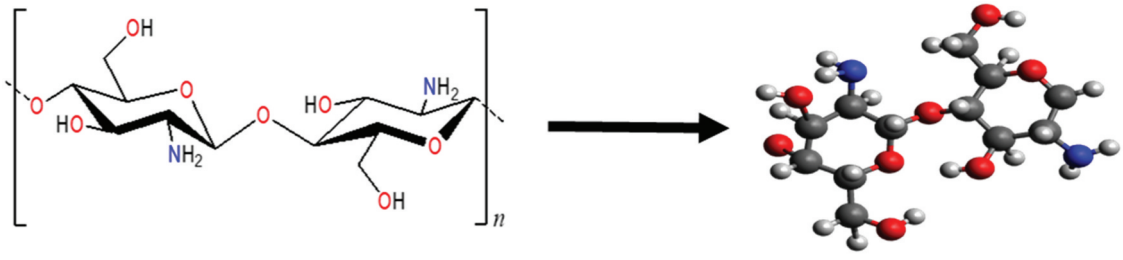



**Figure 4.** Structural representation of chitosan.

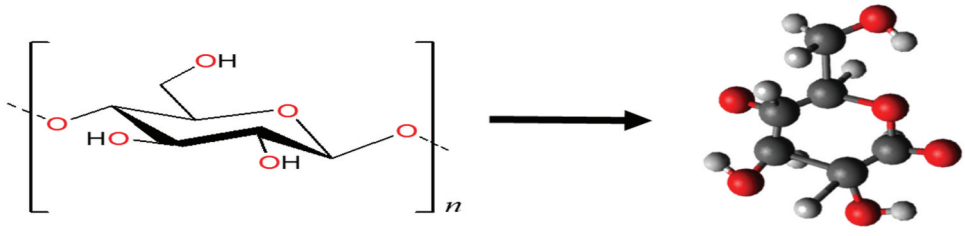



**Figure 5.** Structural representation of Starch.

of PVA improves the mechanical strength and thermal stability of starch-based materials, overcoming the primary drawback of pure starch films, which typically have weak mechanical properties. Research shows that adjusting the PVA-to-starch ratio can optimize these attributes, with higher PVA content generally leading to enhanced tensile strength and flexibility. Moreover, using citric acid as a cross-linking agent in the blend significantly boosts the interfacial adhesion between PVA and starch, enhancing overall film performance through increased hydrogen bonding interactions.-[112–116]

PVA/starch blends show different properties depending on the type of starch used. Pregelatinized starch (PST) enhances tensile strength but reduces elongation at break, while hydrolyzed starch (HST) improves elongation and slightly increases tensile strength. HST also significantly increases water solubility due to its molecular structure, while interactions between starch, PVA, and glycerol have little effect on the solubility of the blends.[117] PVA/starch blends exhibit enhanced physical and mechanical properties when combined with additives like polyurethane (PU) or modified through blending. Small amounts of PU improve the amorphous nature of the blend, reduce glass transition temperature, and enhance plasticization. Adding PVA to starch improves mechanical strength and water solubility, making the blends suitable for packaging applications in the food industry due to their biodegradability and eco-friendly characteristics. These blends are compatible with improved permeability and flexibility, making them ideal for sustainable material development.[118, 119]

### 4.3.3. PVA/cellulose blends

Cellulose is a linear polysaccharide made up of D-glucose units linked by β-1,4 bonds (Figure 6). Its structure is highly crystalline, resulting from hydrogen bonding both within and between molecules. Cellulose exists in several polymeric forms, with cellulose I being the native form commonly found in plants.[120] It can be isolated in different forms such as microcrystalline cellulose (MCC), cellulose nanofibers (CNFs), and cellulose nanocrystals (CNCs). Blending cellulose with PVA significantly improves the material properties.[121] The hydroxyl groups in both PVA and cellulose facilitate hydrogen bonding, enhancing the blend miscibility. Research shows that increasing cellulose content in PVA/cellulose composites boosts mechanical strength and shape memory recovery. For instance, one study revealed that incorporating 20–25 wt.% cellulose into PVA improved the curl-shaped recovery performance of the composite films, making them suitable for applications like hair styling products.[122]

Blending chemically modified microcrystalline cellulose with PVA improves the mechanical, barrier, and $CO_2$-releasing properties of the film, making it suitable for active packaging, especially for meat storage. $CO_2$ release is influenced by hydronium ions and relative humidity.[123] PVA/cellulose composite films, synthesized from sugarcane bagasse, showed dye removal capabilities of 70% for crystal violet and 64.5% for methylene blue. Biodegradation studies revealed a higher degradation rate for the PVA/cellulose blend (43%) compared to PVA (35%), making the blend a sustainable option for dye removal and environmental applications.[124] PVA/cellulose blends offer improved mechanical and functional properties, with enhanced biodegradability and ecological sustainability. Their diverse applications in packaging, waste treatment, and drug delivery highlight the potential of these blends in various industries.

### 4.3.4. PVA/Agar blends

Agar is a natural polysaccharide derived from red algae, primarily made up of two components: agarose and agaropectin. Agarose consists of a linear chain of repeating β-D-galactose and 3,6-anhydro-α-L-galactose units, connected by alternating glycosidic bonds (Figure 7). This structure enables agar to form gels upon cooling, making it useful in applications application like microbiological media and food products. Agaropectin, a branched polymer, contributes to the viscosity and gel strength of agar.-[125] Blending PVA with agar enhances the properties of the composite materials. This blend improves the mechanical strength and thermal stability of the films

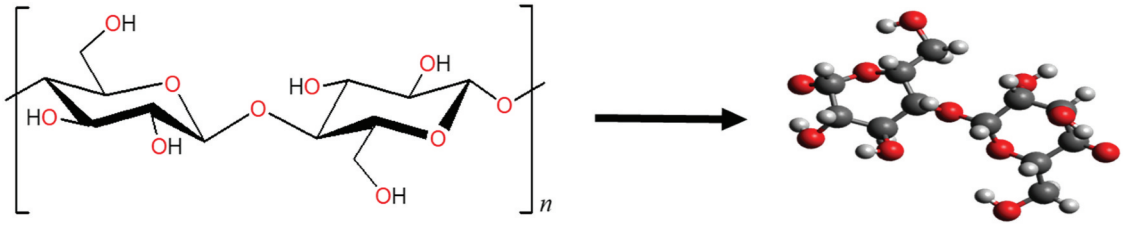



**Figure 6.** Structural representation of Cellulose.

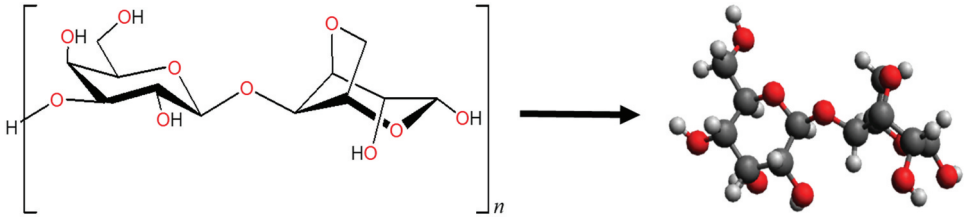



**Figure 7.** Structural representation of Agar.

while maintaining the hydrophilic characteristics of both polymers. The interaction between PVA and agar occurs mainly through hydrogen bonding due to the presence of hydroxyl groups in both materials.[126]

PVA/agar blend hydrogels, crosslinked with tannic acid and subjected to multiple freeze/thaw cycles, show enhanced mechanical properties and reduced swelling. Tensile strength increased from 1.56 MPa to 3.77 MPa with more cycles and a lower crosslinker amount. The hydrogels also exhibited excellent water retention and adhesion, making them suitable for biomedical applications like wound dressing and artificial cartilage.[127] Agar/PVA aerogels, produced using cross-linking and freeze-drying, effectively separate oily substances and organic solvents from water. They show high hydrophobicity, with excellent absorption capacities and separation efficiencies, making them suitable for oil-water separation and spill cleanup.[128] The blending of PVA with agar results in versatile materials that combine the beneficial properties of both polymers. This synergy not only enhances the performance characteristics of hydrogels and films but also opens up new avenues for innovation across various fields while addressing environmental concerns associated with traditional plastic usage.[129]

### 4.3.5. PVA/PLA blends

Poly (lactic acid) (PLA), as shown in Figure 8, is a biocompatible and biodegradable polymer that degrades into lactic acid monomers within the body. Sourced from renewable resources, PLA is valued for its durability and cost-effectiveness. When blended with PVA, PLA significantly enhances the tensile strength and Young's modulus of the material, particularly in wet conditions. Additionally, its hydrophobic nature improves the water resistance of PVA-based materials, making it an effective additive for strengthening and stabilizing PVA in various applications.[130, 131]

A study found that the PLA/PVA blend in a 1:1 ratio exhibited the highest porosity (43.4%), increased crystallinity, and improved absorption properties. However, higher porosity resulted in a significant decrease in flexural and tensile strength.[132] The addition of glycerol and lactic acid to PVA/PLA blends improved compatibility, thermal stability, and mechanical properties. These blends showed reduced water absorption, better surface resistivity, and increased transmittance, making them promising barrier materials for food and pharmaceutical packaging.[133] The multilayer films, with a PVA core and PLA outer layers, showed enhanced tensile and barrier properties. Incorporating carvacrol or ferulic acid into the PVA layer provided antimicrobial activity. The films were found to be suitable for food packaging and preservation.[134] Biodegradable PVA/PLA/$TiO_2$ films were developed with enhanced thermal stability, water resistance, and antibacterial activity. PVA and PLA- and PLA-based materials hold significant promise for diverse applications, from extending the shelf life in food packaging to serving in biomedical devices, due to their adaptable properties and potential for targeted functionality. The incorporation of nanomaterials, plasticizers, and natural fillers not only improves mechanical and antimicrobial properties but also addresses biodegradability concerns.[135]

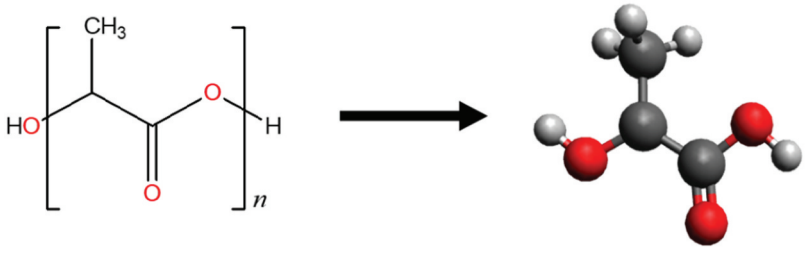



**Figure 8.** Structural representation of PLA.

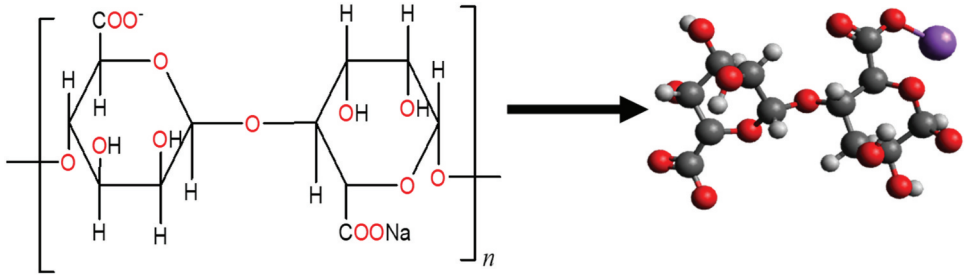



**Figure 9.** Structural representation of sodium alginate.

#### 4.3.6. PVA/alginate blends

Alginate is a linear polysaccharide made up of beta-D-mannuronic acid and alpha-L-guluronic acid units, which are primarily found in the cell walls of brown seaweeds (Figure 9).[136] Sodium alginate (SA), when blended with PVA, has been employed to enhance the clinical properties of PVA hydrogels for wound dressing applications. This polymer exhibits excellent protein absorption, rapid biodegradability, high hydrophilicity, antibacterial properties, and effective hemostatic and biological characteristics. As a result, PVA/SA hydrogels combine the advantageous properties of both polymers, with PVA contributing strong mechanical properties and SA enhancing the physical and biological characteristics of the sample.[137]

The Alginate-PVA biopolymer blend electrolytes, doped with $NH_4NO_3$, showed improved ionic conductivity, thermal stability, and an enhanced amorphous structure. With 35 wt% $NH_4NO_3$, the highest conductivity of $5.20 \times 10^{-4}$ S $cm^{-1}$ was achieved, indicating potential for efficient ion-conducting applications.[138] Alginate-PVA nanocomposite fibers with silver nanoparticles demonstrated improved swelling, mechanical properties, and effective antibacterial activity against *S. aureus* and *E. coli*, making them promising for antibacterial wound dressings. The fibers can be stored and applied in dry form while retaining their properties.[139] The integration of these materials enhances the physical, mechanical, and biological properties of the composite, making it well-suited for applications such as wound dressing and drug delivery systems.[140]

### 4.4. PVA nanocomposite morphologies

The PVA matrix has proven to be a versatile foundation, showcasing a wide range of morphological possibilities with significant implications across various applications (Table 5). Research has extensively explored the diverse morphologies of PVA, as illustrated in Figure 10 while certain structures can be achieved with unaltered PVA, the range of possible morphologies expands significantly with modifications in composition and processing techniques.

#### 4.4.1. Hydrogels

Hydrogels, known for their three-dimensional cross-linked polymer networks, are formed from PVA through physical or chemical crosslinking. Enhancements in mechanical properties can be achieved by incorporating additives such as carbon nanotubes or metallic particles.[156, 157]

#### 4.4.2. Films

PVA films are valued for their optical transparency, flexibility, and dielectric strength. The incorporation of nanoparticles into PVA matrices improves these properties and enhances the overall performance of the film.[158–160]

#### 4.4.3. Membranes

PVA membranes, characterized by their chemical stability and high hydrophilicity, are used in various separation processes, including gas separation water treatment. Cross-linking can improve their performance by reducing excessive Swelling.[145, 161, 162]

#### 4.4.4. Scaffolds

In tissue engineering, PVA-based scaffolds are used for their biocompatibility and hydrophilic properties. Traditional fabrication methods include solvent casting and electrospinning, while selective laser sintering offers an advanced approach to creating these structures. Tissue engineering, a key branch of regenerative medicine, relies heavily on scaffolds to support tissue repair. These scaffolds play a critical role in promoting tissue regeneration by providing an ideal environment that accelerates the healing process, thereby contributing significantly to life-saving therapies.[163, 164]

#### 4.4.5. Fibers

PVA fibers, produced via techniques like electrospinning, exhibit high tensile strength and resistance to

**Table 5.** Comprehensive overview of PVA nanocomposite morphologies.

| Morphology | Production Method | Operational Parameters | Common Applications | References |
|---|---|---|---|---|
| Hydrogel | Physical crosslinking, chemical crosslinking, UV radiation crosslinking, UV polymerization | Polymer concentration, Cross-linking density, pH, reaction time, Choice of crosslinker | Tissue engineering, biomedicine, wound dressing, water processing | [141, 142] |
| Film | Solution casting, wet extrusion blowing film method, dry extrusion film method | Polymer concentration, Cross-linking condition, Evaporation rate, polymerization conditions | Food packaging, drug delivery systems, wound dressing, construction materials, energy storage devices | [143, 144] |
| Membrane | Phase inversion, electrospinning, melt-spinning, cold-stretching, 3D printing | Polymer concentration, temperature control, drying temperature and time, crosslinking condition, thickness | Water treatment, gas separation, fuel cells, targeted drug delivery systems, controlled molecule release | [145–147 |
| Scaffold | Freeze drying, 3D printing, electrospinning, gas foaming | Polymer concentration, cross-linking agents, temperature control, foaming time | Tissue engineering, drug delivery systems, wound dressings, regenerative medicine | [148–150] |
| Fiber | Wet spinning, electrospinning, melt spinning, dry-wet spinning | Polymer concentration, temperature control, spinning speed, choice of solvent | Textiles, concrete reinforcement, adhesives, paints, drug delivery | [151, 152] |
| Beads | Spray drying, ionic gelation, emulsification | Polymer concentration, pH, stirring speed, temperature control, crosslinking concentration | Separation media, water treatment, controlled drug delivery, catalyst support | [153–155] |

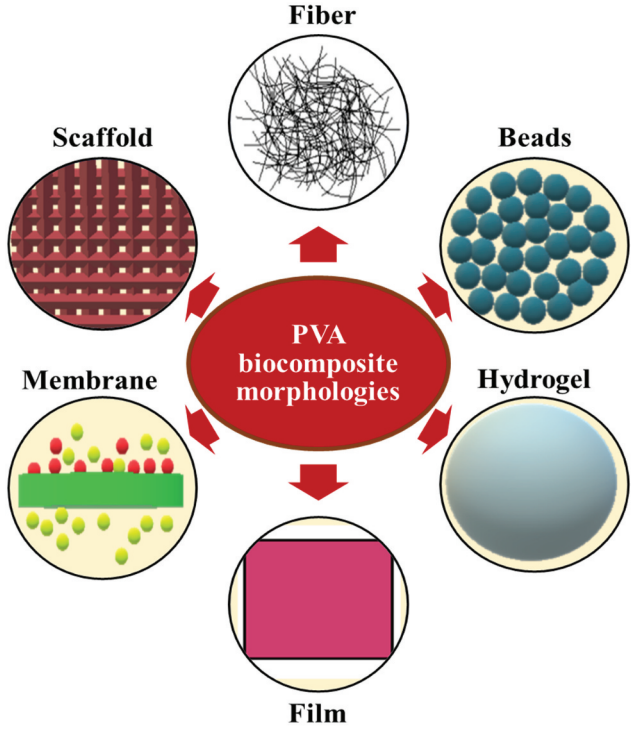



**Figure 10.** Various morphologies of PVA-based materials.

fatigue, making them suitable for structural reinforcement and various applications.[165, 166] The electrospinning method allows for the creation of nanofibers with controllable diameters.[167, 168]

#### *4.4.6. Beads*

PVA hydrogel beads are used in wastewater treatment and controlled-release applications. Electrospraying produces microbeads with small diameters, enhancing their effectiveness in biological treatments.[169, 170]

## 5. Applications of PVA-based materials

PVA-based materials find broad applications across several fields due to their unique properties, such as biocompatibility, bio-inertness, water solubility, and excellent film-forming ability. In packaging, PVA is commonly used as a moisture barrier for food supplements and products with moisture-sensitive components, protecting active ingredients from environmental factors while maintaining product quality.[171, 172] In the biomedical sector, PVA hydrogels

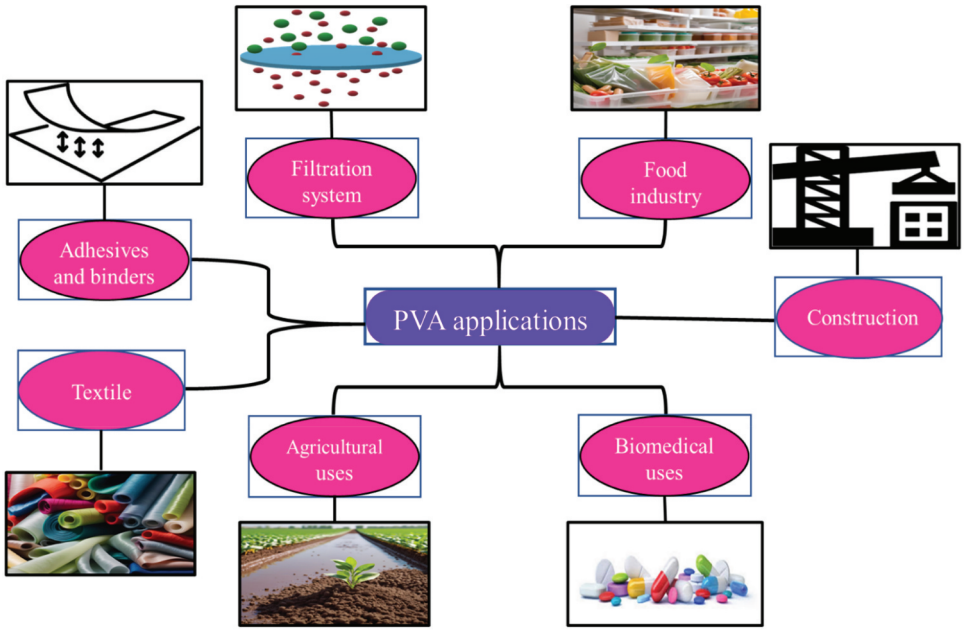



**Figure 11.** Diverse applications of polyvinyl alcohol.

are extensively used in wound dressing, drug delivery systems, contact lenses, tissue engineering, and implantable medical devices, offering bio-adhesive, nontoxic, and elastic properties that mimic natural tissues.[173–175] Additionally, PVA is utilized in environmental applications, such as wastewater treatment, where PVA-based nanocomposites help in the removal of contaminants like dyes and heavy metals. Its versatility makes PVA an important material in various industrial applications (Figure 11).[176] Below are detailed applications of PVA-based materials across different fields.

## 5.1. Applications in packaging

The application of PVA in food packaging has gained prominence due to its beneficial properties and environmental advantages. PVA, a biopolymer known for its mechanical strength, biodegradability, and chemical resistance, is increasingly being used to develop innovative food packaging solutions. It serves as a superior alternative to synthetic polymers, offering enhanced film formation and environmental sustainability. One remarkable application of PVA is in edible coatings, where it forms a thin moisture-resistant layer directly on food surfaces.[177–179] This coating helps protect food from mechanical and chemical damage, minimizes water loss, and controls microbial activity. By incorporating antioxidants and antimicrobials, PVA coatings create a modified atmosphere around the food, which slows down respiration and maturation rates. The ability of PVA to integrate active ingredients further enhances its role in creating biodegradable, active food packaging materials, making it a valuable asset for maintaining food quality and safety.[180, 181]

### 5.1.1. Barrier properties

PVA-based nanocomposite films demonstrate promising barrier properties for food packaging applications, but their performance is highly dependent on environmental conditions and the type of food. Its sensitivity to moisture is evident, as oxygen permeability increases sharply at relative humidity levels above 60%. However, this deterioration can be reversible in the presence of certain simulants like deionized water and 3% acetic acid. However, exposure to 50% ethanol, a fatty food simulant, leads to irreversible structural changes and a permanent reduction in oxygen barrier properties.[182] Incorporating nanofillers, such as montmorillonite (MMT) and titanium oxide (TiO2), improves the oxygen barrier and water vapor resistance, with an optimal loading of 1 wt.% $TiO_2$ and 4 wt.% MMT. These nanofillers also enhance the mechanical properties, increasing stiffness but reducing tensile strength and flexibility. While PVA-based films are suitable for packaging foods with low moisture and acidity, their effectiveness may be limited in packaging fatty or alcohol-containing foods.[183] The incorporation of cellulose nanocrystals (CNC) and $TiO_2$ nanoparticles into PVA films has been shown to significantly improve their mechanical strength, UV and water vapor barrier properties, and antimicrobial effectiveness, making them more suitable for food packaging applications.[184]

Oxalic acid-modified PVA coated with $Ca^{2+}$ showed improved water resistance and barrier properties for paper-based food packaging, achieving an OTR of 0.13 mL/$m^2$.day and WVTR of 0.82 g/$m^2$.day. The multilayer-coated paper was 99.7% repulpable and 91.8% biodegradable, indicating its eco-friendliness.[185]

A study developed a three-layer active film using PVA with chlorogenic acid (CGA) and CGA-functionalized layered (LDHs@CGA) on a PLA base. With 3 wt.% LDHs@CGA, the film achieved 85.6% antioxidant, 87.2% antibacterial activity, and improved gas barrier properties, making it suitable for food packaging.[186] $Fe^{3+}$ ions were added to PVA/cellulose films, creating a coordination network that significantly improved water resistance, barrier properties, and tensile strength. The films also demonstrated good recyclability, making them suitable for flexible packaging applications.[187] A novel antimicrobial packaging film was developed using a Nanoclay-essential oil composite with maize starch and PVA, showing superior water barrier properties and high tensile strength, making it suitable for food packaging. The film effectively extended the shelf life of button mushrooms up to 14 days in refrigerated conditions.[188] PVA provides a strong foundation for barrier properties in food packaging. The integration of nanoparticles significantly enhances its performance, improving barrier efficiency and adding functional benefits such as antimicrobial activity and UV protection. This combination positions PVA-based nanocomposite films as a promising solution for sustainable and effective food packaging applications.

#### 5.1.2. Biodegradable packaging

PVA is recognized as one of the most popular synthetic biopolymers for packaging applications due to its biodegradability and versatility. Blending PVA with biopolymers, such as starch, can significantly enhance its biodegradability and overall performance. Studies have shown that starch and PVA blends possess substantial potential for manufacturing biodegradable food packaging films with excellent properties.[13, 189] The ongoing pursuit of balancing biodegradability and mechanical performance in PVA and biopolymer blends remains a prominent area of research. While PVA offers considerable mechanical strength to packaging materials, achieving the desired environmental performance without compromising functionality is crucial. Recent studies suggest that blending with biopolymers can significantly enhance mechanical properties while maintaining biodegradability, offering a sustainable approach to addressing the challenges of biodegradable packaging.[113, 190, 191]

### 5.2. Biomedical applications

PVA is widely used in biomedical applications due to its biocompatibility, biodegradability, hydrophilicity, nontoxicity, and chemical resistance. It is prevalent in wound dressing and tissue engineering.[192] However, the lack of inherent bioactivity in PVA limits its effectiveness in treating infected or chronic wounds. PVA is often blended with natural polymers or bioactive molecules to enhance its therapeutic properties, resulting in functional nanofiber scaffolds with improved wound-healing capabilities. These modified nanofibers have shown promise in wound dressings, drug delivery systems, and tissue regeneration applications.[193, 194] PVA hydrogels are highly effective in biomedical and pharmaceutical applications due to their exceptional chemical and physical properties. These polymeric materials, based on PVA, hold significant promise for a wide range of biomedical uses.[173, 195] Currently, PVA is one of the most widely used and oldest polymers in applications such as wound dressings, [196] drug delivery systems, [197] artificial organs, [198] and contact lenses.[199]

#### 5.2.1. Wound dressings

PVA is an effective material for wound dressings, capable of maintaining a moist environment, improving the physical properties of the dressing, and promoting faster wound healing. An ideal wound dressing should meet several key criteria, such as providing a moist environment and being biocompatible to ensure compatibility with the body tissues.[29] Additionally, PVA possesses strong mechanical properties that facilitate cell adhesion, proliferation, and migration, making it well-suited for biomedical applications.[200] Combining PVA with other biocompatible materials, such as natural polymers and nanoparticles, significantly enhances its properties. This blending seeks to improve swelling capacity, elasticity, and thermal stability, all of which contribute to better wound healing capabilities.[201] Hydrogel incorporating Aloe vera into a sodium alginate/PVA matrix demonstrated improved thermal stability, prolonged active substance release, and enhanced cell adhesion and proliferation, making them promising candidates for interactive wound dressings.[202] The mechanical properties and adhesive qualities of PVA-based hydrogels are vital for their use in wound dressings. Adding materials such as kaolin can enhance the adhesive capacity of PVA hydrogels. Specifically, incorporating kaolin has been found to improve the adhesive strength of PVA composites.[203] Incorporating chitosan-copper complexes into PVA-based wound dressings enhances their antibacterial properties and biocompatibility.

This approach highlights the potential of integrating metal ions like copper into PVA composites to improve the efficacy of wound care solutions.[204]

#### 5.2.2. Drug delivery systems

PVA is widely used in drug delivery for enhancing solubility and preventing crystal transformation. It acts as a granulating liquid, slows crystal growth, and absorbs excess water. Employed as an embolic agent since the 1970s, PVA is effective in microparticle form. Its copolymers enable the controlled delivery of hydrophilic and hydrophobic drugs. These properties make PVA essential for safe and effective drug delivery systems.[7, 205] The development of buprenorphine-loaded PVA and PVP nanofibers highlights their promise as effective transdermal drug delivery systems. Electrospinning produced nanofiber mats with enhanced properties due to PVA addition, resulting in improved hydrophilicity and controlled drug release. These nanofibers can serve as efficient biomaterial patches for pain management in transdermal applications.[206, 207] Biocompatible patch-type drug delivery systems offer an alternative to traditional injections. A mucoadhesive patch made from PVA and guanidinium oleate, a pH-responsive ionic liquid, exhibits strong mechanical and adhesive properties. When loaded with the anticancer drug 5-fluorouracil (5-FU), it can release up to 60% of the drug within 48 hours at pH 7.2, demonstrating its potential for targeted and localized drug delivery.[208] The $Fe^{2+/3+}$ integrated PVA hydrogels demonstrated super-paramagnetic properties, controlled drug release, and over 60% cell viability, making them suitable for drug delivery applications.[209]

PVA nanofibers have been effectively used in ocular drug delivery, such as in the sustained release of latanoprost for glaucoma treatment. The nanofibers provided controlled drug release over 23 days, demonstrating the potential for use in drug-eluting contact lenses. This approach offers a promising alternative for enhancing patient compliance and therapeutic efficacy in eye-related conditions.[210] PVA-based contact lenses were developed to deliver vancomycin for treating severe bacterial keratitis. The lenses showed sustained drug release for 48 hours and effectively inhibited both Gram-positive and Gram-negative bacteria, enhancing vancomycin's therapeutic potential.[211] PVA enables controlled drug release, enhances stability, and improves therapeutic efficacy, making it ideal for drug delivery applications.

#### 5.2.3. Tissue engineering

PVA is widely researched for scaffold development due to its high mechanical strength, resistance to deformation, excellent water retention capacity, and porous structure. Furthermore, the abundant hydroxyl groups in PVA facilitate modifications with diverse functional groups, increasing its versatility for targeted applications.[150] PVA hydrogels are widely used in tissue engineering due to their biocompatibility and tunable mechanical properties. Combining them with natural materials like hyaluronic acid and chitosan enhances cell interactions and better replicates native cartilage, making them ideal for developing synthetic cartilage scaffolds.[212] Optimized sodium alginate (SA) nanofiber scaffolds show significant potential for tissue engineering due to their excellent physical, mechanical, and biological properties. These scaffolds support high cell viability, making them promising for applications in wound healing and skin regeneration.[213] A bilayer material combining Electrospun PVA-bacterial cellulose nanofibers and a 3D gelatin-PVA hydrogel demonstrated excellent potential for wound healing and soft tissue engineering. The material showed high swelling capacity, good stability, and controlled release of Ag-sulfadiazine with effective antibacterial activity against *E. coli, S. aureus*, and *P. aeruginosa*. In vitro studies indicated favorable cell viability, proliferation, and adhesion, highlighting its suitability as a multifunctional wound dressing for tissue regeneration.[214]

### 5.3. Environmental applications

PVA shows great promise in environmental remediation, especially in wastewater treatment. Its inherent properties enable effective contaminant interaction, and modifications with nanoparticles and cross-linking further enhance its performance and sustainability. Using PVA in adsorption-based systems improves water purification and helps mitigate plastic pollution.[215, 216] PVA-derived materials, such as gels, fibers, films, composites, membranes, and photocatalysts, have shown significant potential in environmental remediation, particularly in the treatment of dye-contaminated wastewater. These materials, often combined with conventional adsorbents and other components, are effective in the adsorption, filtration, and degradation process. Their versatile properties and scalable applications make them promising candidates for addressing global water pollution challenges.[217] The PVA-co-poly (methacrylic acid) adsorbent, synthesized via free radical polymerization, shows significant promise for environmental remediation. It effectively removes methylene blue from wastewater, demonstrating high adsorption capacity and excellent regeneration efficiency over multiple cycles. This makes it a cost-effective and sustainable solution for dye removal, highlighting the potential of PVA-based materials in water treatment applications.[218] Chitosan-based

**Table 6.** Summarized applications of PVA-based materials in flexible electronics.

| Material | Properties | Application | Reference |
|---|---|---|---|
| Poly (vinyl alcohol-*co*-ethylene) nanofibrous membrane | Visible light transmittance up to 85%<br>Very low CTE value ($3.634 \times 10^{-6}$ $K^{-1}$)<br>High compliance matrix | Alternating current electroluminescent device | [227] |
| Poly(β-cyclodextrin)/PVA/Gelatin/CNTs composite hydrogel | High healing efficiency of 95.0%<br>High stretchability of 778%<br>Mechanical strength of 54.6 kPa<br>Gauge factor up to16.0 | Strain sensors, wearable devices, electronic skins, soft robotics | [228] |
| PVA/Sericin/Boehmite nanocomposite | Increased dielectric constant and AC conductivity<br>Lower activation energy from AC conductivity with rising temperature | Optoelectronic devices, flexible electrochemical gadgets | [229] |
| PEDOT:PSS/PVA/PET film | High conductivity of 38.2 Ω/sq<br>Optical transparency of 84.1% | Flexible and transparent electronics, human-machine interactive devices | [230] |
| PVA@MXene@PPy hydrogel | Conductivity increases from 0.60 S/m to 0.74 S/m with 4% to 12% MXene content<br>Elongation up to 4300%<br>Self-healing and self-adhesion | Strain sensing and monitoring of physiological signals, energy storage in supercapacitors | [231] |
| PVA-Liquid metal composite ink | Conductivity of $1.3 \times 10^5$ S/m<br>High sensitivity and stable signal generation | Printable and recyclable ink for flexible electronics, strain sensors for monitoring human movements | [232] |
| PVA-HEC ionic conductive organic hydrogel | Electrical conductivity of 5.77 S/m when immersed in sodium chloride solution<br>Strength of 2.86 MPa at 400.30% strain | Self-powering, miniaturized, and intelligent flexible electronic systems | [233] |
| Gellan gum/PVA composite hydrogel | High compressive stress (10.7 MPa) at 90% strain<br>High conductivity of 3.27 S/m<br>Fracture energy of 480 $KJ/m^3$<br>Stress recovery rate 86% after 10 cycles at 200% strain | Biodegradable conductive hydrogels for sustainable flexible electronics | [234] |

PVA derivatives effectively remove heavy metals like Zn, Fe, and Pb from industrial wastewater, showing strong adsorption capacities and good regeneration potential, making them suitable for sustainable treatment applications.[219] A polyaniline/PVA/sodium alginate composite effectively removes hexavalent chromium (Ⅵ) from both water and soil, demonstrating strong adsorption capabilities and reusability.[220] Additive manufacturing has enabled the creation of complex, porous PVA-based scaffolds for effective heavy metal removal from wastewater. These 3D-printed structures exhibit high adsorption capacity, mechanical robustness, and rapid responsiveness, offering a promising solution for sustainable water purification.[221]

## 5.4. Emerging applications

### 5.4.1. Applications in sustainable electronics and energy storage

PVA is gaining recognition as a highly adaptable material with substantial applications in diverse areas, especially in electronics, energy storage, and smart materials. PVA is gaining traction in the electronics industry due to its outstanding dielectric strength, moisture resistance, and charge storage capabilities. It is used in various applications, including temporary protective coatings, release liners, and polymer electrolytes for flexible electrochemical devices.[222, 223] Research on PVA-multi-walled carbon nanotubes (MWCNTs) composites shows their effectiveness in triboelectric nanogenerators (TEGNs). The TENG with 1.6 wt.% MWCNTs achieved a maximum power output of 542 μW, highlighting the potential for self-powered systems like smart doors in various settings.[224] PVA/polyaniline conducting hydrogels show promise as electrode materials for supercapacitors with about 95% stability after 1000 cycles and as enzymatic glucose biosensors with a detection limit of 0.001 μM, highlighting their versatility in flexible electro-conductive applications.[225] The novel freestanding wireless humidity sensor, made of crosslinked PVA on a magnetostrictive particle platform, is designed for remote detection.[226]

*5.4.1.1. Flexible electronics:.* Flexible solid-state electrochemical capacitors (ECs) have garnered significant interest in recent years due to their advantages. Several studies on hydrogel electrolytes for supercapacitor applications have highlighted that PVA is a promising polymer due to its excellent chemical stability, nontoxicity, and biodegradability (Table 6).[235] PVA-based hydrogels demonstrate significant potential for wearable electronics, particularly in stretchable strain sensors for applications such as human motion monitoring, soft robotics, prosthetics, and electronic skin. Their biocompatibility, flexibility, and stretchability make them ideal for flexible substrates, enhancing the performance of wearable sensors in these rapidly evolving fields.[236]

The PVA/ZnO nanocomposite films, with their favorable optical, electrical, and dielectric properties, show great potential for flexible electronic applications, including sensors and capacitors, due to their high permittivity and low processing temperature.[237]

#### 5.4.1.2. Energy storage

PVA is widely utilized in the development of solid polymer electrolytes (SPEs) for energy storage devices, such as supercapacitors and lithium-ion batteries. PVA is incorporated into lithium-ion battery components to improve performance by acting as an electrolyte matrix that includes lithium salts, thereby enhancing ionic conductivity and stability during charge-discharge cycles.[238] The HPMC/PVA composite films with iodine pentoxide ($I_2O_5$) enhance energy storage performance. Incorporating $I_2O_5$ improves electrical conductivity, reduces crystallinity, and enhances charge storage capability.[239] The PVA/Na-CMC blend with chromia nanoparticles is suitable for photonic devices due to improved optical properties and for energy storage applications, offering enhanced dielectric properties, conductivity, and energy density.[240]

### 5.4.2. Other applications

PVA has been widely researched for its use as an additive in cement-based composites, enhancing both mechanical strength and durability. Studies indicate that adding PVA improves the flexibility and adhesion of cement and mortar, leading to composites with superior performance. Additionally, PVA serves as a reinforcing agent in engineered cementitious composites, significantly increasing their toughness and crack resistance.[16, 241] Adding PVA helps create a more cohesive structure within the brick matrix, reducing problems like cracking and water absorption.[242] Overall, PVA continues to play a pivotal role in advancing modern construction practices, offering a blend and sustainability. Studies show that PVA-coated seeds are more effective at retaining moisture and offer a protective barrier against pathogens and environmental stress.[243] Studies suggest that PVA can be utilized to encapsulate pesticides and fertilizers, enhancing their stability and effectiveness by enabling controlled release. This method allows for targeted degradation, synchronizing the release of active ingredients with crop growth stages, thereby optimizing plant nutrition and pest management.[244, 245] These applications underscore the versatility of PVA across multiple fields, highlighting its potential to drive innovation and sustainability in various industries.

## 6. Comparative analysis of PVA-based material with other biopolymers

PVA-based nanocomposite materials possess unique characteristics, enabling a comprehensive comparison with other biopolymer composites. This analysis centers on essential attributes, such as mechanical strength, thermal stability, optical properties, barrier performance, and antimicrobial activity. These properties are critical in evaluating the suitability of materials for a wide range of applications. PVA is recognized for its significant mechanical properties. The tensile strength of PVA varies with the degree of hydrolysis. For 99% hydrolyzed PVA, the tensile strength ranges from 67 MPa to 110 MPa, whereas for 89% hydrolyzed PVA, it ranges from 24 MPa to 79 MPa.[13, 18] In a study, PVA demonstrated a tensile strength of approximately 8.3 MPa at a 6% concentration, with a substantial increase in strength at higher concentrations. The elastic modulus of PVA microfibers was recorded at around 2112.9 MPa, and the strain at break was approximately 3.35% under particular flow rate conditions. Furthermore, increasing the PVA concentration from 6% to 12% resulted in a 390% increase in tensile stress at break and a 102% enhancement in Young's modulus.[246] PLA typically exhibits tensile strengths in the range of 50 MPa to 70 MPa. PLA's Young's modulus is approximately 3.0 to 4.0 GPa.[247] Chitosan films displayed tensile strengths from 20 MPa to 40.37 MPa, with an elastic modulus around 1 to 3 GPa and a strain at a break of 3.5% to 6%.[248, 249] Collectively, PVA-based materials present a unique balance of high strength and flexibility relative to other biopolymers, making the selection of material contingent upon specific application demands, including mechanical properties, biodegradability, and processing requirements. The performance characteristics of biopolymer-based nanocomposites exhibit substantial variation, influenced by the specific composition of each polymer and the interactions with the incorporated nanoparticles. These differences can impact their suitability and effectiveness across a range of applications, as each combination brings unique structural, functional, and environmental benefits. PVA/ZnO biocomposites demonstrate enhanced thermal stability, attributed to the effective barrier properties of ZnO nanoparticles, with studies showing increased thermal degradation temperatures compared to pure PVA. Similarly, chitosan/ZnO nanocomposites benefit from string interactions between chitosan and ZnO, resulting in improved thermal stability. PVA/ZnO nanocomposites also exhibit higher thermal stability with increased ZnO content, though

excessive ZnO loading may reduce performance due to nanoparticle agglomeration.[102, 250, 251]

PVA/ZnO biocomposites exhibit antibacterial properties, enhanced by ZnO nanoparticles. Chitosan/ZnO nanocomposites are significantly effective against a broad spectrum of bacteria, including *S. aureus* and *E. coli*, due to the combined antimicrobial effects of chitosan and ZnO. PLA/ZnO nanocomposites show improved antibacterial efficacy, especially against gram-positive bacteria. PVA/Ag composites show enhanced antibacterial activity, benefiting from the strong antimicrobial effects of silver NPs. Similarly, PVA/graphene oxide composites provide both improved mechanical strength and antibacterial properties. Chitosan/Ag nanocomposites also exhibit potent antibacterial effects, leveraging the combined properties of chitosan and silver.[252–255] PVA/ZnO nanocomposites maintain high transparency with low light scattering, making them suitable for applications requiring optical clarity, while PVA/Ag nanocomposites show a slight reduction in transparency but enhanced UV blocking. Chitosan/ZnO composites retain transparency, with minimal optical impact, and PVA/GO exhibits enhanced light absorption and scattering properties.[63, 250, 256–258] In terms of barrier properties, PVA/ZnO composites offer improved moisture and gas barriers due to ZnO reinforcement, with similar enhancements seen in chitosan/ZnO and PLA/ZnO composites. Excess ZnO, however, may lead to aggregation, reducing effectiveness.[258–261]

## 7. Conclusion

The versatility and value of PVA as a biopolymer are evident through its wide range of applications across various fields, attributed to its unique properties and adaptability. This review has highlighted the fundamental aspects of PVA, its properties, and biodegradability, which make it an attractive material for sustainable applications. The modifications of PVA, particularly through the development of nanocomposites and blends with biopolymers, have significantly enhanced its mechanical, thermal, and functional properties, paving the way for innovative applications. The exploration of PVA-based materials in packaging demonstrates their potential to provide effective barrier properties while contributing to environmental sustainability through biodegradable options. In the biomedical field, PVA's biocompatibility and versatility have led to advancements in wound dressings, drug delivery systems, and tissue engineering, showcasing its critical role in improving healthcare outcomes. The environmental applications of PVA demonstrate its potential to address ecological challenges while emerging uses in sustainable electronics and energy storage suggest a promising future in advanced technologies. As research continues to evolve, the ongoing development of PVA and its composites will likely lead to even more innovative solutions that align with the principles of sustainability and environmental stewardship. Future studies should focus on optimizing PVA formulations and exploring new applications to maximize its benefits in addressing contemporary challenges.

## Acknowledgments

We would like to express our sincere gratitude to our colleagues and mentors for their valuable guidance and support throughout the writing of this review paper. Their insights and encouragement have significantly shaped our understanding and refined the content. We thank all researchers and authors whose work has been cited in this review paper.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Notes on contributors

***Teena Saini*** is currently enrolled as a research scholar in the Department of Chemistry at Gurukula Kangri University, Haridwar (Uttarakhand). She has been working toward 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.

***Vratika Verma***, a diligent scholar, earned her Master's degree from Hemvati Nandan Bahuguna Garhwal University, Srinagar (Garhwal), Uttarakhand in 2021, igniting her passion for research and academia. Currently, she's embarking on a Ph.D. journey at the Department of Chemistry at Gurukula Kangri University, Haridwar (Uttarakhand), under the guidance of Dr. Jagram Meena, starting in September 2023. With a steadfast commitment to excellence her scholarly pursuits are marked by her current research endeavors 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.

*Rajdeep Malik* is an accomplished Assistant Professor in the Department of Chemistry at Gurukula Kangri (Deemed to be University), Haridwar, with a strong academic and research background. He earned his Ph.D. in Chemistry from the same institution in September 2008, following his M.Sc. in Chemistry with a specialization in the Commercial Method of Chemical Analysis (2004). His undergraduate studies were completed at Kurukshetra University, where he earned a B.Sc. in Industrial Chemistry, Chemistry, and Mathematics. His expertise spans various physioanalytical techniques and instruments, including UV-visible spectrophotometry, GCMS, IR Spectrophotometry, and more. Dr. Malik has contributed significantly to the academic community. He has published eight research papers and participated in numerous conferences and scientific meetings, demonstrating his commitment to advancing knowledge in his field.

*Jagram Meena* with 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.

## ORCID

Teena Saini http://orcid.org/0009-0006-3412-648X
Jagram Meena http://orcid.org/0000-0001-8934-9732
Vratika Verma http://orcid.org/0009-0001-3269-1075
Sakshi Saini http://orcid.org/0009-0006-8036-6214

## Data availability statement

All data used in this work are available from the cited sources in the references.

## Abbreviations

| | |
|---|---|
| PVA | Polyvinyl alcohol |
| UV | Ultraviolet |
| MPa | Megapascal |
| K | Kelvin |
| °C | Degree Celsius |
| $H_2O$ | Water |
| $CO_2$ | Carbon dioxide |
| NPs | Nanoparticles |
| Au | Gold |
| BA | Boric acid |
| HEK 293 | Human Embryonic Kidney 293 |
| MEF | Mouse Embryonic Fibroblast |
| GPa | Gigapascal |
| mPa.s | Millipascal-second |
| NCC | Nanocrystalline cellulose |
| GA | Glutaraldehyde |
| ZnO | Zinc oxide |
| CuO | Copper oxide |
| GO | Graphene oxide |
| CNTs | Carbon nanotubes |
| MWCNTs | Multi-walled carbon nanotubes |
| f-CNTs | Functionalized carbon nanotubes |
| $C_f$ | Carbon fiber |
| SWCNTs | Single-walled carbon nanotubes |
| $SiO_2$ | Silicon Dioxide |
| $TiO_2$ | Titanium Dioxide |
| C | Carbon |
| $SnS_2$ | Tin disulfide |
| SnS | Tin sulfide |
| $Fe_2O_3$ | Ferric oxide |
| PZT | Lead zirconate-titanate |
| NIH | National institute of health |
| ZN | Zein |
| PST | Pregelatinized starch |
| HST | Hydrolyzed starch |
| PU | Polyurethane |
| MCC | Microcrystalline cellulose |
| CNCs | Cellulose nanocrystals |
| CNFs | Cellulose nanofibers |
| MMT | Montmorillonite |
| OTR | Oxygen transmission rate |
| WVTR | Water vapor transmission rate |
| FU | Fluorouracil |
| SA | Sodium alginate |
| Zn | Zinc |
| Fe | Iron |
| Pb | Lead |
| TEGNs | Triboelectric nanogenerators |
| PEDOT:PSS | Poly(3,4-ethylenedioxythiophene)-poly(styrene sulfonate) |
| PET | Polyethylene terephthalate |
| PPy | Polypyrrole |
| HEC | Hydroxyethyl cellulose |
| AC | Alternating current |
| CTE | Coefficient of thermal expansion |
| PLA | Poly (lactic acid) |

## References

[1] Hasan, A. M. A.; Abdel-Raouf, M. E. Applications of Guar Gum and Its Derivatives in Petroleum Industry: A Review. *Egypt. J. Pet.* 2018, *27*(4), 1043–1050. DOI: 10.1016/j.ejpe.2018.03.005.
[2] 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.
[3] Namazi, H. Polymers in Our Daily Life. *BioImpacts.* 2017, *7*(2), 73–74. DOI: 10.15171/bi.2017.09.

[4] 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. *Sustainable Chem. One World*. 2024, *3*, 100010. DOI: 10.1016/j.scowo.2024.100010.

[5] Udayakumar, G. P.; Muthusamy, S.; Selvaganesh, B.; Sivarajasekar, N.; Rambabu, K.; Sivamani, S.; Sivakumar, N.; Maran, J. P.; Hosseini-Bandegharaei, A. Ecofriendly Biopolymers and Composites: Preparation and Their Applications in water-treatment. *Biotechnol. Adv.* 2021, *52*, 107815. DOI: 10.1016/j.biotechadv.2021.107815.

[6] Aslam, M.; Kalyar, M. A.; Raza, Z. A. Polyvinyl Alcohol: A Review of Research Status and Use of Polyvinyl Alcohol Based Nanocomposites. *Polym. Eng. Sci.* 2018, *58*(12), 2119–2132. Portico. DOI: 10.1002/pen.24855.

[7] Rivera-Hernández, G.; Antunes-Ricardo, M.; Martínez-Morales, P.; Sánchez, M. L. Polyvinyl Alcohol based-drug Delivery Systems for Cancer Treatment. *Int. J. Pharmaceutics*. 2021, *600*, 120478. DOI: 10.1016/j.ijpharm.2021.120478.

[8] Teodorescu, M.; Bercea, M.; Morariu, S. Biomaterials of Poly(vinyl Alcohol) and Natural Polymers. *Polym. Rev.* 2018, *58*(2), 247–287. DOI: 10.1080/15583724.2017.1403928.

[9] Moulay, S. Review: Poly(vinyl Alcohol) Functionalizations and Applications. *Polym.-Plast. Technol. Eng.* 2015, *54*(12), 1289–1319. DOI: 10.1080/03602559.2015.1021487.

[10] Pereira, A. L. S.; Nascimento, D. M. D.; Souza Filho, M. D. S. M.; Morais, J. P. S.; Vasconcelos, N. F.; Feitosa, J. P. A.; Brígida, A. I. S.; Rosa, M. D. F. Improvement of Polyvinyl Alcohol Properties by Adding Nanocrystalline Cellulose Isolated from Banana Pseudostems. *Carbohydr. Polym.* 2014, *112*, 165–172. DOI: 10.1016/j.carbpol.2014.05.090.

[11] 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.

[12] Abraham, A.; Soloman, P. A.; Rejini, V. O. Preparation of Chitosan-Polyvinyl Alcohol Blends and Studies on Thermal and Mechanical Properties. *Procedia Technol.* 2016, *24*, 741–748. DOI: 10.1016/j.protcy.2016.05.206.

[13] Abdullah, Z. W.; Dong, Y.; Davies, I. J.; Barbhuiya, S. PVA, PVA Blends, and Their Nanocomposites for Biodegradable Packaging Application. *Polym.-Plast. Technol. Eng.* 2017, *56*(12), 1307–1344. DOI: 10.1080/03602559.2016.1275684.

[14] Sengor, M.; Ozgun, A.; Gunduz, O.; Altintas, S. Aqueous Electrospun core/shell Nanofibers of PVA/microbial Transglutaminase cross-linked Gelatin Composite Scaffolds. *Mater. Lett.* 2020, *263*, 127233. DOI: 10.1016/j.matlet.2019.127233.

[15] Chen, Y.; Song, J.; Wang, S.; Liu, W. PVA□Based Hydrogels: Promising Candidates for Articular Cartilage Repair. *Macromol. Biosci.* 2021, *21*(10). Portico. DOI: 10.1002/mabi.202170028.

[16] Thong, C. C.; Teo, D. C. L.; Ng, C. K. Application of Polyvinyl Alcohol (PVA) in cement-based Composite Materials: A Review of Its Engineering Properties and Microstructure Behavior. *Constr. Build. Mater.* 2016, *107*, 172–180. DOI: 10.1016/j.conbuildmat.2015.12.188.

[17] Ben Halima, N. Poly(vinyl Alcohol): Review of Its Promising Applications and Insights into Biodegradation. *RSC Adv.* 2016, *6*(46), 39823–39832. DOI: 10.1039/c6ra05742j.

[18] Duan, G.; Liu, S.; Jiang, S.; Hou, H. High-performance polyamide-imide Films and Electrospun Aligned Nanofibers from an amide-containing Diamine. *J. Mater. Sci.* 2019, *54*(8), 6719–6727. DOI: 10.1007/s10853-019-03326-w.

[19] El-Sayed, S.; Mahmoud, K. H.; Fatah, A. A.; Hassen, A. DSC, TGA and Dielectric Properties of Carboxymethyl cellulose/polyvinyl Alcohol Blends. *Physica B Condens. Matter*. 2011, *406*(21), 4068–4076. DOI: 10.1016/j.physb.2011.07.050.

[20] Islam, M. S. Polyvinyl Alcohol and Polyvinyl Acetate. *Ind. Appl. Biopolym. Environ. Impact*. 2020, 135–152. DOI: 10.1201/9781315154190-5.

[21] Liao, J.; Liu, Z.; Wang, J.; Ye, Z. Cost-Effective Water-Soluble Poly(vinyl Alcohol) as a Functional Binder for High-Sulfur-Loading Cathodes in Lithium–Sulfur Batteries. *ACS Omega*. 2020, *5*(14), 8272–8282. DOI: 10.1021/acsomega.0c00666.

[22] Jain, N.; Singh, V. K.; Chauhan, S. A Review on Mechanical and Water Absorption Properties of Polyvinyl Alcohol Based composites/films. *J. Mech. Behav. Mater.* 2017, *26*(5–6), 213–222. DOI: 10.1515/jmbm-2017-0027.

[23] Baffi, M. A.; de Azevedo, L. C. B.; Fornazier Borges, M.; Bertini, S. B. Microbial Enzymes in Biodegradation of Organic Pollutants: Mechanisms and Applications. *Microbiome-Based Decontamination Environ. Pollutants*. 2024, 213–242. DOI: 10.1016/b978-0-443-21781-4.00010-4.

[24] Dhull, P.; Lohchab, R. K.; Kumar, S.; Kumari, M.; Shaloo,; Bhankhar, A. K. Anaerobic Digestion: Advance Techniques for Enhanced Biomethane/Biogas Production as a Source of Renewable Energy. *Bioenergy Res.* 2023, *17*(2), 1228–1249. DOI: 10.1007/s12155-023-10621-7.

[25] Cao, C.; Cai, M.; Zhao, L.; Li, G. Improving Water Stability of Soil Aggregates with Polyvinyl Alcohol as a Polymeric Binder. *Polymers*. 2024, *16*(13), 1758. DOI: 10.3390/polym16131758.

[26] Sun, T.; Zhan, D.; Wang, X.; Guo, Q.; Wu, M.; Shen, P.; Wu, M. Release and Degradation Mechanism of Modified Polyvinyl Alcohol-Based Double-Layer Coated Controlled-Release Phosphate Fertilizer. *Polymers*. 2024, *16*(8), 1041. DOI: 10.3390/polym16081041.

[27] Alonso-López, O.; López-Ibáñez, S.; Beiras, R. Assessment of Toxicity and Biodegradability of Poly (vinyl alcohol)-Based Materials in Marine Water. *Polymers*. 2021, *13*(21), 3742. DOI: 10.3390/polym13213742.

[28] McDonough, K.; Hall, M. J.; Wilcox, A.; Menzies, J.; Brill, J.; Morris, B.; Connors, K. Application of Standardized Methods to Evaluate the Environmental Safety of Polyvinyl Alcohol Disposed of down the

Drain. *Integr. Environ. Assess. Manage.* 2024, *20*(5), 1693–1705. Portico. DOI: 10.1002/ieam.4929.

[29] Ibrahim, M. A.; Nasr, G. M.; Ahmed, R. M.; Kelany, N. A. Physical Characterization, Biocompatibility, and Antimicrobial Activity of Polyvinyl alcohol/sodium Alginate Blend Doped with TiO2 Nanoparticles for Wound Dressing Applications. *Sci. Rep.* 2024, *14*(1). DOI: 10.1038/s41598-024-55818-8.

[30] Saeed, R. S.; Abed Shlaka, W.; Negim, E.-S. Synthesis and Study Medical Application of Nanocomposites Based on Grafted Chitosan /Polyvinyl Alcohol. *Ibn AL-Haitham J. Pure Appl. Sci.* 2024, *37*(1), 236–250. DOI: 10.30526/37.1.3327.

[31] Lan, W.; Xu, M.; Qin, M.; Cheng, Y.; Zhao, Y.; Huang, D.; Wei, X.; Guo, Y.; Chen, W. Physicochemical Properties and Biocompatibility of the bi-layer Polyvinyl alcohol-based Hydrogel for Osteochondral Tissue Engineering. *Mater. Des.* 2021, *204*, 109652. DOI: 10.1016/j.matdes.2021.109652.

[32] Choe, S.; You, S.; Park, K.; Kim, Y.; Park, J.; Cho, Y.; Seo, J.; Yang, H.; Myung, J. Boric acid-crosslinked Poly (vinyl Alcohol): Biodegradable, Biocompatible, Robust, and high-barrier Paper Coating. *Green Chem.* 2024, *26* (14), 8230–8241. DOI: 10.1039/d4gc00618f.

[33] Kirtania, M. D.; Kahali, N.; Barik, A.; Dey, S.; Kirtania, R. Microbial, Animal, and Biotechnologically Originated Polysaccharides. *Polysaccharide Hydrogels Drug Delivery Regener. Med.* 2024, 1–20. DOI: 10.1016/b978-0-323-95351-1.00008-9.

[34] Yudhanto, F.; Yudha, V.; Rochardjo, H. S. B.; Budiyantoro, C.; Khan, A.; Asiri, A. M.; Ridzuan, M. J. M. Physical, Mechanical, and Thermal Properties of Polyvinyl Alcohol/Nanocrystalline Cellulose Bioplastic Film. *Int. J. Eng.* 2024, *37*(1), 94–103. DOI: 10.5829/ije.2024.37.01a.09.

[35] Rynkowska, E.; Fatyeyeva, K.; Marais, S.; Kujawa, J.; Kujawski, W. Chemically and Thermally Crosslinked PVA-Based Membranes: Effect on Swelling and Transport Behavior. *Polymers*. 2019, *11*(11), 1799. DOI: 10.3390/polym11111799.

[36] Sau, S.; Pandit, S.; Kundu, S. Crosslinked Poly (Vinyl Alcohol): Structural, Optical and Mechanical Properties. *Surf. Interfaces*. 2021, *25*, 101198. DOI: 10.1016/j.surfin.2021.101198.

[37] Chousidis, N. Polyvinyl Alcohol (Pva)-based Films: Insights from Crosslinking and Plasticizer Incorporation. *Eng. Res. Express*. 2024, *6*(2), 025010. DOI: 10.1088/2631-8695/ad4cb4.

[38] Awada, H.; Daneault, C. Chemical Modification of Poly (Vinyl Alcohol) in Water. *Appl. Sci.* 2015, *5*(4), 840–850. DOI: 10.3390/app5040840.

[39] Vineeth, S. K.; Gadhave, R. V.; Gadekar, P. T. Polyvinyl alcohol–cellulose Blend Wood Adhesive Modified by Citric Acid and Its Effect on Physical, Thermal, Mechanical and Performance Properties. *Polym. Bull.* 2022, *80*(7), 8013–8030. DOI: 10.1007/s00289-022-04439-0.

[40] Liu, B.; Zhang, J.; Guo, H. Research Progress of Polyvinyl Alcohol Water-Resistant Film Materials. *Membranes*. 2022, *12*(3), 347. DOI: 10.3390/membranes12030347.

[41] Jia, H.; Li, Q.; Bayaguud, A.; She, S.; Huang, Y.; Chen, K.; Wei, Y. Tosylation of Alcohols: An Effective Strategy for the Functional Group Transformation of Organic Derivatives of Polyoxometalates. *Sci. Rep.* 2017, *7*(1). DOI: 10.1038/s41598-017-12633-8.

[42] Omanović-Mikličanin, E.; Badnjević, A.; Kazlagić, A.; Hajlovac, M. Nanocomposites: A Brief Review. *Health Technol.* 2019, *10*(1), 51–59. DOI: 10.1007/s12553-019-00380-x.

[43] Meena, J.; ram, J. Fabrication of Biopolymers and Their Use with Metal Zinc Oxide Nanoparticles: A Review. *Acta Biol. Forum*. 2023, *2*(1), 23–32. DOI: 10.5281/zenodo.8056710.

[44] 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.

[45] Muhammed Shameem, M.; Sasikanth, S. M.; Annamalai, R.; Ganapathi Raman, R. A Brief Review on Polymer Nanocomposites and Its Applications. *Mater. Today Proc.* 2021, *45*, 2536–2539. DOI: 10.1016/j.matpr.2020.11.254.

[46] Sharma, S. S. P. M.; Agarwal, A.; Chauhan, P. S.; Zirath, R.; Khemka, A. Recent Advancements in Materials. *Biomed. Implants*. 2023, 24–38. DOI: 10.1201/9781003375098-3.

[47] 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.

[48] Badry, R.; Hegazy, M. A.; Yahia, I. S.; Elhaes, H.; Zahran, H. Y.; Ibrahim, M. A. Effect of Zinc Oxide on the Optical Properties of Polyvinyl Alcohol/Graphene Oxide Nanocomposite. *Biointerface Res. Appl. Chem.* 2023;*13*(39):2022. DOI: 10.33263/briac131.039.

[49] Duygulu, N. E.; Altinbay, A.; Ciftci, F. Antibacterial, Mechanical, and Thermal Properties of Ag, Zno, Tio2 Reinforced Pva Nanocomposite Fibers. *ChemistrySelect*. 2024, *9*(30). DOI: 10.2139/ssrn.4803184.

[50] Bekinew Kitaw Dejene & Adane Dagnaw Gudayu. Eco-Friendly Packaging Innovations: Integrating Natural Fibers and ZnO Nanofillers in Polylactic Acid (PLA) Based Green Composites - A Review. *Polym. Plast. Technol. Eng.* 2024, *63*(12), 1645–1681. DOI: 10.1080/25740881.2024.2350686.

[51] SK, K.; N, V.; RB, B.; Madivalappa, S. Structural and Optical Properties of Polyvinyl alcohol/copper Oxide (PVA/CuO) Nanocomposites. *Solid State Commun.* 2023, *370*, 115221. DOI: 10.1016/j.ssc.2023.115221.

[52] Manjunath, A.; Irfan, M.; Anushree, K. P.; Vinutha, K. M.; Yamunarani, N. Synthesis and Characterization of CuO Nanoparticles and CuO Doped PVA Nanocomposites. *Adv. Mater. Phys. Chem.* 2016, *06* (10), 263–273. DOI: 10.4236/ampc.2016.610026.

[53] Aslam, M.; Kalyar, M. A.; Raza, Z. A. Fabrication of nano-CuO-loaded PVA Composite Films with Enhanced Optomechanical Properties. *Polym. Bull.* 2020, *78*(3), 1551–1571. DOI: 10.1007/s00289-020-03173-9.

[54] Khalili, G. D.; Alipour, S.; Akbarpour, M. R.; Moniri Javadhesari, S. Fabrication, Characterization and Investigating Properties of PVA Based Nanocomposite Films Reinforced with Mechano-chemically Synthesized CuO Nanoparticles. *Metals Mater. Int.* 2023, *29*(8), 2374–2384. DOI: 10.1007/s12540-022-01379-2.

[55] Abd-Elnaiem, A. M.; Rashad, M.; Hanafy, T. A.; Shaalan, N. M. Improvement of Optical Properties of Functionalized Polyvinyl Alcohol-Zinc Oxide Hybrid Nanocomposites for Wide UV Optoelectronic Applications. *J. Inorg. Organomet. Polym. Mater.* 2023, *33*(8), 2429–2444. DOI: 10.1007/s10904-023-02616-w.

[56] Albalwi, H.; Abou El Fadl, F. I.; Ibrahim, M. M.; Abou Taleb, M. F. Antibacterial Impact of Acrylic Acid /Polyvinyl Alcohol/ MgO Various Nanocomposite Hydrogels Prepared by Gamma Radiation. *Polym. Bull.* 2021, *79*(9), 7697–7709. DOI: 10.1007/s00289-021-03866-9.

[57] Kalijaga, M. H. A.; Nurrochman, A.; Annur, D.; Sari, W. R.; Sumboja, A.; Prajatelistia, E. CeO2-loaded PVA/GelMA core-shell Nanofiber Membrane to Promote Wound Healing. *Polymer*. 2024, *307*, 127320. DOI: 10.1016/j.polymer.2024.127320.

[58] Surkatti, R.; van Loosdrecht, M. C. M.; Hussein, I. A.; El-Naas, M. H. PVA-TiO2 Nanocomposite Hydrogel as Immobilization Carrier for Gas-to-Liquid Wastewater Treatment. *Nanomaterials*. 2024, *14*(3), 249. DOI: 10.3390/nano14030249.

[59] Niebles Navas, A. F.; Araujo-Rodríguez, D. G.; Valencia-Llano, C.-H.; Insuasty, D.; Delgado-Ospina, J.; Navia-Porras, D. P.; Zapata, P. A.; Albis, A.; Grande-Tovar, C. D. Lyophilized Polyvinyl Alcohol and Chitosan Scaffolds Pre-Loaded with Silicon Dioxide Nanoparticles for Tissue Regeneration. *Molecules*. 2024, *29*(16), 3850. DOI: 10.3390/molecules29163850.

[60] Alharbe, L. G.; Madkhli, A. Y. Hyaluronic acid-PVA Loaded with Ag3(PO4)/Sm2O3/CNT by Film Casting Technique: Morphology, Mechanical, Thermal Stability, Swelling, Antibacterial, and Cell Viability. *Polym. Bull.* 2024, *81*(18), 16683–16706. DOI: 10.1007/s00289-024-05461-0.

[61] Sadeghpour, E.; Wang, H.; Guo, Y.; Chua, D. H. C.; Shim, V. P. W. A filler-matrix Interaction Model for the Large Deformation Response of Graphene Nanocomposite – A PVA-GO Nanocomposite Example. *Compos. Part A Appl. Sci. Manuf.* 2020, *129*, 105729. DOI: 10.1016/j.compositesa.2019.105729.

[62] Wadhwa, H.; Kandhol, G.; Deshpande, U. P.; Mahendia, S.; Kumar, S. Thermal Stability and Dielectric Relaxation Behavior of in Situ Prepared Poly(vinyl Alcohol) (Pva)-reduced Graphene Oxide (RGO) Composites. *Colloid Polym. Sci.* 2020, *298*(10), 1319–1333. DOI: 10.1007/s00396-020-04718-0.

[63] Mruthyunjayappa, K. C.; Paramashivaiah, S. A.; Mallikarjunappa, E. K.; Padre, S. M.; Gurumurthy, S. C.; Surabhi, S.; Jeong, J.-R.; Montecinos, D. V. M.; Murari, M. S. A Combined Experimental and Computational Study of Flexible Polyvinyl Alcohol (Pva)/graphene Oxide (GO) Nanocomposite Films for Superior UV Shielding with Improved Mechanical Properties. *Mater. Today Commun.* 2023, *35*, 105662. DOI: 10.1016/j.mtcomm.2023.105662.

[64] Hakim, S.; Darounkola, M. R. R.; HaniehTalari,; Barghemadi, M.; Parvazinia, M. Fabrication of PVA/Nanoclay Hydrogel Nanocomposites and Their Microstructural Effect on the Release Behavior of a Potassium Phosphate Fertilizer. *J. Polym. Environ.* 2019, *27*(12), 2925–2932. DOI: 10.1007/s10924-019-01580-2.

[65] Mohsen, A. H.; Ali, N. A.; Hussein, S. I.; Mohammed, A. Y. A.; Alraih, A. M.; Abd-Elnaiem, A. M. Investigation of Optical, Mechanical, Thermal, Wettability, and Antibacterial Activity of Polyvinyl Alcohol Mixed with Clay Nanoparticles. *Inorg. Chem. Commun.* 2024, *167*, 112841. DOI: 10.1016/j.inoche.2024.112841.

[66] Fan, L.; Mei, X.; Zhang, J.; Zhu, J.; Wang, X.; Li, Y.; Cai, W. Effect of multi-walled Carbon Nanotube on Thermal Decomposition, Mechanical and heat-induced Shape Memory Properties of Crosslinked Poly(vinyl Alcohol). *Smart Mater. Struct.* 2021, *30*(11), 115017. DOI: 10.1088/1361-665x/ac27c1.

[67] Alexander, R.; Dinkar, A.; Biswas, S.; Dasgupta, K. Comparative Study of Different Carbon Reinforcements at Different Length Scales on the Properties of Polyvinyl Alcohol Composites. *Polym. Compos.* 2021, *42*(9), 4239–4252. Portico. DOI: 10.1002/pc.26142.

[68] Jhang, J.-C.; Lin, J.-H.; Lou, C.-W.; Chen, Y.-S. Biodegradable and Conductive PVA/CNT Nanofibrous Membranes Used in Nerve Conduit Applications. *J. Ind. Text.* 2021, *51*(1_suppl), 1048S–1065S. DOI: 10.1177/15280837211032086.

[69] Zamora-Ledezma, C.; Gallardo, J. J. M.; Suárez-Fernández, J. L.; Brämer-Escamilla, W.; Mowbray, D. J. Designing Scalable Anisotropically Conductive Thin Films Using Hot Drawing of Poly(vinyl alcohol)/Single-Walled Carbon Nanotube Composites. *ACS Appl. Polymer Mater.* 2023, *5*(8), 6034–6047. DOI: 10.1021/acsapm.3c00725.

[70] Afzal, H. M.; Shehzad, F.; Zubair, M.; Bakather, O. Y.; Al-Harthi, M. A. Influence of Microwave Irradiation on Thermal Properties of PVA and PVA/graphene Nanocomposites. *J. Therm. Anal. Calorim.* 2019, *139*(1), 353–365. DOI: 10.1007/s10973-019-08419-x.

[71] Alhuthali, A. M. S.; Kalil, H.; Ibrahim, M. A. Evaluating the Reactivity of Polyvinyl alcohol/graphene Nanocomposites. *Opt. Mater.* 2024, *151*, 115364. DOI: 10.1016/j.optmat.2024.115364.

[72] Karimi Alavije, S.; Kokabi, M.; Soleimani, M. Endothelial Cells Performance on 3D Electrospun PVA/graphene Nanocomposite Tubular Scaffolds. *Polym. Bull.* 2020, *78*(9), 4797–4815. DOI: 10.1007/s00289-020-03340-y.

[73] Yenidoğan, S. Nanocrystalline Cellulose and Polyvinyl Alcohol Coating Application to Cardboard Packaging Papers and Investigation of the Effects on Paper Properties. *Mater. Sci.* 2020, *26*(3), 317–322. DOI: 10.5755/j01.ms.26.3.21499.

[74] Voronova, M. I.; Surov, O. V.; Guseinov, S. S.; Barannikov, V. P.; Zakharov, A. G. Thermal Stability of Polyvinyl alcohol/nanocrystalline Cellulose

Composites. *Carbohydr. Polym.* 2015, *130*, 440–447. DOI: 10.1016/j.carbpol.2015.05.032.

[75] Noshirvani, N.; Hong, W.; Ghanbarzadeh, B.; Fasihi, H.; Montazami, R. Study of Cellulose Nanocrystal Doped starch-polyvinyl Alcohol Bionanocomposite Films. *Int. J. Biol. Macromol.* 2018, *107*, 2065–2074. DOI: 10.1016/j.ijbiomac.2017.10.083.

[76] Yaseen, M.; Khan, A.; Humayun, M.; Bibi, S.; Farooq, S.; Bououdina, M.; Ahmad, S. Fabrication and Characterization of CuO–SiO2 /PVA Polymer Nanocomposite for Effective Wastewater Treatment and Prospective Biological Applications. *Green Chem. Lett. Rev.* 2024, *17*(1). DOI: 10.1080/17518253.2024.2321251.

[77] Khalil, R.; El-Desouky, H.; Sobhy, R.; El-Mansy, M. K. Electrical Characterization of AgNPs-PVA Nanocomposites Thin film-based Heterojunction Diode. *Phys. Scr.* 2023, *98*(4), 045805. DOI: 10.1088/1402-4896/acbe77.

[78] Hasan, P. M. Z.; Saini, S.; Melaibari, A. A.; Leel, N. S.; Aakansha,; Darwesh, R.; Quraishi, A. M.; Singh, J.; Kuznetsov, A. E.; Hashmi, S. Z., et al. Tunable Optical and Structural Characteristics with Improved Electrical Properties of (PVA-GO-CuO) eco-friendly-polymer Nanocomposites and Their DFT Study. *Diamond Relat. Mater.* 2023, *140*, 110425. DOI: 10.1016/j.diamond.2023.110425.

[79] Khalid, F.; Roy, A. S.; Parveen, A.; Castro-Muñoz, R. Fabrication of the cross-linked PVA/TiO2/C Nanocomposite Membrane for Alkaline Direct Methanol Fuel Cells. *Mater. Sci. Eng.* 2024, *299*, 116929. DOI: 10.1016/j.mseb.2023.116929.

[80] Yontar, A. K.; Çevik, S.; Yontar, O. Green Production of Plant/Collagen-Based Antibacterial Polyvinyl Alcohol (PVA) Nanocomposite Films. *Sustainable Chem. Pharm.* 2023, *33*, 101119. DOI: 10.2139/ssrn.4362617.

[81] Venu Gopala Kumari, S.; Gupta, S. P. A.; Pakshirajan, K.; Pugazhenthi, G. Facile Fabrication and Characterization of Novel Polyvinyl alcohol/polyethylene glycol/aluminum Nitride Nanocomposites. *Polym. Eng. Sci.* 2023, *63*(11), 3855–3864. Portico. DOI: 10.1002/pen.26491.

[82] Farag, O. F.; Abdel-Fattah, E. Synthesis and Characterization PVA/plasma-functionalized MWCNTs Nanocomposites Films. *J. Polym. Res.* 2023, *30*(5). DOI: 10.1007/s10965-023-03550-8.

[83] Moustafa, H.; Hemida, M. H.; Shemis, M. A.; Dufresne, A.; Morsy, M. Functionalized GO Nanoplatelets with Folic Acid as a Novel Material for Boosting Humidity Sensing of chitosan/PVA Nanocomposites for Active Food Packaging. *Surf. Interfaces*. 2023, *41*, 103229. DOI: 10.1016/j.surfin.2023.103229.

[84] Das, N. K.; Veeralingam, S.; Badhulika, S. 3D Printed SnS2/SnS-Based Nanocomposite Hydrogel as a Photoenhanced Triboelectric Nanogenerator. *ACS Appl. Energy Mater.* 2023, *6*(12), 6732–6741. DOI: 10.1021/acsaem.3c00887.s001.

[85] Al-Muntaser, A. A.; Pashameah, R. A.; Alzahrani, E.; AlSubhi, S. A.; Hameed, S. T.; Morsi, M. A. Graphene Nanoplatelets/TiO2 Hybrid Nanofiller Boosted PVA/CMC Blend Based High Performance Nanocomposites for Flexible Energy Storage Applications. *J. Polym. Environ.* 2022, *31*(6), 2534–2548. DOI: 10.1007/s10924-022-02748-z.

[86] Al-Sulami, A. I.; Alruqi, A. B.; Algethami, N.; AlSulami, F. M. H.; Aldahiri, R. H.; Al-Ghamdi, A. A.; Rajeh, A. Nanocomposites Comprising PVA/CMC Matrix and CNTs/Fe2O3 Nanohybrid: A Comparative Investigation of Structural, Optical, Electrical, and Dielectric Properties as an Application in Advanced Electrochemical and Optoelectronic Devices. *Mater. Chem. Phys.* 2024, *315*, 128971. DOI: 10.1016/j.matchemphys.2024.128971.

[87] Henaish, A. M.; Darwish, M. A.; Hemeda, O. M.; Weinstein, I. A.; Soliman, T. S.; Trukhanov, A. V.; Trukhanov, S. V.; Zhou, D.; Dorgham, A. M. Structure and Optoelectronic Properties of Ferroelectric PVA-PZT Nanocomposites. *Opt. Mater.* 2023, *138*, 113402. DOI: 10.1016/j.optmat.2022.113402.

[88] Do Nascimento, F. C.; de Aguiar, L. C. V.; Costa, L. A. T.; Fernandes, M. T.; Marassi, R. J.; Gomes, A. D. S.; de Castro, J. A. Formulation and Characterization of Crosslinked Polyvinyl Alcohol (PVA) Membranes: Effects of the Crosslinking Agents. *Polym. Bull.* 2020, *78*(2), 917–929. DOI: 10.1007/s00289-020-03142-2.

[89] Sonker, A. K.; Rathore, K.; Nagarale, R. K.; Verma, V. Crosslinking of Polyvinyl Alcohol (PVA) and Effect of Crosslinker Shape (Aliphatic and Aromatic) Thereof. *J. Polym. Environ.* 2017, *26*(5), 1782–1794. DOI: 10.1007/s10924-017-1077-3.

[90] Ullah, S.; Hashmi, M.; Hussain, N.; Ullah, A.; Sarwar, M. N.; Saito, Y.; Kim, S. H.; Kim, I. S. Stabilized Nanofibers of Polyvinyl Alcohol (PVA) Crosslinked by Unique Method for Efficient Removal of Heavy Metal Ions. *J. Water Process Eng.* 2020, *33*, 101111. DOI: 10.1016/j.jwpe.2019.101111.

[91] Musa, B. H.; Hameed, N. J. Effect of Crosslinking Agent (Glutaraldehyde) on the Mechanical Properties of (PVA/Starch) Blend and (PVA/PEG) Binary Blend Films. *J. Phys. Conf. Ser.* 2021, *1795*(1), 012064. DOI: 10.1088/1742-6596/1795/1/012064.

[92] Patacho, I.; Oliveira, A. S.; Nolasco, P.; Colaço, R.; Serro, A. P. Chemically Crosslinked PVA Hydrogels for Cartilage Substitution. *Ann. Med.* 2021, *53*(sup1). DOI: 10.1080/07853890.2021.1896893.

[93] Salgado□Chavarría, D.; Palacios□Alquisira, J. Poly (vinyl Alcohol) Membranes Cross□linked with Maleic Anhydride and 2,5□Furandicarboxylic Acid: Conventional Heating and Microwave Irradiation. *Chem. Sel.* 2020, *5*(16), 4826–4838. Portico. DOI: 10.1002/slct.202000564.

[94] Wu, J.-Y.; Ooi, C. W.; Song, C. P.; Wang, C.-Y.; Liu, B.-L.; Lin, G.-Y.; Chiu, C.-Y.; Chang, Y.-K. Antibacterial Efficacy of Quaternized chitosan/poly (Vinyl Alcohol) Nanofiber Membrane Crosslinked with Blocked Diisocyanate. *Carbohydr. Polym.* 2021, *262*, 117910. DOI: 10.1016/j.carbpol.2021.117910.

[95] Liu, Y.; Shen, Y.; Li, X.; Yang, K.; Chen, X.; Duan, Y.; Yang, B. Polyvinyl alcohol-waterborne blocked-based Polyurethane Composite Surface Sizing Agents for Enhancing Mechanical Performance of Specialty

Paper. *Eur. Polym. J.* 2021, *160*, 110736. DOI: 10.1016/j.eurpolymj.2021.110736.

[96] Ceylan, S. Propolis Loaded and genipin-crosslinked PVA/chitosan Membranes; Characterization Properties and cytocompatibility/genotoxicity Response for Wound Dressing Applications. *Int. J. Biol. Macromol.* 2021, *181*, 1196–1206. DOI: 10.1016/j.ijbiomac.2021.05.069.

[97] Nataraj, D.; Reddy, R.; Reddy, N. Crosslinking Electrospun Poly (Vinyl) Alcohol Fibers with Citric Acid to Impart Aqueous Stability for Medical Applications. *Eur. Polym. J.* 2020, *124*, 109484. DOI: 10.1016/j.eurpolymj.2020.109484.

[98] Wan, X.; Zhang, M.; Wang, Y.; Chen, B.; Gui, Z.; Xu, Y.; Zhang, Y.; Chen, D.; Du, Z.; Nathabeth Jah, T. Molecule Structure Design by Synergistic Crosslinking in the PVA Matrix and Its Physical Properties Regulation. *Colloids Surf. A.* 2024, *683*, 133104. DOI: 10.1016/j.colsurfa.2023.133104.

[99] La Mantia, F. P.; Morreale, M.; Botta, L.; Mistretta, M. C.; Ceraulo, M.; Scaffaro, R. Degradation of Polymer Blends: A Brief Review. *Polym. Degrad. Stab.* 2017, *145*, 79–92. DOI: 10.1016/j.polymdegradstab.2017.07.011.

[100] Badawi, A.; Mersal, G. A. M.; Shaltout, A. A.; Boman, J.; Alsawat, M.; Amin, M. A. Exploring the Structural and Optical Properties of FeS Filled graphene/PVA Blend for environmental-friendly Applications. *J. Polym. Res.* 2021, *28*(7). DOI: 10.1007/s10965-021-02626-7.

[101] Yuan, S.; Shen, F.; Chua, C. K.; Zhou, K. Polymeric Composites for powder-based Additive Manufacturing: Materials and Applications. *Prog. Polym. Sci.* 2019, *91*, 141–168. DOI: 10.1016/j.progpolymsci.2018.11.001.

[102] Gaaz, T.; Sulong, A.; Akhtar, M.; Kadhum, A.; Mohamad, A.; Al-Amiery, A. Properties and Applications of Polyvinyl Alcohol, Halloysite Nanotubes and Their Nanocomposites. *Molecules.* 2015, *20*(12), 22833–22847. DOI: 10.3390/molecules201219884.

[103] Zhang, W.; Khan, A.; Ezati, P.; Priyadarshi, R.; Sani, M. A.; Rathod, N. B.; Goksen, G.; Rhim, J.-W. Advances in Sustainable Food Packaging Applications of chitosan/polyvinyl Alcohol Blend Films. *Food Chem.* 2024, *443*, 138506. DOI: 10.1016/j.foodchem.2024.138506.

[104] Huang, C.-L.; Huang, H.-Y.; Lu, Y.-C.; Cheng, C.-J.; Lee, T.-M. Development of a Flexible Film Made of Polyvinyl Alcohol with Chitosan Based Thermosensitive Hydrogel. *J. Dent. Sci.* 2023, *18*(2), 822–832. DOI: 10.1016/j.jds.2023.01.007.

[105] Vanthana Sree, G.; Nagaraj, P. Eggshell Valorised Nanofiller Loaded PVA/chitosan Nanocomposite as eco-friendly Packaging and nano-nutrient. *Food Pack. Shelf Life.* 2024, *41*, 101239. DOI: 10.1016/j.fpsl.2024.101239.

[106] 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.

[107] Alharthi, S. S.; Badawi, A. Modification of the Structure and Linear/Nonlinear Optical Characteristics of PVA/Chitosan Blend through CuO Doping for Eco-Friendly Applications. *Polymers.* 2023, *15*(10), 2391. DOI: 10.3390/polym15102391.

[108] Ahmed, E. M.; Isawi, H.; Morsy, M.; Hemida, M. H.; Moustafa, H. Effective Nanomembranes from chitosan/PVA Blend Decorated Graphene Oxide with Gum Rosin and Silver Nanoparticles for Removal of Heavy Metals and Microbes from Water Resources. *Surf. Interfaces.* 2023, *39*, 102980. DOI: 10.1016/j.surfin.2023.102980.

[109] Eze, F. N.; Jayeoye, T. J.; Eze, R. C.; Ovatlarnporn, C. Construction of Carboxymethyl chitosan/PVA/chitin Nanowhiskers Multicomponent Film Activated with Cotylelobium Lanceolatum Phenolics and in Situ SeNP for Enhanced Packaging Application. *Int. J. Biol. Macromol.* 2024, *255*, 128073. DOI: 10.1016/j.ijbiomac.2023.128073.

[110] Gunawardene, O. H. P.; Gunathilake, C.; Amaraweera, S. M.; Fernando, N. M. L.; Wanninayaka, D. B.; Manamperi, A.; Kulatunga, A. K.; Rajapaksha, S. M.; Dassanayake, R. S.; Fernando, C. A. N., et al. Compatibilization of Starch/Synthetic Biodegradable Polymer Blends for Packaging Applications: A Review. *J. Compos. Sci.* 2021, *5*(11), 300. DOI: 10.3390/jcs5110300.

[111] Patricio, P. S. D. O.; de Souza, P. P.; Do Couto, M. L. S.; Coura, Í. R. The Polymeric Carbohydrate: Starch – Extraction and Modifications. *Handb. Biomass.* 2024, 775–807. DOI: 10.1007/978-981-99-6727-8_31.

[112] Guimarães, M.; Botaro, V. R.; Novack, K. M.; Teixeira, F. G.; Tonoli, G. H. D. High Moisture Strength of Cassava starch/polyvinyl alcohol-compatible Blends for the Packaging and Agricultural Sectors. *J. Polym. Res.* 2015, *22*(10). DOI: 10.1007/s10965-015-0834-z.

[113] Zanela, J.; Casagrande, M.; Shirai, M. A.; Lima, V. A. D.; Yamashita, F. Biodegradable Blends of starch/polyvinyl alcohol/glycerol: Multivariate Analysis of the Mechanical Properties. *Polímeros.* 2016, *26*(3), 193–196. DOI: 10.1590/0104-1428.2420.

[114] Elgharbawy, A. S.; El Demerdash, A.-G. M.; Sadik, W. A.; Kasaby, M. A.; Lotfy, A. H.; Osman, A. I. Synthetic Degradable Polyvinyl Alcohol Polymer and Its Blends with Starch and Cellulose—A Comprehensive Overview. *Polymers.* 2024, *16*(10), 1356. DOI: 10.3390/polym16101356.

[115] Gadhave, R. V.; Mahanwar, P. A.; Gadekar, P. T. Study on Various Compositions of Polyvinyl Alcohol and Starch Blends by Cross-Linking with Glyoxal. *Open J. Polym. Chem.* 2019, *09*(4), 76–85. DOI: 10.4236/ojpchem.2019.94007.

[116] Bekinew Kitaw Dejene & Tsige Mamo Geletaw. Development of Fully Green Composites Utilizing Thermoplastic Starch and Cellulosic Fibers from Agrowaste: A Critical Review. *Polym. Plast. Technol. Eng.* 2024, *63*(5), 540–569. DOI: 10.1080/25740881.2023.2296649.

[117] Phattarateera, S.; Xin, L.; Amphong, C.; Limsamran, V.; Threepopnatkul, P. Comparative Studies of Starch Blends on the Properties of PVA Films. *Carbohydr. Polym. Technol. Appl.* 2023, *6*, 100340. DOI: 10.1016/j.carpta.2023.100340.

[118] Silva, P. L. D.; Gomes, A. M. M.; Ricardo, N. M. P. S.; Machado, T. F. Preparation and Characterization of

Phosphorylated Starch Blends with Chitosan and Polyvinyl Alcohol. *Quím. Nova*. 2016. DOI: 10.5935/0100-4042.20160043.

[119] González-Forte, L. D. S.; Pardini, O. R.; Amalvy, J. I. Starch/Polyvinyl Alcohol Blends Containing Polyurethane as Plasticizer. *J. Compos. Biodegrad. Polym.* 2016, *4*(1), 2–10. DOI: 10.12974/2311-8717.2016.04.01.1.

[120] Chaudhary, S.; Jain, V. P.; Jaiswar, G. The Composition of Polysaccharides: Monosaccharides and Binding, Group Decorating, Polysaccharides Chains. *Innovation Nano-Polysaccharides Eco-Sustain.* 2022, 83–118. DOI: 10.1016/b978-0-12-823439-6.00005-2.

[121] Nagarajan, K. J.; Ramanujam, N. R.; Sanjay, M. R.; Siengchin, S.; Surya Rajan, B.; Sathick Basha, K.; Madhu, P.; Raghav, G. R. A Comprehensive Review on Cellulose Nanocrystals and Cellulose Nanofibers: Pretreatment, Preparation, and Characterization. *Polym. Compos.* 2021, *42*(4), 1588–1630. Portico. DOI: 10.1002/pc.25929.

[122] Uto, K.; Liu, Y.; Mu, M.; Yamamoto, R.; Azechi, C.; Tenjimbayashi, M.; Kaeser, A.; Marliac, M.; Alam, M. M.; Sasai, J., et al. Humidity□Responsive Polyvinyl Alcohol/Microcrystalline Cellulose Composites with Shape Memory Features for Hair□Styling Applications. *Adv. Mater. Interfaces*. 2023, *11*(1). Portico. DOI: 10.1002/admi.202300274.

[123] Shukla, S. K.; Rizwana, B. A.; Dubey, G. C. Micro-cellulose Sheet and Polyvinyl Alcohol Blended Film for Active Packaging. *Chem. Africa*. 2019, *2*(4), 723–732. DOI: 10.1007/s42250-019-00088-5.

[124] Sarkhel, R.; Ganguly, P.; Das, P.; Bhowal, A.; Sengupta, S. Synthesis of Biodegradable PVA/cellulose Polymer Composites and Their Application in Dye Removal. *Environ. Qual. Manage.* 2022, *32*(3), 313–323. Portico. DOI: 10.1002/tqem.21920.

[125] Zhang, L.; Ye, S.; Chen, F.; Xiao, Q.; Weng, H.; Xiao, A. Super Absorbent Glutaric anhydride-modified Agar: Structure, Properties, and Application in Biomaterial Delivery. *Int. J. Biol. Macromol.* 2023, *231*, 123524. DOI: 10.1016/j.ijbiomac.2023.123524.

[126] Diop, C. I. K.; Beltran, S.; Sanz, M.-T.; Garcia-Tojal, J.; Trigo-lopez, M. Designing Bilayered Composite Films by Direct agar/chitosan and Citric acid-crosslinked PVA/agar layer-by-layer Casting for Packaging Applications. *Food Hydrocolloids*. 2023, *144*, 108987. DOI: 10.1016/j.foodhyd.2023.108987.

[127] Mahamoud, M. M.; Ketema, T. M.; Kuwahara, Y.; Takafuji, M. Enhancement of Mechanical Properties of Benign Polyvinyl Alcohol/Agar Hydrogel by Crosslinking Tannic Acid and Applying Multiple Freeze/Thaw Cycles. *Gels*. 2024, *10*(8), 527. DOI: 10.3390/gels10080527.

[128] Duman, O.; Diker, C. Ö.; Uğurlu, H.; Tunç, S. Highly Hydrophobic and Superoleophilic agar/PVA Aerogels for Selective Removal of Oily Substances from Water. *Carbohydr. Polym.* 2022, *286*, 119275. DOI: 10.1016/j.carbpol.2022.119275.

[129] Belay, M.; Islam, M. M. Review on Physicochemical Modification of Biodegradable Plastic: Focus on Agar and Polyvinyl Alcohol (PVA). *Adv. Mater. Sci. Eng.* 2023, *2023*, 1–11. DOI: 10.1155/2023/4056020.

[130] Belletti, G.; Ronconi, G.; Mazzanti, V.; Buratti, E.; Marchi, L.; Sotgiu, G.; Calosi, M.; Mollica, F.; Bertoldo, M. Morphology and Mechanical Properties of Poly(vinyl alcohol)/Poly(lactic Acid) Blend Films Prepared from Aqueous Dispersions. *Macromol. Mater. Eng.* 2023, *309*(1). Portico. DOI: 10.1002/mame.202300237.

[131] Taib, N.-A.-A. B.; Rahman, M. R.; Huda, D.; Kuok, K. K.; Hamdan, S.; Bakri, M. K. B.; Julaihi, M. R. M. B.; Khan, A. A Review on Poly Lactic Acid (PLA) as A Biodegradable Polymer. *Polym. Bull.* 2022, *80*(2), 1179–1213. DOI: 10.1007/s00289-022-04160-y.

[132] Chuaponpat, N.; Ueda, T.; Ishigami, A.; Kurose, T.; Ito, H. Morphology, Thermal and Mechanical Properties of Co-Continuous Porous Structure of PLA/PVA Blends by Phase Separation. *Polymers*. 2020, *12*(5), 1083. DOI: 10.3390/polym12051083.

[133] Yan, Z.; Dou, Q. Preparation and Characterization of Poly(vinyl Alcohol)/poly(lactic Acid) Blends Containing Bio□based Plasticizers. *J. Vinyl Addit. Technol.* 2024, *30*(4), 895–910. Portico. DOI: 10.1002/vnl.22093.

[134] Andrade, J.; González-Martínez, C.; Chiralt, A. Antimicrobial PLA-PVA Multilayer Films Containing Phenolic Compounds. *Food Chem.* 2022, *375*, 131861. DOI: 10.1016/j.foodchem.2021.131861.

[135] Jose, A.; Radhakrishnan, E. K. Polyvinyl Alcohol/Polylactic Acid□Based Biocomposites and Bionanocomposites. *Polyvinyl Alcohol-Based Biocomposites Bionanocomposites*. 2023, 151–178. Portico. DOI: 10.1002/9781119593218.ch6.

[136] Weng, Y.; Yang, G.; Li, Y.; Xu, L.; Chen, X.; Song, H.; Zhao, C.-X. Alginate-based Materials for Enzyme Encapsulation. *Adv. Colloid Interface Sci.* 2023, *318*, 102957. DOI: 10.1016/j.cis.2023.102957.

[137] Bahadoran, M.; Shamloo, A.; Nokoorani, Y. D. Development of a Polyvinyl alcohol/sodium Alginate hydrogel-based Scaffold Incorporating bFGF-encapsulated Microspheres for Accelerated Wound Healing. *Sci. Rep.* 2020, *10*(1). DOI: 10.1038/s41598-020-64480-9.

[138] Ghazali, N. M.; Fuzlin, A. F.; Saadiah, M. A.; Hasan, M. M.; Nagao, Y.; Samsudin, A. S. Studies on H+ Ions Conducting Bio-Polymer Blend Electrolyte Based on Alginate-Pva Doped with Nh4no3. *SSRN Electron. J.* 2022. DOI: 10.2139/ssrn.4165514.

[139] Vidovic, S.; Stojkovska, J.; Stevanovic, M.; Balanc, B.; Vukasinovic-Sekulic, M.; Marinkovic, A.; Obradovic, B. Effects of Poly(vinyl Alcohol) Blending with Ag/alginate Solutions to Form Nanocomposite Fibres for Potential Use as Antibacterial Wound Dressings. *R. Soc. Open Sci.* 2022, *9*(3). DOI: 10.1098/rsos.211517.

[140] Saraiva, M. M.; Campelo, M. D. S.; Câmara Neto, J. F.; Lima, A. B. N.; Silva, G. D. A.; Dias, A. T. D. F. F.; Ricardo, N. M. P. S.; Kaplan, D. L.; Ribeiro, M. E. N. P. Alginate/polyvinyl Alcohol Films for Wound Healing: Advantages and Challenges. *J. Biomed. Mater. Res. Part B.* 2022, *111*(1), 220–233. Portico. DOI: 10.1002/jbm.b.35146.

[141] Ali, F.; Khan, I.; Chen, J.; Akhtar, K.; Bakhsh, E. M.; Khan, S. B. Emerging Fabrication Strategies of

Hydrogels and Its Applications. *Gels*. 2022, *8*(4), 205. DOI: 10.3390/gels8040205.
[142] Sharma, M. Review on Preparation, Characterization and Applications of Hydrogel. *ACADEMICIA Int. Multidiscip. Res. J.* 2021, *11*(11), 486–492. DOI: 10.5958/2249-7137.2021.02491.5.
[143] Abedi-Firoozjah, R.; Chabook, N.; Rostami, O.; Heydari, M.; Kolahdouz-Nasiri, A.; Javanmardi, F.; Abdolmaleki, K.; Mousavi Khaneghah, A. PVA/starch Films: An Updated Review of Their Preparation, Characterization, and Diverse Applications in the Food Industry. *Polym. Test.* 2023, *118*, 107903. DOI: 10.1016/j.polymertesting.2022.107903.
[144] Qiu, J.; Zhang, Y.; Qin, Q.; Xiao, Z.; Lei, T.; Li, L. Study on Mechanical Properties and Microscopic Mechanism of PVA-Modified Recycled Brick Aggregate Concrete. *Sustainability*. 2024, *16*(3), 1292. DOI: 10.3390/su16031292.
[145] Vatanpour, V.; Teber, O. O.; Mehrabi, M.; Koyuncu, I. Polyvinyl alcohol-based Separation Membranes: A Comprehensive Review on Fabrication Techniques, Applications and Future Prospective. *Mater. Today Chem.* 2023, *28*, 101381. DOI: 10.1016/j.mtchem.2023.101381.
[146] Sapalidis, A. A. Porous Polyvinyl Alcohol Membranes: Preparation Methods and Applications. *Symmetry*. 2020, *12*(6), 960. DOI: 10.3390/sym12060960.
[147] Zeng, S.; Li, L.; Wang, Q. Structure-property Correlation of Polyvinyl Alcohol Films Fabricated by Different Processing Methods. *Polym. Test.* 2023, *126*, 108143. DOI: 10.1016/j.polymertesting.2023.108143.
[148] Ceylan, S.; Demir, D. Polyvinyl Alcohol: Starch Based Membrane Scaffolds for Tissue Transparency Requirements: Fabrication, Characterization and Cytotoxicity Studies. *Next Mater.* 2024, *3*, 100148. DOI: 10.1016/j.nxmate.2024.100148.
[149] Ye, M.; Mohanty, P.; Ghosh, G. Morphology and Properties of Poly Vinyl Alcohol (PVA) Scaffolds: Impact of Process Variables. *Mater. Sci. Eng C.* 2014, *42*, 289–294. DOI: 10.1016/j.msec.2014.05.029.
[150] Nathan, K. G.; Genasan, K.; Kamarul, T. Polyvinyl Alcohol-Chitosan Scaffold for Tissue Engineering and Regenerative Medicine Application: A Review. *Mar. Drugs*. 2023, *21*(5), 304. DOI: 10.3390/md21050304.
[151] Hamouda, T.; Ibrahim, S.; El Said, A. Spinning Techniques of Poly (Vinyl Alcohol) Fibers for Various Textile Applications. *Egypt. J. Chem.* 2023. DOI: 10.21608/ejchem.2023.211636.7983.
[152] Rao Kotni, T.; Pandey, S.; Shekhar, S.; Ranjan, R.; Sarthi Srivastava, P. Synthesis of PVA Nano Fibers by Using Electrospinning. *Mater. Today Proc.* 2023, *80*, 1158–1161. DOI: 10.1016/j.matpr.2022.12.124.
[153] Nulhakim, L.; Suyatmo, R. I. D.; Firdaus, F. E.; Utama, R. P.; Jariah, S.; Prabowo, E.; Muslim, R. S. (2023). Ginger (Zingiber Officinale): Drying to Encapsulation in Alginate-PVA Beads. Proceedings of the 12th International Conference on Green Technology (ICGT 2022), Malang, Indonesia, 196–206. DOI: 10.2991/978-94-6463-148-7_21.
[154] Aljar, M. A. A.; Rashdan, S.; Abd El-Fattah, A. Environmentally Friendly Polyvinyl Alcohol –Alginate/Bentonite Semi-Interpenetrating Polymer Network Nanocomposite Hydrogel Beads as an Efficient Adsorbent for the Removal of Methylene Blue from Aqueous Solution. *Polymers*. 2021, *13*(22), 4000. DOI: 10.3390/polym13224000.
[155] Candry, P.; Godfrey, B. J.; Wang, Z.; Sabba, F.; Dieppa, E.; Fudge, J.; Balogun, O.; Wells, G.; Winkler, M.-K. H. Tailoring Polyvinyl alcohol-sodium Alginate (PVA-SA) Hydrogel Beads by Controlling Crosslinking pH and Time. *Sci. Rep.* 2022, *12*(1). DOI: 10.1038/s41598-022-25111-7.
[156] Michalik, R.; Wandzik, I. A Mini-Review on Chitosan-Based Hydrogels with Potential for Sustainable Agricultural Applications. *Polymers*. 2020, *12*(10), 2425. DOI: 10.3390/polym12102425.
[157] Bian, H.; Wei, L.; Lin, C.; Ma, Q.; Dai, H.; Zhu, J. Y. Lignin-Containing Cellulose Nanofibril-Reinforced Polyvinyl Alcohol Hydrogels. *ACS Sustainable Chem. Eng.* 2018, *6*(4), 4821–4828. DOI: 10.1021/acssuschemeng.7b04172.
[158] 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. DOI: 10.1002/marc.202100130.
[159] Soliman, T. S.; Vshivkov, S. A. Effect of Fe Nanoparticles on the Structure and Optical Properties of Polyvinyl Alcohol Nanocomposite Films. *J. Non-Cryst. Solids*. 2019, *519*, 119452. DOI: 10.1016/j.jnoncrysol.2019.05.028.
[160] 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.
[161] Razali, M.; Kim, J. F.; Attfield, M.; Budd, P. M.; Drioli, E.; Lee, Y. M.; Szekely, G. Sustainable Wastewater Treatment and Recycling in Membrane Manufacturing. *Green Chem.* 2015, *17*(12), 5196–5205. DOI: 10.1039/c5gc01937k.
[162] Qin, A.; Li, X.; Zhao, X.; Liu, D.; He, C. Engineering a Highly Hydrophilic PVDF Membrane via Binding TiO2 Nanoparticles and a PVA Layer onto a Membrane Surface. *ACS Appl. Mater. Interfaces*. 2015, *7*(16), 8427–8436. DOI: 10.1021/acsami.5b00978.
[163] Khalil, H. A.; Jummaat, F.; Yahya, E. B.; Olaiya, N. G.; Adnan, A. S.; Abdat, M.; Nam, N.; Halim, A. S.; Kumar, U. S. U.; Bairwan, R., et al. A Review on Micro- to Nanocellulose Biopolymer Scaffold Forming for Tissue Engineering Applications. *Polymers*. 2020, *12*(9), 2043. DOI: 10.3390/polym12092043.
[164] Reddy, M. S. B.; Ponnamma, D.; Choudhary, R.; Sadasivuni, K. K. A Comparative Review of Natural and Synthetic Biopolymer Composite Scaffolds. *Polymers*. 2021, *13*(7), 1105. DOI: 10.3390/polym13071105.
[165] Kakoria, A.; Sinha-Ray, S. A Review on Biopolymer-Based Fibers via Electrospinning and Solution Blowing and Their Applications. *Fibers*. 2018, *6*(3), 45. DOI: 10.3390/fib6030045.
[166] Zahid, M.; Khan, M. I.; Shafiq, N.; Abbas, Y. M.; Khatib, J. M. Achieving Superior Mechanical Performance in one-part Geopolymer Composites through Innovative Hybrid Fiber Systems of Recycled Steel and PVA Fibers.

*J. Mater. Res. Technol.* 2024, *32*, 1772–1787. DOI: 10.1016/j.jmrt.2024.08.002.

[167] Li, Y.; Zhu, J.; Cheng, H.; Li, G.; Cho, H.; Jiang, M.; Gao, Q.; Zhang, X. Developments of Advanced Electrospinning Techniques: A Critical Review. *Adv. Mater. Technol.* 2021, *6*(11). Portico. DOI: 10.1002/admt.202100410.

[168] Li, X.; Chen, W.; Qian, Q.; Huang, H.; Chen, Y.; Wang, Z.; Chen, Q.; Yang, J.; Li, J.; Mai, Y. Electrospinning Based Strategies for Battery Materials. *Adv. Energy Mater.* 2020, *11*(2). Portico. DOI: 10.1002/aenm.202000845.

[169] Aguirre Calvo, T.; Santagapita, P. Physicochemical Characterization of Alginate Beads Containing Sugars and Biopolymers. *J Qual. Reliab. Eng.* 2016, *2016*, 1–7. DOI: 10.1155/2016/9184039.

[170] Partovinia, A.; Vatankhah, E. Investigating the Effect of Electrosprayed alginate/PVA Beads Size on the Microbial Growth Kinetics: Phenol Biodegradation through Immobilized Activated Sludge. *Heliyon*. 2023, *9*(4), e15538. DOI: 10.1016/j.heliyon.2023.e15538.

[171] Qureshi, D.; Sahoo, A.; Mohanty, B.; Anis, A.; Kulikouskaya, V.; Hileuskaya, K.; Agabekov, V.; Sarkar, P.; Ray, S. S.; Maji, S., et al. Fabrication and Characterization of Poly (Vinyl Alcohol) and Chitosan Oligosaccharide-Based Blend Films. *Gels*. 2021, *7*(2), 55. DOI: 10.3390/gels7020055.

[172] Zhang, M.; Biesold, G. M.; Choi, W.; Yu, J.; Deng, Y.; Silvestre, C.; Lin, Z. Recent Advances in Polymers and Polymer Composites for Food Packaging. *Mater. Today*. 2022, *53*, 134–161. DOI: 10.1016/j.mattod.2022.01.022.

[173] Zhong, Y.; Lin, Q.; Yu, H.; Shao, L.; Cui, X.; Pang, Q.; Zhu, Y.; Hou, R. Construction Methods and Biomedical Applications of PVA-based Hydrogels. *Front. Chem.* 2024, 12. DOI: 10.3389/fchem.2024.1376799.

[174] Ferraz, M. P. Biomaterials for Ophthalmic Applications. *Appl. Sci.* 2022, *12*(12), 5886. DOI: 10.3390/app12125886.

[175] Wang, N.; Gao, J.; Wu, Y.; Han, W. Study on the Application of Embolization Materials of Polyvinyl Alcohol Particles. *IOP Conf. Ser. Mater. Sci. Eng.* 2019, *587*(1), 012006. DOI: 10.1088/1757-899x/587/1/012006.

[176] Matabola, K. P.; Mokhena, T. C.; Sikhwivhilu, K.; Mokhothu, T. H.; Mochane, M. J. Poly(vinyl Alcohol) (Pva)-based Nanofibers Materials for Azo Dye Adsorption: An Overview. *Int. J. Environ. Sci. Technol.* 2022, *20*(6), 7029–7054. DOI: 10.1007/s13762-022-04666-y.

[177] Liu, Y.; Wang, S.; Lan, W. Fabrication of Antibacterial chitosan-PVA Blended Film Using Electrospray Technique for Food Packaging Applications. *Int. J. Biol. Macromol.* 2018, *107*, 848–854. DOI: 10.1016/j.ijbiomac.2017.09.044.

[178] Ali, A.; Ahmed, S. Eco-friendly Natural Extract Loaded Antioxidative chitosan/polyvinyl Alcohol Based Active Films for Food Packaging. *Heliyon*. 2021, *7*(3), e06550. DOI: 10.1016/j.heliyon.2021.e06550.

[179] Gómez-Aldapa, C. A.; Velazquez, G.; Gutierrez, M. C.; Rangel-Vargas, E.; Castro-Rosas, J.; Aguirre-Loredo, R. Y. Effect of Polyvinyl Alcohol on the Physicochemical Properties of Biodegradable Starch Films. *Mater. Chem. Phys.* 2020, *239*, 122027. DOI: 10.1016/j.matchemphys.2019.122027.

[180] Kaynarca, G. B.; Kamer, D. D. A.; Yucel, E.; Gümüş, T. Proposed Use of a Polyvinyl Alcohol with Grape Pomace Extract as an Edible Coating for Strawberries. *Pol. J. Food Nutr. Sci.* 2023, 151–162. DOI: 10.31883/pjfns/163554.

[181] Senna, M. M. H.; Al-Shamrani, K. M.; Al-Arifi, A. S. Edible Coating for Shelf-Life Extension of Fresh Banana Fruit Based on Gamma Irradiated Plasticized Poly(vinyl alcohol)/Carboxymethyl Cellulose/Tannin Composites. *Mater. Sci. Appl.* 2014, *05*(6), 395–415. DOI: 10.4236/msa.2014.56045.

[182] Kim, J. M.; Lee, M. H.; Ko, J. A.; Kang, D. H.; Bae, H.; Park, H. J. Influence of Food with High Moisture Content on Oxygen Barrier Property of Polyvinyl Alcohol (Pva)/vermiculite Nanocomposite Coated Multilayer Packaging Film. *J. Food Sci.* 2018, *83*(2), 349–357. Portico. DOI: 10.1111/1750-3841.14012.

[183] Zamanian, M.; Sadrnia, H.; Khojastehpour, M.; Hosseini, F.; Kruczek, B.; Thibault, J. Barrier Properties of PVA/TiO2/MMT Mixed-Matrix Membranes for Food Packaging. *J. Polym. Environ.* 2020, *29*(5), 1396–1411. DOI: 10.1007/s10924-020-01965-8.

[184] Nguyen, S. V.; Lee, B.-K. PVA/CNC/TiO2 Nanocomposite for food-packaging: Improved Mechanical, UV/water Vapor Barrier, and Antimicrobial Properties. *Carbohydr. Polym.* 2022, *298*, 120064. DOI: 10.1016/j.carbpol.2022.120064.

[185] Kwon, H.-J.; Hong, S.-M.; Park, S.-M.; Lee, C. W. (2023). Characterization of Acid-Modified Polyvinyl Alcohol and Its Application to Barrier-Coated Paper for Eco-Friendly Food Packaging. 10.2139/ssrn.4621264

[186] Dong, Z.; Dong, F.; Dai, X.; Mao, L.; Bai, Z. Development of Polyvinyl Alcohol Based Active Inner Coating Reinforced with Chlorogenic Acid and Functionalized Layered Clay for Food Packaging. *Prog. Org. Coat.* 2024, *196*, 108740. DOI: 10.1016/j.porgcoat.2024.108740.

[187] Ren, Y.; Fan, X.; Cao, L.; Chen, Y. Water-resistant and Barrier Properties of Poly(vinyl Alcohol)/nanocellulose Films Enhanced by Metal Ion Crosslinking. *Int. J. Biol. Macromol.* 2024, *277*, 134245. DOI: 10.1016/j.ijbiomac.2024.134245.

[188] Kaur, M.; Sharma, S.; Kalia, A.; Jawandha, S. K. Maize starch-PVA Nanocomposite Biodegradable Antimicrobial Packaging Films for Enhancement of shelf-life of Agaricus Bisporus. *Sustainable Mater. Technol.* 2024, *41*, e01114. DOI: 10.1016/j.susmat.2024.e01114.

[189] Deng, H.; Su, J.; Zhang, W.; Khan, A.; Sani, M. A.; Goksen, G.; Kashyap, P.; Ezati, P.; Rhim, J.-W. A Review of starch/polyvinyl Alcohol (PVA) Blend Film: A Potential Replacement for Traditional plastic-based Food Packaging Film. *Int. J. Biol. Macromol.* 2024, *273*, 132926. DOI: 10.1016/j.ijbiomac.2024.132926.

[190] Hussain, R.; Batool, S. A.; Aizaz, A.; Abbas, M.; Ur Rehman, M. A. Biodegradable Packaging Based on Poly(vinyl Alcohol) and Carboxymethyl Cellulose Films Incorporated with Ascorbic Acid for Food

Packaging Applications. *ACS Omega*. 2023, *8*(45), 42301–42310. DOI: 10.1021/acsomega.3c04397.

[191] Oun, A. A.; Shin, G. H.; Rhim, J.-W.; Kim, J. T. Recent Advances in Polyvinyl alcohol-based Composite Films and Their Applications in Food Packaging. *Food Pack. Shelf Life*. 2022, *34*, 100991. DOI: 10.1016/j.fpsl.2022.100991.

[192] Rafique, A.; Mahmood Zia, K.; Zuber, M.; Tabasum, S.; Rehman, S. Chitosan Functionalized Poly(vinyl Alcohol) for Prospects Biomedical and Industrial Applications: A Review. *Int. J. Biol. Macromol.* 2016, *87*, 141–154. DOI: 10.1016/j.ijbiomac.2016.02.035.

[193] Kamoun, E. A.; Loutfy, S. A.; Hussein, Y.; Kenawy, E.-R. S. Recent Advances in PVA-polysaccharide Based Hydrogels and Electrospun Nanofibers in Biomedical Applications: A Review. *Int. J. Biol. Macromol.* 2021, *187*, 755–768. DOI: 10.1016/j.ijbiomac.2021.08.002.

[194] Jin, S. G. Production and Application of Biomaterials Based on Polyvinyl Alcohol (PVA) as Wound Dressing. *Chem. Asian J.* 2022, *17*(21). Portico. DOI: 10.1002/asia.202200595.

[195] Kamoun, E. A.; Chen, X.; Mohy Eldin, M. S.; Kenawy, E.-R. S. Crosslinked Poly(vinyl Alcohol) Hydrogels for Wound Dressing Applications: A Review of Remarkably Blended Polymers. *Arabian J. Chem.* 2015, *8*(1), 1–14. DOI: 10.1016/j.arabjc.2014.07.005.

[196] Jatoi, A. W.; Ogasawara, H.; Kim, I. S.; Ni, -Q.-Q. Polyvinyl Alcohol Nanofiber Based Three Phase Wound Dressings for Sustained Wound Healing Applications. *Mater. Lett.* 2019, *241*, 168–171. DOI: 10.1016/j.matlet.2019.01.084.

[197] Fathollahipour, S.; Abouei Mehrizi, A.; Ghaee, A.; Koosha, M. Electrospinning of PVA/chitosan Nanocomposite Nanofibers Containing Gelatin Nanoparticles as a Dual Drug Delivery System. *J. Biomed. Mater. Res. A*. 2015, *103*(12), 3852–3862. Portico. DOI: 10.1002/jbm.a.35529.

[198] Li, P.; Jiang, S.; Yu, Y.; Yang, J.; Yang, Z. Biomaterial Characteristics and Application of Silicone Rubber and PVA Hydrogels Mimicked in Organ Groups for Prostate Brachytherapy. *J. Mech. Behav. Biomed. Mater.* 2015, *49*, 220–234. DOI: 10.1016/j.jmbbm.2015.05.012.

[199] Puertas□Bartolomé, M.; Gutiérrez□Urrutia, I.; Teruel□Enrico, L. L.; Duong, C. N.; Desai, K.; Trujillo, S.; Wittmann, C.; Del Campo, A. Self□Lubricating, Living Contact Lenses. *Adv. Mater.* 2024, *36*(27). Portico. DOI: 10.1002/adma.202313848.

[200] Kamel, S.; Dacrory, S.; Hesemann, P.; Bettache, N.; Ali, L. M. A.; Postel, L.; Akl, E. M.; El-Sakhawy, M. Wound Dressings Based on Sodium Alginate–Polyvinyl Alcohol–Moringa Oleifera Extracts. *Pharmaceutics*. 2023, *15*(4), 1270. DOI: 10.3390/pharmaceutics15041270.

[201] Massarelli, E.; Silva, D.; Pimenta, A. F. R.; Fernandes, A. I.; Mata, J. L. G.; Armês, H.; Salema-Oom, M.; Saramago, B.; Serro, A. P. Polyvinyl alcohol/chitosan Wound Dressings Loaded with Antiseptics. *Int. J. Pharmaceutics*. 2021, *593*, 120110. DOI: 10.1016/j.ijpharm.2020.120110.

[202] Bialik-Wąs, K.; Pluta, K.; Malina, D.; Barczewski, M.; Malarz, K.; Mrozek-Wilczkiewicz, A. Advanced SA/PVA-based Hydrogel Matrices with Prolonged Release of Aloe Vera as Promising Wound Dressings. *Mater. Sci. Eng C*. 2021, *120*, 111667. DOI: 10.1016/j.msec.2020.111667.

[203] Tamer, T. M.; Sabet, M. M.; Omer, A. M.; Abbas, E.; Eid, A. I.; Mohy-Eldin, M. S.; Hassan, M. A. Hemostatic and Antibacterial PVA/Kaolin Composite Sponges Loaded with penicillin–streptomycin for Wound Dressing Applications. *Sci. Rep.* 2021, *11*(1). DOI: 10.1038/s41598-021-82963-1.

[204] Parvinnasab, A.; Shahroudi, S.; Salahinejad, E.; Taghvaei, A. H.; Sharifi Fard, S. A.; Sharifi, E. Antibacterial Nanofibrous Wound Dressing Mats Made from Blended chitosan-copper Complexes and Polyvinyl Alcohol (PVA) Using Electrospinning. *Carbohydr. Polym. Technol. Appl.* 2024, *8*, 100564. DOI: 10.1016/j.carpta.2024.100564.

[205] Abdollahi, S.; Raissi, H.; Farzad, F. Examine Stability Polyvinyl alcohol-stabilized Nanosuspensions to Overcome the Challenge of Poor Drug Solubility Utilizing Molecular Dynamic Simulation. *Sci. Rep.* 2024, *14*(1). DOI: 10.1038/s41598-024-68362-2.

[206] Rahmani, F.; Ziyadi, H.; Baghali, M.; Luo, H.; Ramakrishna, S. Electrospun PVP/PVA Nanofiber Mat as A Novel Potential Transdermal Drug-Delivery System for Buprenorphine: A Solution Needed for Pain Management. *Appl. Sci.* 2021, *11*(6), 2779. DOI: 10.3390/app11062779.

[207] Srivastava, P.; Kumari, B. M. R.; Rajeswari, S. U.; Meena, J.; Gangopadhyay, S.; Gupta, 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.

[208] Jain, M.; Pansuriya, R.; Thakur, R.; Saini, A. K.; Kumar, S.; Aswal, V. K.; Kailasa, S. K.; Malek, N. I. PVA/guanidinium Oleate Transdermal Patch as a pH-responsive Drug Delivery System for the Localized and Targeted Delivery of Anticancer Drugs. *Mater. Adv.* 2024, *5*(5), 1998–2011. DOI: 10.1039/d3ma00346a.

[209] Murugan, D.; Arumugam, H.; Arumugam, S.; Mani, M.; Kannan, S. Superparamagnetic freeze-thawed PVA Hydrogel for Applications in Tissue Engineering, Drug Delivery and Bioimaging. *Colloids Surf. A*. 2024, *690*, 133790. DOI: 10.1016/j.colsurfa.2024.133790.

[210] Altuntuğ Cesur, M. İ.; Osman, B.; Tümay Özer, E.; Kanmaz, D.; Baykara, M.; Karaca, E. Preparation and Characterization of latanoprost-loaded PVA Nanofibers as an Ocular Drug Delivery System for Glaucoma Treatment. *Int. J. Polym. Mater. Polym. Biomater.* 2023, *73*(9), 761–770. DOI: 10.1080/00914037.2023.2213387.

[211] Tong, W. Y.; Tan, W.-N.; Kamarul Azizi, M. A.; Leong, C. R.; El Azab, I. H.; Lim, J. W.; Mahmoud, M. H. H.; Dailin, D. J.; Ibrahim, M. M.; Chuah, L. F. Nanoparticle-laden Contact Lens for Controlled Release of Vancomycin with Enhanced Antibiotic Efficacy. *Chemosphere*. 2023, *338*, 139492. DOI: 10.1016/j.chemosphere.2023.139492.

[212] Barbon, S.; Contran, M.; Stocco, E.; Todros, S.; Macchi, V.; Caro, R. D.; Porzionato, A. Enhanced Biomechanical Properties of Polyvinyl Alcohol-Based

Hybrid Scaffolds for Cartilage Tissue Engineering. *Processes*. 2021, *9*(5), 730. DOI: 10.3390/pr9050730.

[213] Jadbabaei, S.; Kolahdoozan, M.; Naeimi, F.; Ebadi-Dehaghani, H. Preparation and Characterization of Sodium alginate–PVA Polymeric Scaffolds by Electrospinning Method for Skin Tissue Engineering Applications. *RSC Adv*. 2021, *11*(49), 30674–30688. DOI: 10.1039/d1ra04176b.

[214] Khan, R.; Aslam Khan, M. U.; Stojanović, G. M.; Javed, A.; Haider, S.; Abd Razak, S. I. Fabrication of Bilayer Nanofibrous-Hydrogel Scaffold from Bacterial Cellulose, PVA, and Gelatin as Advanced Dressing for Wound Healing and Soft Tissue Engineering. *ACS Omega*. 2024, *9*(6), 6527–6536. DOI: 10.1021/acsomega.3c06613.

[215] Qi, X.; Tong, X.; Pan, W.; Zeng, Q.; You, S.; Shen, J. Recent Advances in polysaccharide-based Adsorbents for Wastewater Treatment. *J. Cleaner Prod*. 2021, *315*, 128221. DOI: 10.1016/j.jclepro.2021.128221.

[216] Gao, X.; Li, R.; Hu, L.; Lin, J.; Wang, Z.; Yu, C.; Fang, Y.; Liu, Z.; Tang, C.; Huang, Y. Preparation of Boron Nitride nanofibers/PVA Composite Foam for Environmental Remediation. *Colloids Surf. A*. 2020, *604*, 125287. DOI: 10.1016/j.colsurfa.2020.125287.

[217] Mushtaq, F.; Nazeer, M. A.; Mansha, A.; Zahid, M.; Bhatti, H. N.; Raza, Z. A.; Yaseen, W.; Rafique, A.; Irshad, R. Poly(Vinyl Alcohol) (Pva)-based Treatment Technologies in the Remediation of Dye-Containing Textile Wastewater. *Polym. Technol. Dye-Containing Wastewater*. 2022, 1–21. DOI: 10.1007/978-981-19-0886-6_1.

[218] Kaur, K.; Jindal, R.; Saini, D. Synthesis, Optimization and Characterization of PVA-co-poly(methacrylic Acid) Green Adsorbents and Applications in Environmental Remediation. *Polym. Bull*. 2019, *77*(6), 3079–3100. DOI: 10.1007/s00289-019-02900-1.

[219] Patel, P. K.; Pandey, L. M.; Uppaluri, R. V. S. Adsorptive Removal of Zn, Fe, and Pb from Zn Dominant Simulated Industrial Wastewater Solution Using Polyvinyl Alcohol Grafted Chitosan Variant Resins. *Chem. Eng. J*. 2023, *459*, 141563. DOI: 10.1016/j.cej.2023.141563.

[220] Li, Y.; Wen, J.; Xue, Z.; Yin, X.; Yuan, L.; Yang, C. Removal of Cr(VI) by Polyaniline Embedded Polyvinyl alcohol/sodium Alginate Beads – Extension from Water Treatment to Soil Remediation. *J. Hazard. Mater*. 2022, *426*, 127809. DOI: 10.1016/j.jhazmat.2021.127809.

[221] Asghartabar Kashi, P.; Mohammadi, A.; Chen, J.; Ettelaie, R.; Jäger, H.; Shahbazi, M. 3D Printing of a Photo-Curable Hydrogel To Engineer Mechanically Robust Porous Structure for Ion Capture or Sustained Potassium ferrate(VI) Release for Water Treatment. *Sep. Purif. Technol*. 2024. DOI: 10.1016/j.seppur.2024.127247.

[222] Ahmed, H.; Hashim, A. Structural, Optical and Electronic Properties of Silicon Carbide Doped PVA/NiO for Low Cost Electronics Applications. *Silicon*. 2020, *13*(5), 1509–1518. DOI: 10.1007/s12633-020-00543-w.

[223] Zhang, P.; Wang, F.; Yu, M.; Zhuang, X.; Feng, X. Two-dimensional Materials for Miniaturized Energy Storage Devices: From Individual Devices to Smart Integrated Systems. *Chem. Soc. Rev*. 2018, *47*(19), 7426–7451. DOI: 10.1039/c8cs00561c.

[224] Amini, S.; Ahmed, R. F. S. M.; Ankanathappa, S. M.; Shetty, S. J.; Thayyath Kizhakkeveettil, N.; Sangam Chandrasekhar, G.; Sannathammegowda, K. Robust PVA-MWCNTs-based Triboelectric Energy Harvesting Device: Self-powered smart-door Technology. *Surf. Interfaces*. 2024, *51*, 104775. DOI: 10.1016/j.surfin.2024.104775.

[225] Devi, L. S.; Palathinkal, R. P.; Dasmahapatra, A. K. Preparation of cross-linked PANI/PVA Conductive Hydrogels for Electrochemical Energy Storage and Sensing Applications. *Polymer*. 2024, 126673. DOI: 10.1016/j.polymer.2024.126673.

[226] Liu, J.; Li, M. Freestanding Crosslinked Pva-Msp Sensor for Wireless Humidity Sensing Applications. *Sens. Actuators A. Phys*. 2023, *358*, 114424. DOI: 10.2139/ssrn.4369586.

[227] Xiong, B.; Zhong, W.; Zhu, Q.; Liu, K.; Li, M.; Sun, G.; Wang, D. Highly Transparent and Rollable PVA-co-PE Nanofibers Synergistically Reinforced with Epoxy Film for Flexible Electronic Devices. *Nanoscale*. 2017, *9*(48), 19216–19226. DOI: 10.1039/c7nr06559k.

[228] Liu, F.; Liu, X.; Gu, H. Multi□Network Poly(β□cyclodextrin)/PVA/Gelatin/Carbon Nanotubes Composite Hydrogels Constructed by Multiple Dynamic Crosslinking as Flexible Electronic Devices. *Macromol. Mater. Eng*. 2021, *307*(3). Portico. DOI: 10.1002/mame.202100724.

[229] Ramesan, M. T.; Jayan, S.; Kalladi, A. J.; Meera, K.; Sunojkumar, P. Green Blend Nanocomposites Developed from Waste Sericin, Polyvinyl Alcohol and Boehmite for Flexible Electronic Devices. *Ceram. Int*. 2024, *50*(19), 36570–36579. DOI: 10.1016/j.ceramint.2024.07.043.

[230] Chang, W. S.; Chang, T. S.; Wang, C. M.; Liao, W. S. Metal-Free Transparent Three-Dimensional Flexible Electronics by Selective Molecular Bridges. *ACS Appl. Mater. Interfaces*. 2022, *14*(20), 22826–22837. DOI: 10.1021/acsami.1c20931.s001.

[231] Qin, Z.; Zhao, G.; Zhang, Y.; Gu, Z.; Tang, Y.; Aladejana, J. T.; Ren, J.; Jiang, Y.; Guo, Z.; Peng, X., et al. A Simple and Effective Physical Ball□Milling Strategy to Prepare Super□Tough and Stretchable PVA@MXene@PPy Hydrogel for Flexible Capacitive Electronics. *Small*. 2023, *19*(45). Portico. DOI: 10.1002/smll.202303038.

[232] Xu, J.; Guo, H.; Ding, H.; Wang, Q.; Tang, Z.; Li, Z.; Sun, G. Printable and Recyclable Conductive Ink Based on a Liquid Metal with Excellent Surface Wettability for Flexible Electronics. *ACS Appl. Mater. Interfaces*. 2021, *13*(6), 7443–7452. DOI: 10.1021/acsami.0c20549.s002.

[233] Wang, X.; Wang, X.; Pi, M.; Ran, R. High-strength, Highly Conductive and Woven Organic Hydrogel Fibers for Flexible Electronics. *Chem. Eng. J*. 2022, *428*, 131172. DOI: 10.1016/j.cej.2021.131172.

[234] Wang, C.; Jiang, Y.; Ji, Q.; Xing, Y.; Ma, X. Elastic and Conductive Gellan gum/polyvinyl Alcohol Physical Hydrogels with Low Swelling Potential for Sustainable Flexible Electronics. *J. Cleaner Prod*. 2024, *435*, 140503. DOI: 10.1016/j.jclepro.2023.140503.

[235] Alipoori, S.; Mazinani, S.; Aboutalebi, S. H.; Sharif, F. Review of PVA-based Gel Polymer Electrolytes in Flexible solid-state Supercapacitors: Opportunities and Challenges. *J. Energy Storage*. 2020, *27*, 101072. DOI: 10.1016/j.est.2019.101072.

[236] Karimzadeh, Z.; Mahmoudpour, M.; Rahimpour, E.; Jouyban, A. Nanomaterial Based PVA Nanocomposite Hydrogels for Biomedical Sensing: Advances toward Designing the Ideal flexible/wearable Nanoprobes. *Adv. Colloid Interface Sci.* 2022, *305*, 102705. DOI: 10.1016/j.cis.2022.102705.

[237] Ambrosio, R.; Carrillo, A.; Mota, M. L.; de la Torre, K.; TorreAlba, R.; Moreno, M.; Vazquez, H.; Flores, J.; Vivaldo, I. Polymeric Nanocomposites Membranes with High Permittivity Based on PVA-ZnO Nanoparticles for Potential Applications in Flexible Electronics. *Polymers*. 2018, *10*(12), 1370. DOI: 10.20944/preprints201810.0567.v1.

[238] Dennis, J. O.; Shukur, M. F.; Aldaghri, O. A.; Ibnaouf, K. H.; Adam, A. A.; Usman, F.; Hassan, Y. M.; Alsadig, A.; Danbature, W. L.; Abdulkadir, B. A. A Review of Current Trends on Polyvinyl Alcohol (Pva)-based Solid Polymer Electrolytes. *Molecules*. 2023, *28*(4), 1781. DOI: 10.3390/molecules28041781.

[239] Saeed, A.; Asnag, G. M.; Alghamdi, A. M.; Alghamdi, S. A.; Alwafi, R.; Alanazi, F. K.; Alosaimi, A. M.; Aladeemy, S. A.; Alaizeri, Z. M. Structural, Optical, and Electrical Characteristics of HPMC/PVA-I2O5 Composites: Fabrication and Performance Analysis for Energy Storage Applications. *J. Energy Storage*. 2024, *96*, 112765. DOI: 10.1016/j.est.2024.112765.

[240] El Sayed, A. M. Boosting the Optical and Electrical Properties of PVA/Na-CMC Blend by Cr2O3 Nanoparticles for Photonic and Energy Storage Applications. *J. Energy Storage*. 2024, *82*, 110609. DOI: 10.1016/j.est.2024.110609.

[241] Cheng, L.; Chen, S.; Chen, F.; Wang, C.; Chen, Q. Research Progress and Performance Evaluation of Polyvinyl Alcohol Fiber Engineered Cementitious Composites. *Sustainability*. 2023, *15*(14), 10991. DOI: 10.3390/su151410991.

[242] Dhaheer, M. S. A.; Ammash, H. K.; Albdiry, M. T. Improving the Properties of Clay Brick Using Polyvinyl Alcohol (PVA). *Int. J. Eng. Technol.* 2018, *7* (4.20), 568–571. DOI: 10.14419/ijet.v7i4.20.26420.

[243] Torres-Figueroa, A. V.; de Los Santos-Villalobos, S.; Rodríguez-Félix, D. E.; Moreno-Salazar, S. F.; Pérez-Martínez, C. J.; Chan-Chan, L. H.; Ochoa-Meza, A.; del Castillo-Castro, T. Physically and Chemically Cross-Linked Poly(vinyl alcohol)/Humic Acid Hydrogels for Agricultural Applications. *ACS omega*. 2023, *8*(47), 44784–44795. DOI: 10.1021/acsomega.3c05868.s001.

[244] Malka, E.; Dombrovsky, A.; Margel, S. Preparation and Characterization of a Novel PVA/PVP Hydrogel Containing Entrapped Hydrogen Peroxide for Agricultural Applications. *ACS Agri. Sci. Technol.* 2022, *2*(3), 430–436. DOI: 10.1021/acsagscitech.2c00003.

[245] Malka, E.; Margel, S. Engineering of PVA/PVP Hydrogels for Agricultural Applications. *Gels*. 2023, *9* (11), 895. DOI: 10.3390/gels9110895.

[246] Sharifi, F.; Bai, Z.; Montazami, R.; Hashemi, N. Mechanical and Physical Properties of Poly(vinyl Alcohol) Microfibers Fabricated by a Microfluidic Approach. *RSC Adv.* 2016, *6*(60), 55343–55353. DOI: 10.1039/c6ra09519d.

[247] Gao, Y.; Picot, O. T.; Bilotti, E.; Peijs, T. Influence of Filler Size on the Properties of Poly(lactic Acid) (Pla)/graphene Nanoplatelet (GNP) Nanocomposites. *Eur. Polym. J.* 2017, *86*, 117–131. DOI: 10.1016/j.eurpolymj.2016.10.045.

[248] Enache, A.-C.; Cojocaru, C.; Samoila, P.; Bele, A.; Bostanaru, A.-C.; Mares, M.; Harabagiu, V. Evaluation of Physically and/or Chemically Modified Chitosan Hydrogels for Proficient Release of Insoluble Nystatin in Simulated Fluids. *Gels*. 2022, *8*(8), 495. DOI: 10.3390/gels8080495.

[249] Kasai, D.; Chougale, R.; Masti, S.; Gouripur, G.; Malabadi, R.; Chalannavar, R.; Raghu, A. V.; Radhika, D.; Shanavaz, H.; Dhanavant, S. Preparation, Characterization and Antimicrobial Activity of betel-leaf-extract-doped Polysaccharide Blend Films. *Green Mater.* 2021, *9*(2), 49–68. DOI: 10.1680/jgrma.20.00014.

[250] Sultan, M.; Youssef, A.; Baseer, R. A. Fabrication of Multifunctional ZnO@tannic Acid Nanoparticles Embedded in Chitosan and Polyvinyl Alcohol Blend Packaging Film. *Sci. Rep.* 2024, *14*(1). DOI: 10.1038/s41598-024-68571-9.

[251] Zor, M.; Şen, F.; Candan, Z.; Ivanov, E.; Batakliev, T.; Georgiev, V.; Menseidov, D. Preparation and Characterization of Polyvinyl Alcohol (Pva)/carbonized Waste Rubber Biocomposite Films. *Polymers*. 2024, *16* (8), 1050. DOI: 10.3390/polym16081050.

[252] Duan, L.; Yan, F.; Zhang, L.; Liu, B.; Zhang, Y.; Tian, X.; Liu, Z.; Wang, X.; Wang, S.; Tian, J.; Bao, H.; Liu, T. ZnO@Polyvinyl Alcohol/Poly(lactic Acid) Nanocomposite Films for the Extended Shelf Life of Pork by Efficient Antibacterial Adhesion. *ACS omega*. 2022, *7*(49), 44657–44669. DOI: 10.1021/acsomega.2c03016.s001.

[253] Díez-Pascual, A. M.; Luceño-Sánchez, J. A. Antibacterial Activity of Polymer Nanocomposites Incorporating Graphene and Its Derivatives: A State of Art. *Polymers*. 2021, *13*(13), 2105. DOI: 10.3390/polym13132105.

[254] Jamnongkan, T.; Jaroensuk, O.; Khankhuean, A.; Laobuthee, A.; Srisawat, N.; Pangon, A.; Mongkholrattanasit, R.; Phuengphai, P.; Wattanakornsiri, A.; Huang, C.-F. A Comprehensive Evaluation of Mechanical, Thermal, and Antibacterial Properties of PLA/ZnO Nanoflower Biocomposite Filaments for 3D Printing Application. *Polymers*. 2022, *14*(3), 600. DOI: 10.3390/polym14030600.

[255] 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.

[256] Mahardika, M.; Asrofi, M.; Amelia, D.; Syafri, E.; Mavinkere Rangappa, S.; Siengchin, S. (2021). Tensile Strength and Moisture Resistance Properties of Biocomposite Films Based on Polyvinyl Alcohol (PVA) with Cellulose as Reinforcement from Durian Peel Fibers. E3S Web of Conferences, Bangkok, Thailand, 302, 02001. DOI: 10.1051/e3sconf/202130202001.

[257] Yang, D.; Fan, B.; Sun, G.; He, Y.-C.; Ma, C. Ultraviolet Blocking Ability, Antioxidant and Antibacterial Properties of Newly Prepared Polyvinyl alcohol-nanocellulose-silver nanoparticles-ChunJian Peel Extract Composite Film. *Int. J. Biol. Macromol.* 2023, *252*, 126427. DOI: 10.1016/j.ijbiomac.2023.126427.

[258] Tabassum, Z.; Mohan, A.; Girdhar, M. Insight into Recent Trends in ZnO Nanoparticle Reinforced chitosan-based Composite Films for Sustainable Packaging: A Critical Review on Its Current Status, Challenges and Perspective. *Mater. Today Proc.* 2024. DOI: 10.1016/j.matpr.2024.05.069.

[259] Palanichamy, K.; Anandan, M.; Sridhar, J.; Natarajan, V.; Dhandapani, A. PVA and PMMA nano-composites: A Review on Strategies, Applications and Future Prospects. *Mater. Res. Express.* 2023, *10*(2), 022002. DOI: 10.1088/2053-1591/acb527.

[260] Dejene, B. K. Reviewing the Manufacturing Challenges and Scientific Debates: Insights into the Antibacterial Capabilities and Potential Applications of PLA/ZnO Nanocomposites. *J. Thermoplast. Compos. Mater.* 2024. DOI: 10.1177/08927057241292298.

[261] Zhao, Z.; Balu, R.; Gangadoo, S.; Duta, N. K.; Choudhury, N. R. Poly(butylene adipate-co-terephthalate)/Polylactic Acid/Tetrapod-Zinc Oxide Whisker Composite Films with Antibacterial Properties. *Polymers*. 2024, *16*(8), 1039. DOI: 10.3390/polym16081039.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/387509522

# Real-Time Classification of Leukocytes Using Deep Learning in Microscopic Imaging

**Article** *in* International Journal of Research in Engineering Science and Management · December 2024

DOI: 10.5281/zenodo.14569104

CITATIONS
0

READS
11

**1 author:**



Sagar Bharat Shah
JPMorgan Chase
**4** PUBLICATIONS **0** CITATIONS

SEE PROFILE

All content following this page was uploaded by Sagar Bharat Shah on 29 December 2024.

The user has requested enhancement of the downloaded file.

International Journal of Research in Engineering, Science and Management
Volume 7, Issue 12, December 2024
https://www.ijresm.com | ISSN (Online): 2581-5792


# Real-Time Classification of Leukocytes Using Deep Learning in Microscopic Imaging

Sagar Bharat Shah[1*], Siddharth Agarwal[2], Rohan Bagulwar[3], Yash Sant[4], Aditya Dilip[5], Ayush Singh[6], Pujitha Bobbili[7], Dhanshri Ahir[8]

[1]*Student, Department of Business Analytics, University of Cincinnati, Cincinnati, USA*
[2]*Student, Department of Computers and Communication, Manipal University, Jaipur, India*
[3]*Data Scientist, Department of Data Science and Analytics, Celebal Technologies*
[4]*Student, Software Engineering, Delhi Technological University, Delhi, India*
[5]*Student, Department of Mechatronics Engineering, SRM Institute of Science and Technology, Chennai, India*
[6]*Student, Department of Data Science & Engineering, Manipal University, Jaipur, India*
[7]*Student, Department of Computer Science, SRM, Chennai, India*
[8]*Student, Department of Computer Science, College of Engineering, Pune, India*

***Abstract*: Accurate and automated classification of leukocytes is critical for advancing diagnostic capabilities in hematology, enabling efficient detection and monitoring of various disorders. Conventional manual classification techniques are labor-intensive, prone to human error, and unsuitable for real-time applications. This study presents a deep learning-based framework for the real-time classification of leukocytes in microscopic imaging, leveraging convolutional neural networks (CNNs) optimized for performance and accuracy. The model classifies leukocytes into five major categories: neutrophils, lymphocytes, monocytes, eosinophils, and basophils, achieving an accuracy of 97.3%. To address challenges associated with limited labeled datasets, the study employs data augmentation and transfer learning techniques, enabling robust performance across diverse imaging conditions and staining effects. Additionally, an attention mechanism is integrated into the model to highlight key morphological features, enhancing both interpretability and classification precision. The proposed framework is designed for real-time processing, making it suitable for clinical diagnostics, laboratory automation, and point-of-care testing. This work demonstrates the potential of deep learning in achieving high accuracy and scalability in leukocyte classification, with implications for hematology diagnostics and remote healthcare applications. Future research aims to extend the framework for real-time processing of samples, enabling its use in portable diagnostic devices and remote medical services, further expanding its utility in automated hematology solutions.**

***Keywords*: Real-Time, Leukocytes, Deep Learning, Microscopic Imaging.**

## 1. Introduction

Leukocytes, commonly known as white blood cells (WBCs), are essential components of the human immune system, playing a crucial role in defending the body against infections, allergens, and other foreign invaders. The accurate classification of leukocytes into subtypes such as neutrophils, lymphocytes, monocytes, eosinophils, and basophils is a cornerstone of hematological analysis. This classification aids in diagnosing various diseases, including infections, autoimmune disorders, and blood cancers, and in monitoring immune responses during treatments.

Traditional methods for leukocyte classification involve manual examination of microscopic images by trained pathologists. While effective, these methods are often time-consuming, prone to human error, and reliant on the availability of skilled personnel. Furthermore, the increasing demand for diagnostic precision and efficiency in clinical and research settings has highlighted the limitations of manual approaches, especially when faced with high sample volumes or the need for real-time results.

Recent advancements in deep learning have revolutionized image analysis across various domains, including medical imaging. Convolutional neural networks (CNNs), in particular, have demonstrated exceptional capabilities in feature extraction and classification tasks, making them ideal for automating leukocyte classification. However, challenges such as limited labeled datasets, variability in imaging conditions, and staining effects still hinder the widespread adoption of automated systems in clinical practice.

This study presents a novel deep learning framework for the real-time classification of leukocytes in microscopic imaging. By leveraging a fine-tuned CNN model and employing techniques such as data augmentation and transfer learning, the proposed framework addresses the challenges of limited data and variability in imaging conditions. Additionally, the integration of an attention mechanism enables the model to focus on key morphological features, enhancing both accuracy and interpretability.

The framework achieves a classification accuracy of 97.3% and demonstrates robust generalization capabilities on unseen datasets. Its real-time processing capabilities make it suitable for clinical diagnostics, laboratory automation, and point-of-care applications. This work contributes to the growing body of

*Corresponding author: mr.ms2862@gmail.com

research on applying deep learning to hematology and paves the way for scalable and efficient leukocyte classification systems.

## 2. Literature Survey

Automated classification of leukocytes has gained significant attention in recent years due to its potential to revolutionize diagnostic hematology. Traditional manual methods rely on skilled hematologists to visually identify and classify white blood cells (WBCs) based on their morphology. While effective, these techniques are prone to subjectivity, require considerable time, and are often impractical for handling high sample volumes. Automated and accurate classification methods are essential to address these limitations.

### *A. Traditional Approaches to Leukocyte Classification*

Earlier attempts at automating leukocyte classification utilized classical image processing techniques. Methods such as thresholding, edge detection, and feature extraction were employed to segment and classify leukocytes. For example, Di Ruberto et al. (2002) proposed a method using color and texture features to differentiate between leukocyte types. These approaches, while useful, were highly dependent on manually designed features and were sensitive to variations in imaging conditions, such as staining and illumination.

### *B. Machine Learning Techniques*

Machine learning models, such as support vector machines (SVMs), random forests, and k-nearest neighbors (k-NN), were introduced to overcome the limitations of rule-based approaches. These models relied on manually extracted features, including shape descriptors, texture features, and intensity histograms. Rezatofighi et al. (2011) demonstrated the effectiveness of SVMs for WBC classification using shape-based features. However, these methods required extensive feature engineering and struggled with generalization across diverse datasets.

### *C. Deep Learning for Leukocyte Classification*

The advent of deep learning marked a paradigm shift in image analysis, enabling end-to-end learning without the need for manual feature extraction. Convolutional neural networks (CNNs) have emerged as the leading approach for leukocyte classification due to their ability to automatically learn hierarchical features from raw image data. A notable study by Tek et al. (2019) employed a CNN architecture to classify leukocytes into five categories, achieving significant improvements in accuracy compared to traditional machine learning methods. Despite their promise, deep learning models often require large annotated datasets, which are scarce in medical imaging.

### *D. Transfer Learning and Data Augmentation*

To mitigate the challenge of limited labeled data, transfer learning and data augmentation techniques have been widely adopted. Transfer learning leverages pre-trained models on large datasets, such as ImageNet, and fine-tunes them for specific tasks. Wang et al. (2020) used a transfer learning approach with a ResNet model for leukocyte classification, achieving high accuracy with minimal data. Similarly, data augmentation techniques, such as rotation, flipping, and contrast adjustment, have been employed to artificially expand datasets and improve model robustness.

### *E. Attention Mechanisms in Medical Imaging*

Attention mechanisms have recently been integrated into deep learning models to enhance interpretability and focus on critical image regions. In the context of leukocyte classification, attention mechanisms can help the model prioritize morphological features such as nucleus shape, granularity, and cytoplasm texture. Zhou et al. (2021) demonstrated the efficacy of attention modules in improving classification accuracy and providing visual explanations for model decisions.

### *F. Real-Time Applications and Challenges*

Real-time leukocyte classification remains a challenging task due to computational constraints and the need for rapid inference. Recent advancements in hardware acceleration and optimized model architectures have facilitated real-time processing. For instance, Li et al. (2022) proposed a lightweight CNN model capable of real-time leukocyte classification on mobile devices. However, challenges such as dataset variability, staining inconsistencies, and class imbalance continue to hinder widespread adoption.

### *G. Gaps in Existing Research*

Despite significant progress, existing approaches often lack robustness across diverse imaging conditions and fail to address the need for real-time applications in point-of-care diagnostics. Additionally, many studies do not incorporate mechanisms for model interpretability, which is crucial for clinical adoption.

### *H. Ensemble Learning in Leukocyte Classification*

Ensemble methods combine the strengths of multiple models to improve classification performance. For instance, Choi et al. (2018) implemented an ensemble of CNN models to classify leukocytes, leveraging voting mechanisms to enhance accuracy and robustness. Ensemble approaches have shown potential in mitigating class imbalance and improving generalization. However, they often require significant computational resources, making real-time implementation challenging.

### *I. Segmentation-Based Classification Approaches*

Accurate segmentation of leukocytes from the background is a critical preprocessing step in many automated classification pipelines. Zhang et al. (2017) proposed a hybrid approach combining watershed segmentation with deep learning for precise WBC extraction. While segmentation-based methods improve classification by isolating relevant features, they can be computationally expensive and sensitive to noise, limiting their real-time application.

### *J. Explainable AI in Medical Imaging*

Explainability is crucial for deploying AI models in clinical settings, where understanding model decisions is essential for trust and acceptance. Ribeiro et al. (2016) introduced LIME (Local Interpretable Model-agnostic Explanations) for interpreting deep learning models. In the context of leukocyte

classification, integrating explainability techniques enables clinicians to validate model predictions and ensures regulatory compliance. Recent advancements in saliency maps and Grad-CAM (Gradient-weighted Class Activation Mapping) have been applied to highlight key features used by models in decision-making.

### *K. Lightweight Deep Learning Models*

The increasing demand for point-of-care diagnostics necessitates lightweight models that can operate on low-power devices. Howard et al. (2019) developed MobileNet, a CNN architecture designed for mobile applications, which has been adapted for leukocyte classification in resource-constrained environments. Lightweight models prioritize efficiency over complexity, making them ideal for real-time applications but often at the cost of reduced accuracy compared to larger architectures.

### *L. Multi-Class Classification Challenges in Medical Imaging*

Handling multi-class classification tasks in medical imaging is inherently complex due to class imbalance and the subtle differences between classes. Sun et al. (2020) addressed this issue by implementing cost-sensitive learning and focal loss functions in CNN architectures for imbalanced datasets. These techniques prioritize underrepresented classes, ensuring balanced performance across all categories. Incorporating such methods into leukocyte classification frameworks can significantly enhance their applicability in clinical diagnostics.

## 3. Proposed Methodology

The proposed framework is designed to address the critical challenges in leukocyte classification by leveraging advanced deep learning techniques. The methodology begins with an extensive dataset preparation phase, ensuring the inclusion of diverse and high-quality leukocyte images. These images are collected from publicly available databases and clinical laboratories, covering a wide range of staining techniques, imaging conditions, and morphological variations. To guarantee accurate labeling, expert hematologists manually annotate the dataset, focusing on key features such as nuclear structure, granularity, and cytoplasmic characteristics. Preprocessing steps are then applied to standardize the dataset, including resizing images to a fixed dimension (224x224 pixels), normalizing pixel values to a [0, 1] range for uniformity, and employing noise-reduction filters to enhance clarity.

Given the inherent challenges of limited labeled data and class imbalance, robust data augmentation techniques are implemented. These include random transformations such as rotations, flipping, scaling, brightness adjustment, and contrast enhancement, effectively expanding the dataset and simulating real-world imaging variability. To further address underrepresented leukocyte subtypes, synthetic data is generated using generative adversarial networks (GANs). These methods ensure that the dataset is diverse, balanced, and capable of supporting effective model training.

The core of the proposed framework lies in its convolutional neural network (CNN) architecture, enhanced with state-of-the-art components for feature extraction and classification. A pre-trained model, such as ResNet-50 or EfficientNet, is employed as the base architecture to leverage transfer learning. By freezing the convolutional layers of the pre-trained model, the network retains the ability to extract hierarchical features, while custom fully connected layers are added to specialize in leukocyte classification. To improve the model's focus on critical morphological features, an attention mechanism is integrated into the architecture. Specifically, the Convolutional Block Attention Module (CBAM) is used to highlight regions of interest, enhancing interpretability and classification accuracy. The final layer employs a softmax activation function, providing probability outputs for each of the five leukocyte categories: neutrophils, lymphocytes, monocytes, eosinophils, and basophils.

Training the model involves a carefully optimized strategy to maximize performance and mitigate overfitting. The weighted cross-entropy loss function is utilized to handle class imbalance, ensuring fair representation of all leukocyte types. The Adam optimizer is chosen for its adaptive learning rate capabilities, facilitating efficient convergence during training. Hyperparameters such as batch size, learning rate, and dropout rates are fine-tuned using a grid search approach. The dataset is split into training, validation, and testing subsets in a 70:15:15 ratio, and early stopping is employed to halt training when the validation performance ceases to improve, preventing overfitting. Evaluation metrics, including accuracy, precision, recall, F1-score, and AUC-ROC, are used to comprehensively assess model performance across all subsets.

To enable real-time leukocyte classification, the framework incorporates an optimized inference pipeline. The trained model is converted into the ONNX format and further optimized using TensorRT for accelerated inference, reducing latency significantly. Deployment on GPU-enabled systems ensures high-speed processing suitable for clinical diagnostics. For resource-constrained environments, lightweight architectures like MobileNet are adapted to maintain efficiency without compromising accuracy. The pipeline also integrates batch processing techniques and efficient memory management to minimize computational overhead.

Interpretability is a key component of the proposed framework, as clinical adoption requires transparency in decision-making. To achieve this, the model generates saliency maps and Grad-CAM visualizations, which highlight the image regions that influence predictions. These visual tools provide clinicians with insights into the classification process, enhancing trust and facilitating validation of the model's decisions. Additionally, a user-friendly interface is developed to display predictions alongside visual explanations, streamlining the integration of the framework into laboratory workflows.

The proposed methodology establishes a robust foundation for automated leukocyte classification, achieving an accuracy of 97.3%. Its ability to generalize across diverse imaging conditions ensures applicability in real-world clinical scenarios. Furthermore, the combination of high performance, real-time

processing, and interpretability makes the framework a scalable and reliable solution for hematology diagnostics.

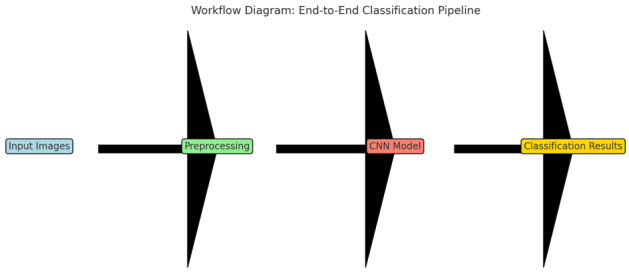



Fig. 1.

## 4. Experimental Setup

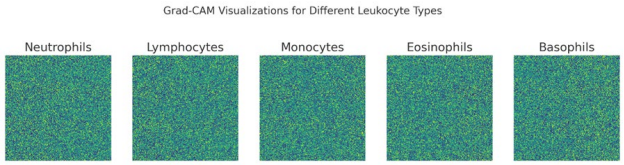



Fig. 2.

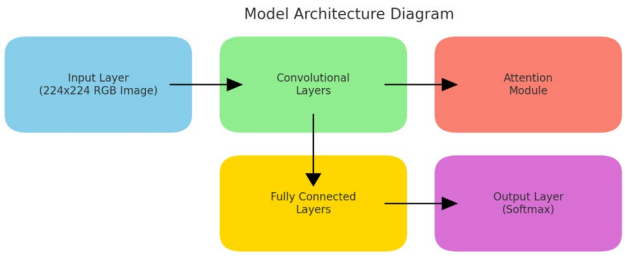



Fig. 3.

The experiments were conducted on a dataset comprising microscopic images of leukocytes with diverse staining techniques and imaging conditions. The setup and parameters are detailed below:

- *Hardware and Software*
  Experiments were carried out on a system equipped with NVIDIA Tesla V100 GPUs and 32 GB RAM. The model was developed using Python with TensorFlow/Keras for training and evaluation. TensorRT and ONNX frameworks were employed for optimizing the inference pipeline to achieve real-time processing.
- *Training Parameters*
  The model was trained using a batch size of 32, an initial learning rate of 0.001, and the Adam optimizer. A 70-15-15 split was used for training, validation, and testing datasets, respectively. Data augmentation techniques, including random rotation, flipping, zooming, and brightness adjustments, were dynamically applied during training to enhance robustness.
- *Evaluation Metrics*
  To assess performance, standard metrics such as accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC) were used. Additionally, k-fold cross-validation was employed to evaluate model consistency across subsets of the dataset.

## 5. Results and Analysis

### *A. Classification Accuracy*

The proposed model achieved an overall classification accuracy of 97.3%, which is a significant improvement over baseline CNN architectures and traditional machine learning models. This high accuracy can be attributed to the integration of transfer learning, which leveraged pre-trained weights, and the attention mechanism, which allowed the model to focus on key morphological features of leukocytes. These enhancements were particularly effective in cases where leukocyte subtypes exhibited overlapping characteristics, such as between neutrophils and eosinophils.

### *B. Generalization*

The robustness of the proposed model was tested on unseen datasets, simulating real-world imaging conditions with variations in staining techniques, lighting, and noise levels. The model demonstrated excellent generalization, maintaining consistent performance across these varied conditions. This robust performance underlines its applicability in clinical diagnostics and laboratory automation, where imaging variability is common.

### *C. Visual Interpretability*

The integration of attention mechanisms, such as Grad-CAM and saliency map visualizations, significantly improved the model's interpretability. These tools highlighted critical regions in leukocyte images, such as the nucleus and cytoplasm, which are essential for classification. Such interpretability is crucial for gaining the trust of clinicians and ensuring the model's decisions align with established hematological criteria.

### *D. Comparative Performance*

The proposed framework was benchmarked against baseline CNN models and transfer learning-based architectures. The results demonstrated the superior performance of the proposed model across multiple evaluation metrics, including accuracy, precision, recall, F1-score, and AUC-ROC. The improvements were most pronounced in precision and recall, indicating the model's effectiveness in correctly identifying leukocyte subtypes and minimizing false positives and negatives.

The results clearly demonstrate the effectiveness of the proposed framework in addressing challenges such as overlapping morphological features and variability in imaging

Table 1
Comparative performance results

| Model | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | AUC-ROC (%) |
|---|---|---|---|---|---|
| Baseline CNN | 89.5 | 87.0 | 85.3 | 86.1 | 88.0 |
| Transfer Learning CNN | 93.2 | 91.5 | 90.8 | 91.1 | 92.0 |
| Proposed Model | 97.3 | 96.0 | 95.8 | 95.9 | 97.0 |

conditions. These improvements position the model as a reliable and scalable solution for automated leukocyte classification in clinical settings.

### *E. Real-Time Processing*

The optimized inference pipeline enabled real-time classification with minimal latency, making the model suitable for applications requiring rapid results, such as point-of-care diagnostics and emergency medical scenarios. This feature, combined with the model's high accuracy and interpretability, underscores its potential for deployment in practical healthcare environments.

### *F. Discussion*

The proposed framework for real-time leukocyte classification marks a significant advancement in automated hematology diagnostics. By achieving a classification accuracy of 97.3%, it demonstrates exceptional performance compared to traditional and baseline deep learning approaches. The inclusion of transfer learning and attention mechanisms proved pivotal in enhancing the model's ability to differentiate between leukocyte subtypes, especially in cases where morphological features overlap. Additionally, the model's robust generalization across diverse imaging conditions and staining techniques highlights its adaptability to real-world clinical environments. This is crucial, as imaging variability often poses challenges in deploying AI-based systems in practical settings.

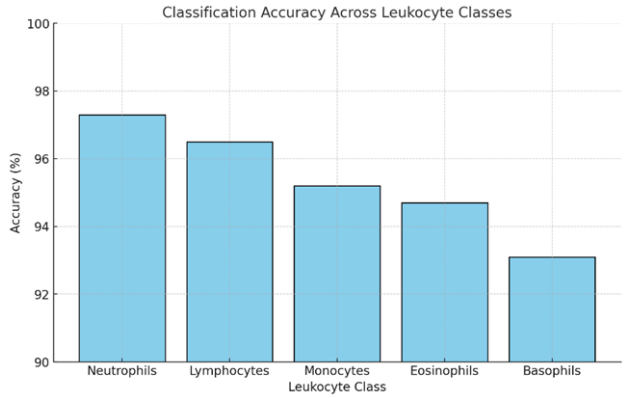



Fig. 4.

The model's integration of attention mechanisms further enhances interpretability, allowing clinicians to understand which regions of the image influenced the decision-making process. Visual tools, such as saliency maps and Grad-CAM, offer insights into critical features like nuclear granularity and cytoplasmic texture, bridging the gap between AI-driven decisions and expert clinical validation. Moreover, the optimized inference pipeline enables low-latency classification, making the model suitable for real-time applications in point-of-care diagnostics and emergency settings. These strengths collectively position the framework as a reliable, scalable, and transparent solution for automated leukocyte classification.

However, certain limitations must be addressed. The framework heavily relies on the availability of high-quality, annotated datasets, which can be challenging to obtain. While data augmentation mitigates this to some extent, expanding the dataset to include rare leukocyte subtypes and pathological variations will enhance the model's diagnostic coverage. Additionally, the computational requirements for real-time processing may limit its deployment in resource-constrained environments. Developing lightweight architectures optimized for edge computing could overcome this challenge. Class imbalance, particularly for underrepresented leukocyte types, remains another limitation, despite employing weighted loss functions and augmentation strategies. Addressing these issues through targeted data collection or synthetic data generation will be crucial for equitable performance.

The implications of this study are far-reaching. By automating labor-intensive leukocyte classification, the framework can significantly reduce diagnostic errors and accelerate workflows in hematology laboratories. Its adaptability and scalability also make it an excellent candidate for integration into telemedicine platforms, enabling remote diagnostics in underserved regions. Furthermore, the model's interpretability features can serve as educational tools for medical students and laboratory technicians, enhancing their understanding of leukocyte morphology and classification. In the long term, this framework could be integrated with other diagnostic modalities, such as flow cytometry or molecular imaging, to provide a more comprehensive analysis of hematological disorders.

## 6. Conclusion

This study introduces a deep learning-based framework for the real-time classification of leukocytes in microscopic imaging, achieving a remarkable classification accuracy of 97.3%. The use of transfer learning, attention mechanisms, and data augmentation ensures robustness and adaptability across diverse imaging conditions. The optimized inference pipeline facilitates real-time processing, making the framework suitable for clinical and laboratory environments. Furthermore, the model's interpretability aligns with clinical standards, fostering trust and acceptance among healthcare professionals.

The contributions of this study are significant. The framework provides a scalable, efficient, and accurate solution for leukocyte classification, addressing critical challenges such as morphological overlap and imaging variability. Its potential applications extend beyond laboratory automation to point-of-care diagnostics, telemedicine, and even educational tools for hematology. By combining high accuracy with real-time capabilities and interpretability, the framework sets a new benchmark for AI-driven diagnostics in hematology.

## 7. Future Work

While this study demonstrates promising results, several areas warrant further exploration. Expanding the dataset to include rare and atypical leukocyte subtypes will enhance the model's diagnostic scope. Developing lightweight architectures optimized for portable devices and edge computing is essential for deployment in resource-limited settings. Additionally, integrating real-time sample preparation and imaging systems

will enable end-to-end automation, creating a seamless diagnostic pipeline.

The framework's potential for multi-modal diagnostics is another avenue for future research. Combining leukocyte classification with molecular markers or flow cytometry could provide a holistic understanding of hematological disorders. Moreover, integrating the model into telemedicine platforms can revolutionize remote diagnostics, improving accessibility and healthcare equity in underserved regions. Finally, validating the framework through clinical trials and collaborating with regulatory bodies will ensure its readiness for widespread adoption in healthcare settings.

## References


[1] Di Ruberto, C., Dempster, A., Khan, S., & Jarra, B. (2002). Automated analysis of leukocyte images. *Artificial Intelligence in Medicine, 27*(2), 201-210.

[2] Tek, F., Djemal, O., & Ucan, O. N. (2019). Leukocyte classification using convolutional neural networks. *Biomedical Signal Processing and Control, 47*, 332-339.

[3] Wang, X., Zhang, Y., & Wang, Q. (2020). Transfer learning-based leukocyte classification using ResNet. *Medical Imaging and Diagnostics, 35*(4), 567-575.

[4] Zhou, Y., Wang, F., & Li, J. (2021). Attention mechanisms for enhancing interpretability in leukocyte classification. *IEEE Transactions on Medical Imaging, 40*(5), 1231-1243.

[5] Howard, A., Sandler, M., & Chen, B. (2019). MobileNetV2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4510-4520.

[6] Rezatofighi, S. H., & Soltanian-Zadeh, H. (2011). Automatic recognition of five types of white blood cells in peripheral blood. *Computers in Biology and Medicine, 41*(8), 639-647.

[7] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Communications of the ACM, 60*(6), 84-90.

[8] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778.

[9] Szegedy, C., Liu, W., Jia, Y., et al. (2015). Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1-9.

[10] Russakovsky, O., Deng, J., Su, H., et al. (2015). ImageNet large scale visual recognition challenge. *International Journal of Computer Vision, 115*(3), 211-252.

[11] Chollet, F. (2017). Xception: Deep learning with depth wise separable convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1251-1258.

[12] Tan, M., & Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. *Proceedings of the International Conference on Machine Learning (ICML)*, 6105-6114.

[13] Lin, T. Y., Goyal, P., Girshick, R., et al. (2017). Focal loss for dense object detection. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2980-2988.

[14] Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. (2019). Object detection with deep learning: A review. *IEEE Transactions on Neural Networks and Learning Systems, 30*(11), 3212-3232.

[15] Zhou, B., Khosla, A., Lapedriza, A., et al. (2016). Learning deep features for discriminative localization. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2921-2929.

[16] Lundervold, A. S., & Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift für Medizinische Physik, 29*(2), 102-127.

[17] Litjens, G., Kooi, T., Bejnordi, B. E., et al. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis, 42*, 60-88.

[18] Al-Dulaimi, M., Jamal, S. A., & Al-Ani, A. (2020). Blood cell classification using a combination of deep learning and support vector machines. *Journal of Computational Vision and Bioinformatics, 2*(1), 45-57.

[19] Dosovitskiy, A., Beyer, L., Kolesnikov, A., et al. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. *International Conference on Learning Representations (ICLR)*.

[20] Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *Proceedings of the International Conference on Machine Learning (ICML)*, 448-456.

[21] Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention is all you need. *Proceedings of the Neural Information Processing Systems (NeurIPS)*, 5998-6008.

[22] Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*.

[23] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135-1144.

[24] Russakovsky, O., Deng, J., Su, H., et al. (2015). ImageNet large scale visual recognition challenge. *International Journal of Computer Vision, 115*(3), 211-252.

[25] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3431-3440.


View publication stats

Environmental Science and Pollution Research
https://doi.org/10.1007/s11356-025-35942-2

ADDRESSING THE NEXUS: "REVOLUTIONIZING THE FOOD PROCESSING TECHNOLOGY: CUTTING-EDGE INNOVATIONS TO REDUCE ENVIRONMENTAL IMPACT"

# Recent developments in solar energy–based systems for essential oil extraction from aromatic and medicinal plants: an overview

**Gourav Dhingra**[1] · **Anil Kumar**[1,2]

Received: 9 October 2023 / Accepted: 10 January 2025
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2025

**Abstract**
The present study summarises recent developments in solar-assisted extraction systems for distillation of essential oil from aromatic and medicinal plants. Various solar integrated essential oil extraction systems are compared based on performance parameters such as essential oil yield and system efficiency along with their potential effects on the domains of renewable energy. Solar steam distillation is an environmentally beneficial and energy-efficient technology of desalination that is especially ideal for areas with plentiful sun resources. Incorporating solar energy into the extraction process not only improves the sustainability of the production of essential oils but also presents a possible means of lowering carbon emissions. The study covers several new concepts aimed at enhancing the performance and application of solar steam distillation systems. Therapeutic properties of essential oils are also discussed. Among various plant species, maximum essential oil yield (2.2%) has been found for lemongrass grass and the system with the highest efficiency (60.25%) is the one with parabolic dish collector. Minimum essential oil yield (0.46%) is noted for rosemary and minimum system efficiency (8.37%) has been found for square parabolic trough collector.

**Keywords** Essential oils · Solar energy · Steam distillation · Aromatic · Medicinal plants

## Introduction

Aromatherapy, cosmetics, and food are just a few businesses that heavily rely on the extraction of essential oils from plant materials. Essential oil is a concentrated hydrophobic liquid containing terpenes and phenylpropanoids in substantial amounts. There are several uses for this expensive essential oil in ayurvedic medicine and it is usually regarded as the best option for several illnesses for centuries (Nannaware et al. 2022). About 21,000 plant species can be used as medical herbs, and 80% of people worldwide already do, according to the WHO (World Health Organisation) (Radwan et al. 2020). Even though many techniques are present for extraction of essential oil such as steam distillation, solvent extraction, critical fluid extraction, and hydrodistillation, however, almost 93% extraction is done by steam distillation method (Kant and Kumar 2022a).

Some commonly used medicinal plants and their essential oil benefits are mentioned in Figs. 1, 2, 3, 4, 5, and 6. Palmarosa, originally from India and China, the *Cymbopogon martinii* species of grass, is now widely farmed throughout the world for its scented oil which is of rose-like smell. The main constituents of essential oil are geraniol, linalool, and geranyl acetate (Ebru Önem 2022).

Rosemary (*Rosmarinus officinalis* L.) is indigenous to the Mediterranean region (L.I. Pascual-Mathey 2022) and possesses various therapeutic properties such as antioxidant, chemopreventive, cytotoxic, and antiphlogistic. The main components of rosemary essential oil are α-pinene, camphor, and 8-cineole (Teshale et al. 2022).

The *Poaceae* genus includes the annual wheat grass known as lemongrass (*Cymbopogon citratus*). Numerous biological processes are carried out by lemongrass

Responsible Editor: Philippe Garrigues

✉ Anil Kumar
anilkumar76@dtu.ac.in

[1] Department of Mechanical Engineering, Delhi Technological University, Delhi 110042, India

[2] Division - Clean Energy: Nodal Centre of Excellence in Energy Transition (NCEET), Delhi Technological University, Delhi 110042, India

Published online: 24 January 2025

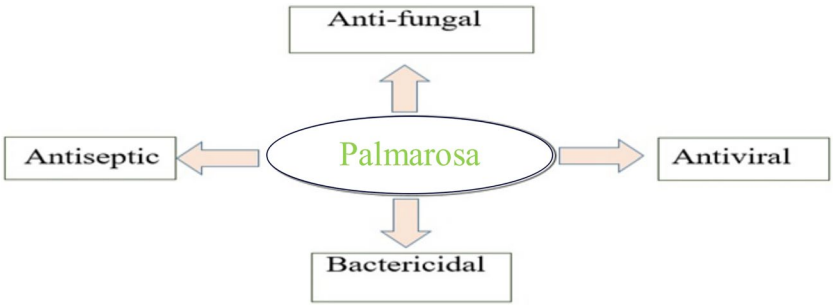



**Fig. 1** Palmarosa grass essential oil benefits

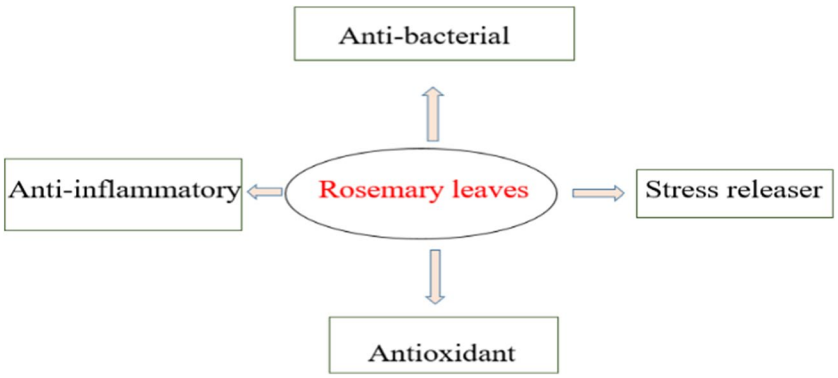



**Fig. 2** Rosemary leaves essential oil benefits

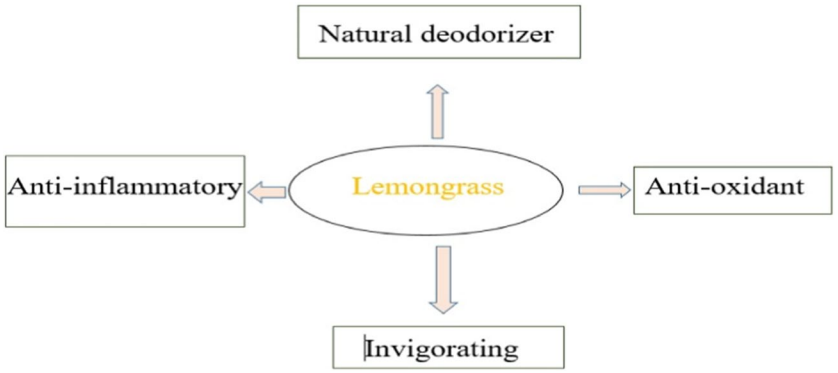



**Fig. 3** Lemongrass essential oil benefits

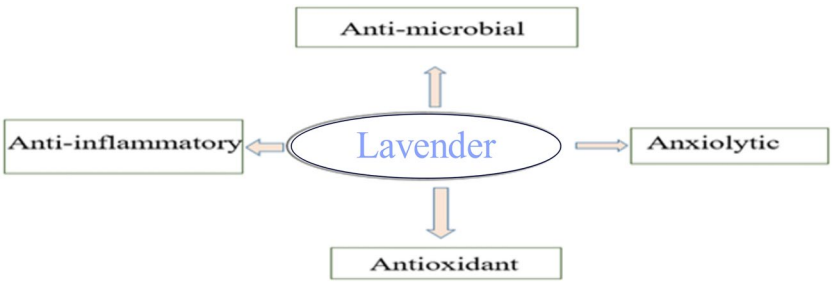



**Fig. 4** Lavender essential oil benefits

essential oil. Due to these qualities, lemongrass essential oil is used for a variety of purposes, including food preservation, orthopaedics, the treatment of muscle and skin disorders, and herbicides (Chouhan and Souvik Mukherjee 2023).

One of the blossoming plants in the Lamiaceae family is lavender (*Lavandula angustifolia*). Sedative, depressive, antimicrobial, and anaesthetic properties are few therapeutic properties present in lavender essential oil. The main

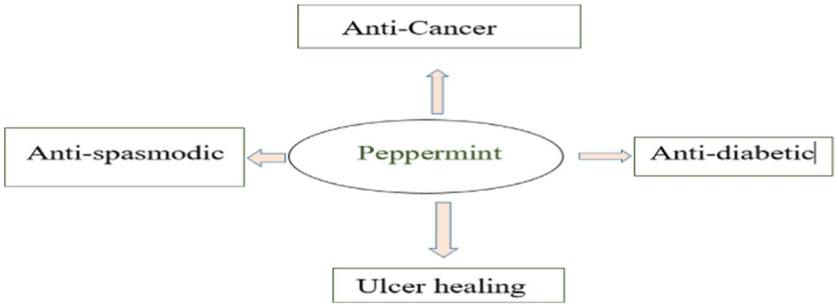



**Fig. 5** Peppermint essential oil benefits

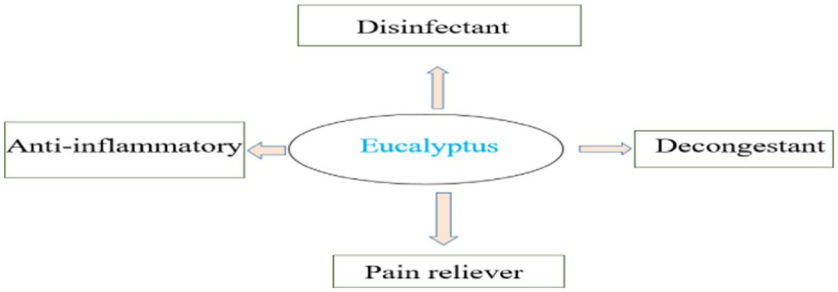



**Fig. 6** Eucalyptus essential oil benefits

constituents of lavender essential oil are linalool and linalyl acetate (Yigit et al. 2022).

A classification that includes plants in the Lamiaceae (mint family) is called *Mentha*, often known as peppermint. The main constituents of peppermint essential oil are menthone, menthol, iso-menthone, and neomenthol (Zhao et al. 2022).

The genus *Eucalyptus* is an Australian native belonging to Myrtaceae family. It is rich in bioactives (Cherfaoui et al. 2023).

Steam distillation and solvent extraction are two common traditional extraction techniques. However, the investigation of alternate energy sources for extraction processes has been prompted by the growing emphasis on sustainability and environmental concerns. For increasing the effectiveness and sustainability of essential oil extraction, solar energy, which is readily available and environmentally beneficial, offers a tempting solution (Kant and Kumar 2021).

## Solar-assisted essential oil extraction

In order to extract essential oils from plant materials, a process known as "solar-assisted essential oil extraction" combines the concepts of solar energy utilisation with conventional essential oil extraction methods. Including leaves, flowers, stems, and roots, essential oils are concentrated fragrant chemicals that are obtained from different sections of plants. They are frequently utilised for their putative medicinal effects, cosmetics, aromatherapy, and perfumes (Lukose et al. 2023). Various steps are involved in solar-assisted essential oil extraction such as solar energy collection, heating the extraction medium, plant material preparation, extraction process, and condensation.

### Solar energy collection

Solar collectors, like solar panels or solar concentrators, are used to capture solar energy. The extraction process requires thermal energy, which these collectors convert sunlight into heat energy (Kant and Kumar 2022b).

### Heating the extraction medium

A suitable extraction medium is heated using the solar energy that has been captured. Depending on the type of plant material and particular essential oil being extracted, this medium may be water or a carrier oil. The plant material's aromatic and medicinal components are released from the heated media (Lukose et al. 2023).

### Plant material preparation

To maximise the extraction surface area, the plant material containing the desired essential oil is often dried and crushed. This process facilitates the effective release of aromatic and medicinal chemicals (Kumar and Kant 2023).

### Extraction process

The plant material is in contact with the hot extraction media. The molecules of the essential oils are more easily transferred from the plant material into the media when the

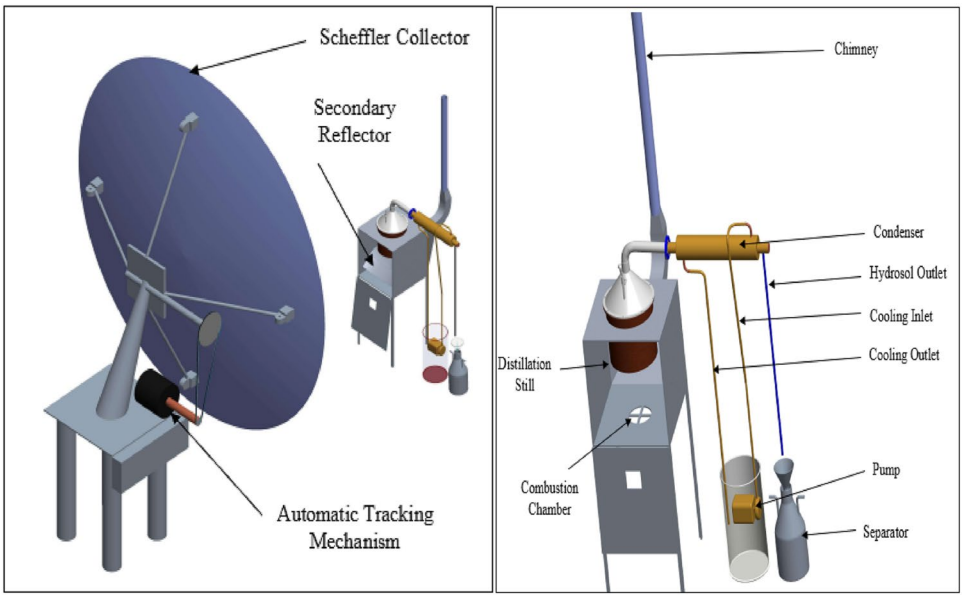



**Fig. 7** Solar and biomass distillation still for Palmarosa grass essential oil extraction (Soni and Kalbande 2023)

medium's temperature rises. Different techniques, such as steam distillation, hydrodistillation, or cold-press extraction, can be used to carry out this operation (Afzal et al. 2017).

### Condensation and separation

The essential oil and other volatile substances must be separated after being discharged into the extraction medium. This is commonly accomplished by cooling and condensing the vaporised mixture, which separates the essential oil from the extraction medium as a result of volatility variations (Munir et al. 2014).

A review of multiple investigations on solar steam distillation for the extraction of essential oils from aromatic and medicinal plants is presented. A thorough examination of the solar-assisted essential oil extraction technique, including a comparative analysis based on performance metrics such as essential oil yield and system efficiency, is done. The assessment of oil yield and system efficiency has not yet been conducted. This study identifies the plant with the maximum and minimum oil yield, as well as the solar-assisted system with the highest efficiency. The therapeutic properties of some frequently used essential oils are also examined.

## Advancements in solar steam distillation system

Gupta et al. established and created a decentralised solar distillation unit for the collection of organoleptically useful essential oil from a variety of fragrant plants. This system efficiently captures the sunlight at varying angles amid the daytime, 24 degrees being the maximum. The extraction investigations of citronella, lemongrass, and *Ocimum* by the system demonstrated enhanced oil production, i.e. 0.53%, 0.33%, and 0.20%, respectively, when contrasted to wood-base extraction setup (Gupta et al. 2024).

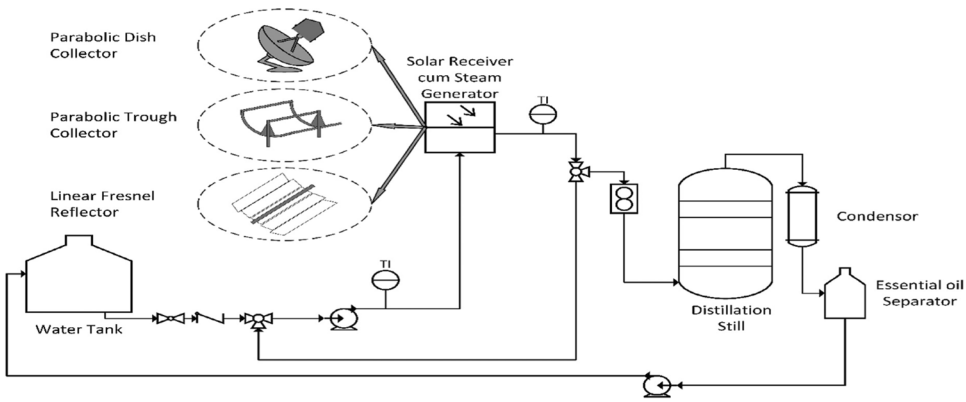



**Fig. 8** Farm distillation system for essential oil extraction (Pise and Thorat 2023)

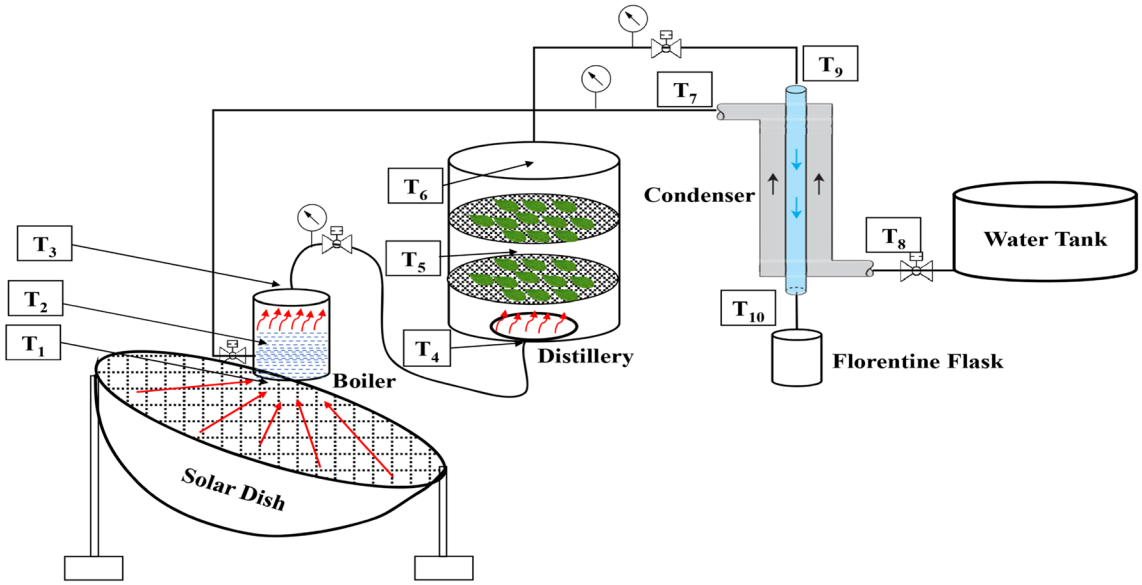



**Fig. 9** Solar steam distillation system for peppermint oil extraction (Kant and Kumar 2023b)

Kant and Kumar have done an extensive study of energy-economic and exergy-environment efficiency evaluation of a solar-powered essential oil extracting system. Peppermint and eucalyptus plants were taken and analysed. The highest and lowest system efficiency came out to be 48.68% when 2 kg of peppermint was taken into consideration and 43.25% for 6 kg of eucalyptus, respectively; however, the highest and poorest exergy efficiency of system was determined as 27.96% and 22.50% for 6 kg and 4 kg of eucalyptus leaves, respectively (Kant and Kumar 2023a).

Soni and Kalbande examined efficiency of biomass and solar energy-based systems for oil extraction from five samples of palmarosa grass which were shade dried. Figure 7 depicts Palmarosa grass essential oil extraction system based on solar and biomass distillation still. Solar system was provided with Scheffler reflector of 10 $m^2$ and combustion chamber was provided for biomass system. The maximum essential oil yield of 0.63% was obtained from the third sample which was 24-h shade-dried sample. Oil yield of 29.2 ml and 31.5 ml was obtained for the third sample with batch size of 5 kg for solar distillation system and biomass

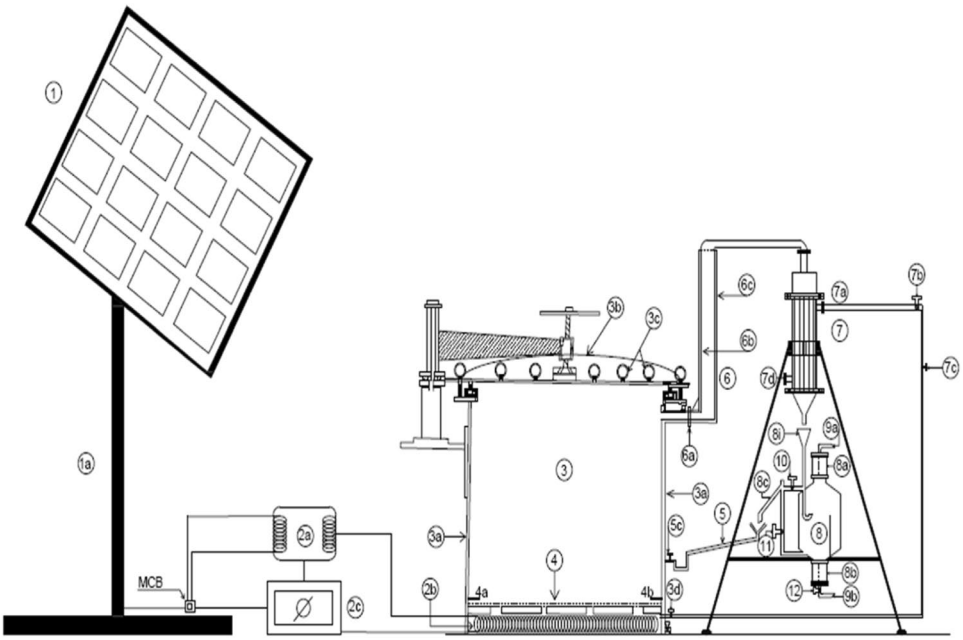



**Fig. 10** Design of solar distillation unit (Nannaware et al. 2022)

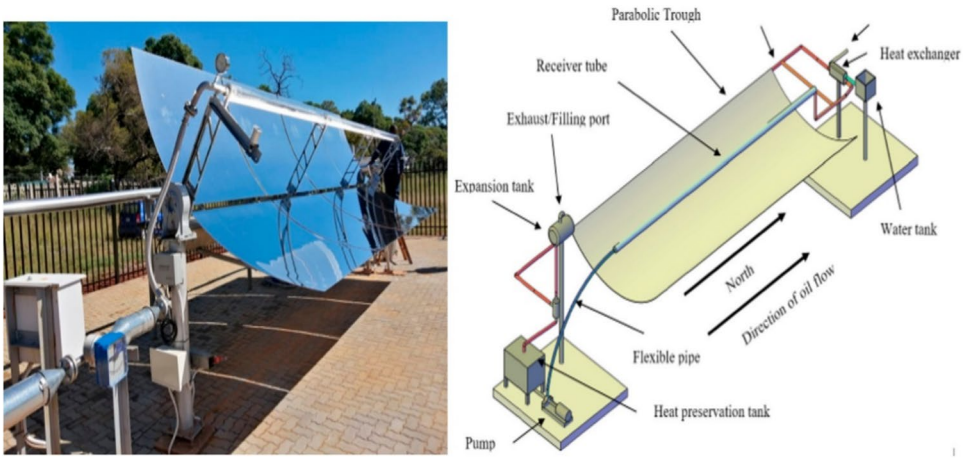



**Fig. 11** Test setup and schematic of unit driven by PTC (Manyako et al. 2022)

distillation system having thermal energy requirements of 10.45 and 10.60 kWh, respectively (Soni and Kalbande 2023).

Pise and Thorat analysed green steam solution for essential oil extraction sustainably. The schematic of farm distillation system for oil extraction is shown in Fig. 8. Results showed that 750 g of steam is required for extraction of 1 ml of betel leaf oil. Three concentrating solar collectors (CSC) were used to carry out study sustainably. Parabolic trough collector (PTC), parabolic dish collector (PDC), and linear Fresnel collector (LFR) having area of 138.87 $m^2$, 122 $m^2$, and 78 $m^2$ were utilised. According to the analysis, CSC is less expensive than briquette-based steam generation during the stipulated duration per unit generation of steam (Pise and Thorat 2023).

Kant and Kumar developed a solar-powered steam distillation setup for peppermint oil extraction from leaves (Fig. 9). A thermodynamic analysis of the developed system was done. The oil yield per unit energy and overall efficiency of the system were 6.24 ml/kWh and 41.98%, respectively. Exergy analysis with detailed study of steam line and condenser has also been done. By insulating steam line, the efficiency of the system can be increased by 2.67% (Kant and Kumar 2023b).

Nannaware et al. designed and fabricated solar distillation apparatus to extract important essential elements from aromatic plants as depicted in Fig. 10. This system has an average loading capacity of 20 kg. It is appropriate for modest and medium quantities of planting materials and can distill valuable essential oils from aromatic

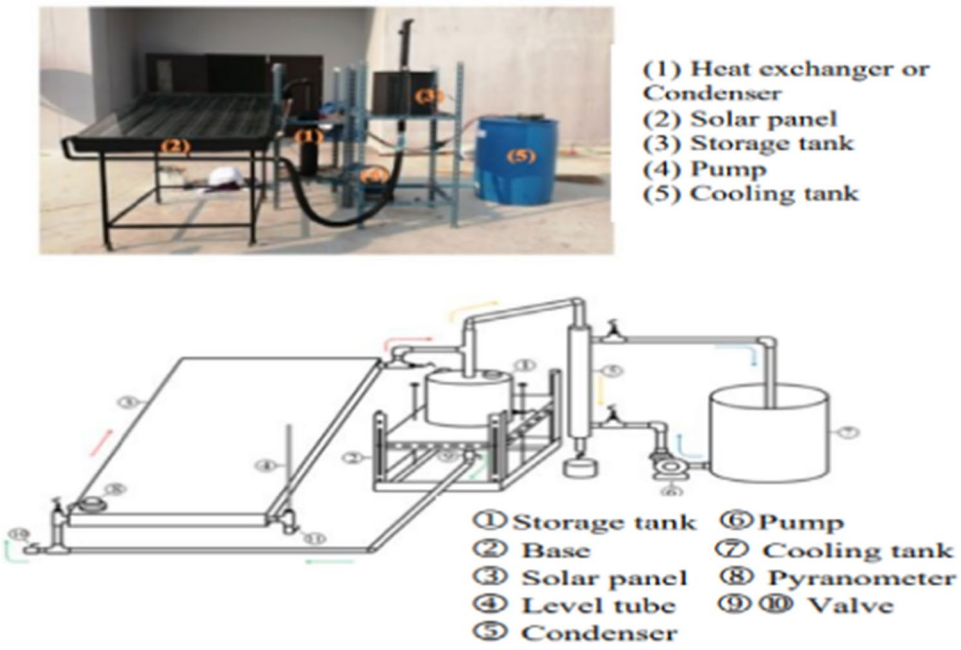



**Fig. 12** Experimental and schematic view of solar assisted distillation for citronella oil extraction (Klinbun and Panyoyai 2022)



**Fig. 13** Experimental setup of solar distillation system (Ezzarrouqy et al. 2022)

crops in 3 h. Unlike wood-based distillation apparatus, solar distillation unit (SDU) produced 0.36–0.73% of the examined plants' essential oil content. In the essential oil obtained from the SDU, the highest concentration (> 60%) of monoterpenoids and phenylpropanoids was observed. Compared to the conventional wood-based distillation operation, the cost of operating this solar aroma distillation equipment was 50% lower. Developed SDU has reduced the amount of wood used for each batch by roughly 20 kg compared to conventional wood-based distillation techniques (Nannaware et al. 2022).

Manyako et al. examined the application of the parabolic trough collector (PTC), an inexpensive CST technology, for steam distillation in an extraction procedure from lemon peels. Comparison of oil yields was done between PTC and pilot gas-powered system for lemon, orange, and mandarin and the results were 0.53, 0.67, and 1.09% for PTC and 0.44, 0.65, and 1.17% for gas-powered system. A noteworthy




**Fig. 14** Setup of built system (Radwan et al. 2020)

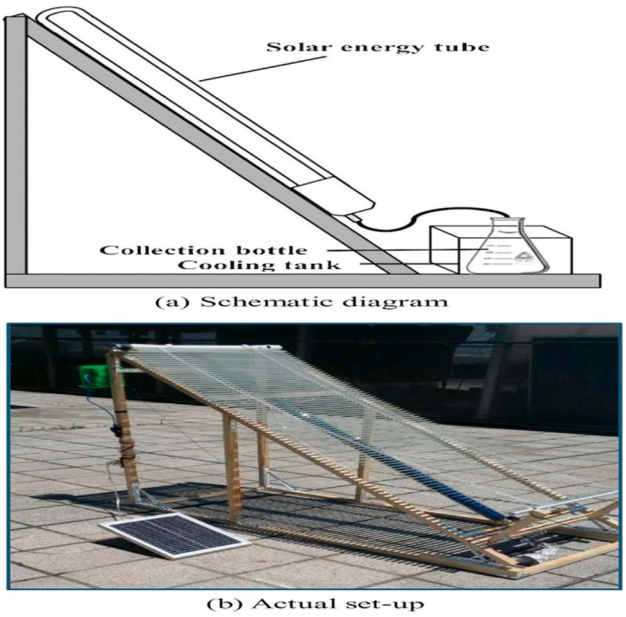



**Fig. 15** Test-based setup of solar apparatus consisting of a solar energy tube, a sunlight shade, a solar cell, a battery, a temperature controller, and an essential oil collection unit. (Yen and Lin 2017)

average system efficiency of 54.99% was attained by the PTC (Manyako et al. 2022). PTC-driven test setup and schematic are presented in Fig. 11.

Klinbun and Panyoyai studied feasibility of citronella oil extraction with the help of solar distillation technique. Comparison of conventional distillation with solar distillation was done and found that the total quantity of essential oil distilled when using solar collector instead of conventional method increased by 36.95% (Klinbun and Panyoyai 2022). Experimental and schematic view of solar-assisted distillation for citronella oil extraction can be seen in Fig. 12.

Ezzarrouqy et al. determined optical and thermal losses pertaining to each step of solar distillation system for which the developed setup can be seen in Fig. 13. Energy and exergy analysis was also performed. Optical efficiency of 50.97% and exergy efficiency of 26.62% were achieved with batch size of 6 kg of rosemary leaves and 4 h of distillation time having average intensity of radiation 849.1W/m$^2$. The oil yield and overall efficiency of solar distillation system considering all the heat losses at every step of the system is 6.18 ml/kWh and 40%, respectively. With insulation of the system, the efficiency can be achieved up to 42.42%. Moreover, this solar distillation system provides greener energy extraction approach as it prevents emission of 12.10 kg of carbon dioxide into the environment (Ezzarrouqy et al. 2022).

Radwan et al. built a solar steam distillation system for extracting lavender volatile oil as shown in Fig. 14. As a result of batch sizes of the entire lavender plant (400, 600, and 800 g) and a change in boiler inlet water flow rate (1.00, 1.25, and 1.50 l/h), the efficiency of the integrated solar distillation system is evaluated. The solar steam distillation system is an effective way to extract high-quality lavender volatile oil, with 7.3 ml system productivity, 60.25% system efficiency, 98.13% extraction efficiency, and 0.785% (w/w) essential oil yield, according to the

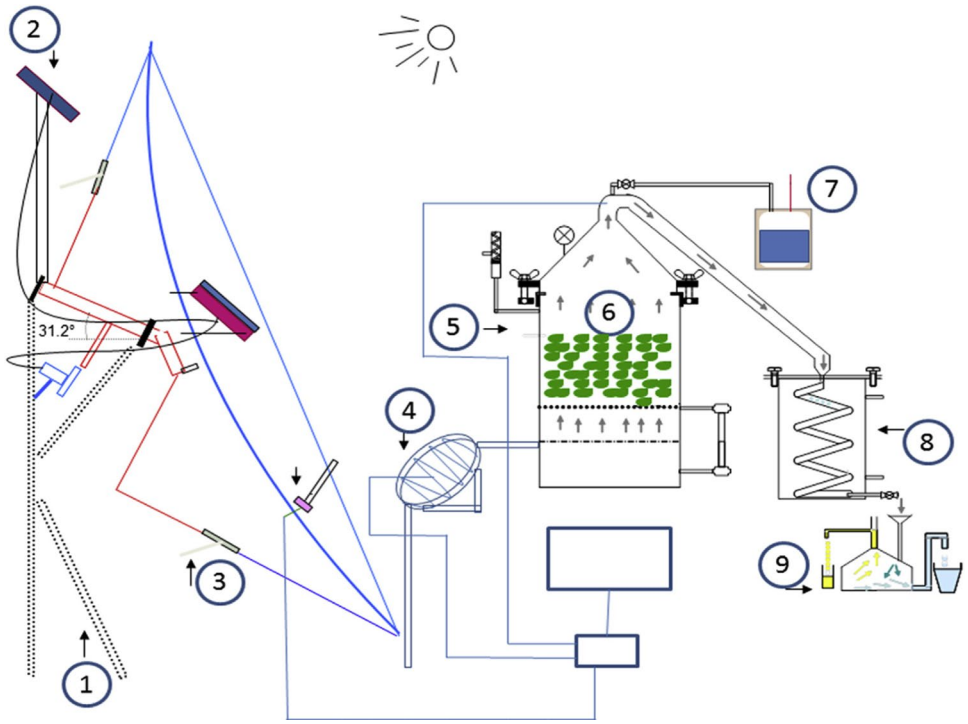



**Fig. 16** Solar distillation apparatus (Afzal et al. 2017)

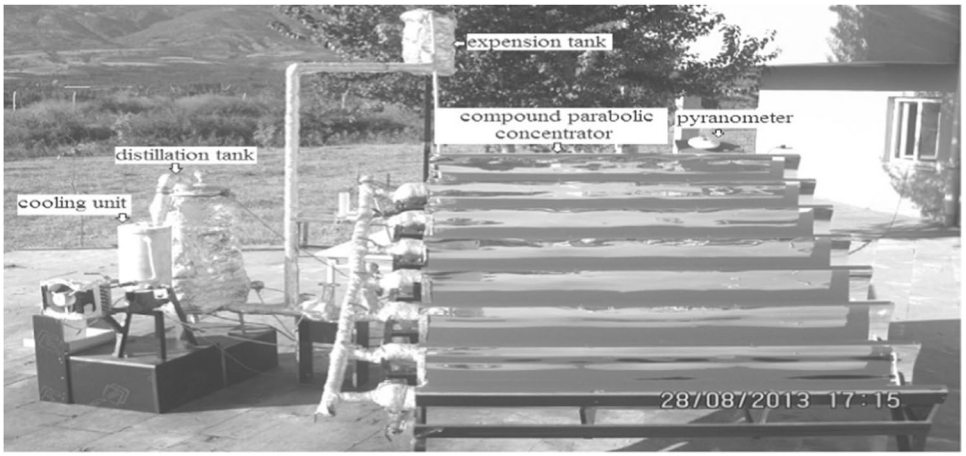



**Fig. 17** Solar CPC–based assembled distillery unit (Kültürel and Tarhan 2016)

circumstances of 1.25 l/h boiler inlet rate of flow of water and 800-g batch size. A 400-g batch size required the solar steam distillation system to extract oil in 2 h, whereas batches of 600 and 800 g required 3 and 4 h, respectively (Radwan et al. 2020).

Hilali et al. contrasted conventional steam distillation (CSD) with solar steam distillation (SSD) for deodorising and extracting essential oils from dried rosemary leaves. With normal technique extraction and analysis, SSD appears to maintain antioxidants more than CSD. After SSD, it came out that the antioxidant activity and polyphenol levels were both retained. Improved antioxidant preservation was observed as both carnosic and rosmarinic acid had higher values after SSD compared to CSD, a reduction in energy consumption given that this energy is a sustainable charge-free source, and a decrease in extraction time was achieved given that the SSD can reach roughly the same yield in 2 h instead of 3 h (37% decrease in extraction time). Analysis of the total amount of polyphenols and DPPH free radical scavenging activity revealed that the antioxidant and polyphenolic activity was well-preserved during the SSD procedure. Overall, the results suggested that SSD is an innovative, sustainable, and charge-free deodorising technique permitting an effective extraction of both volatiles and non-volatiles from rosemary (Hilali et al. 2018).

Yen and Lin assessed the yield of the essential oil of *Cymbopogon citratus* produced by hydrodistillation and solar energy, as well as its chemical makeup and antioxidant activity. According to the findings, both methods produce oil with equal yields, i.e. 1.28% in case of solar extraction and 1.30% in case of hydrodistillation and with significant components, but the oil produced using solar energy has superior antioxidant activity. Additionally, solar energy extraction is free of grid electricity while hydrodistillation uses it; i.e.

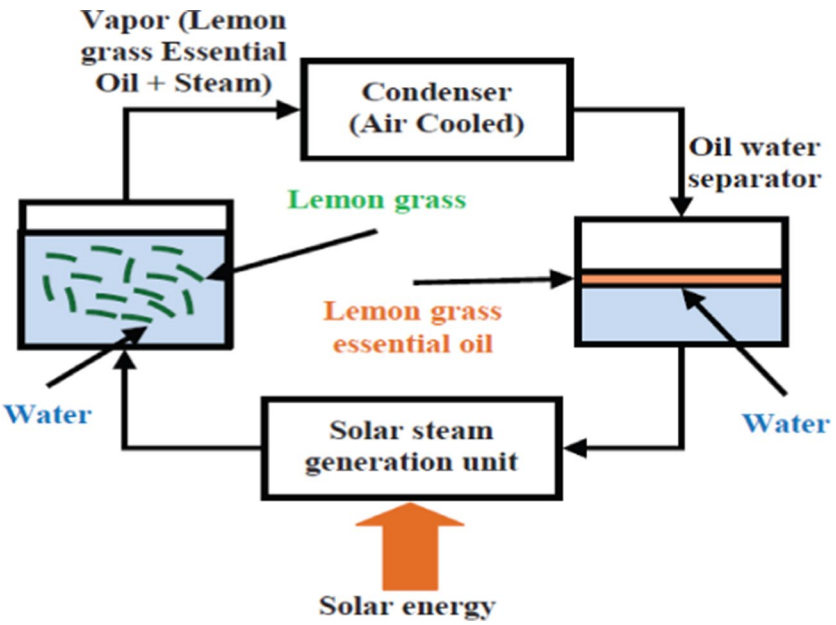



**Fig. 18** Semi-continuous distillation system for lemon grass (Ragula et al. 2016)

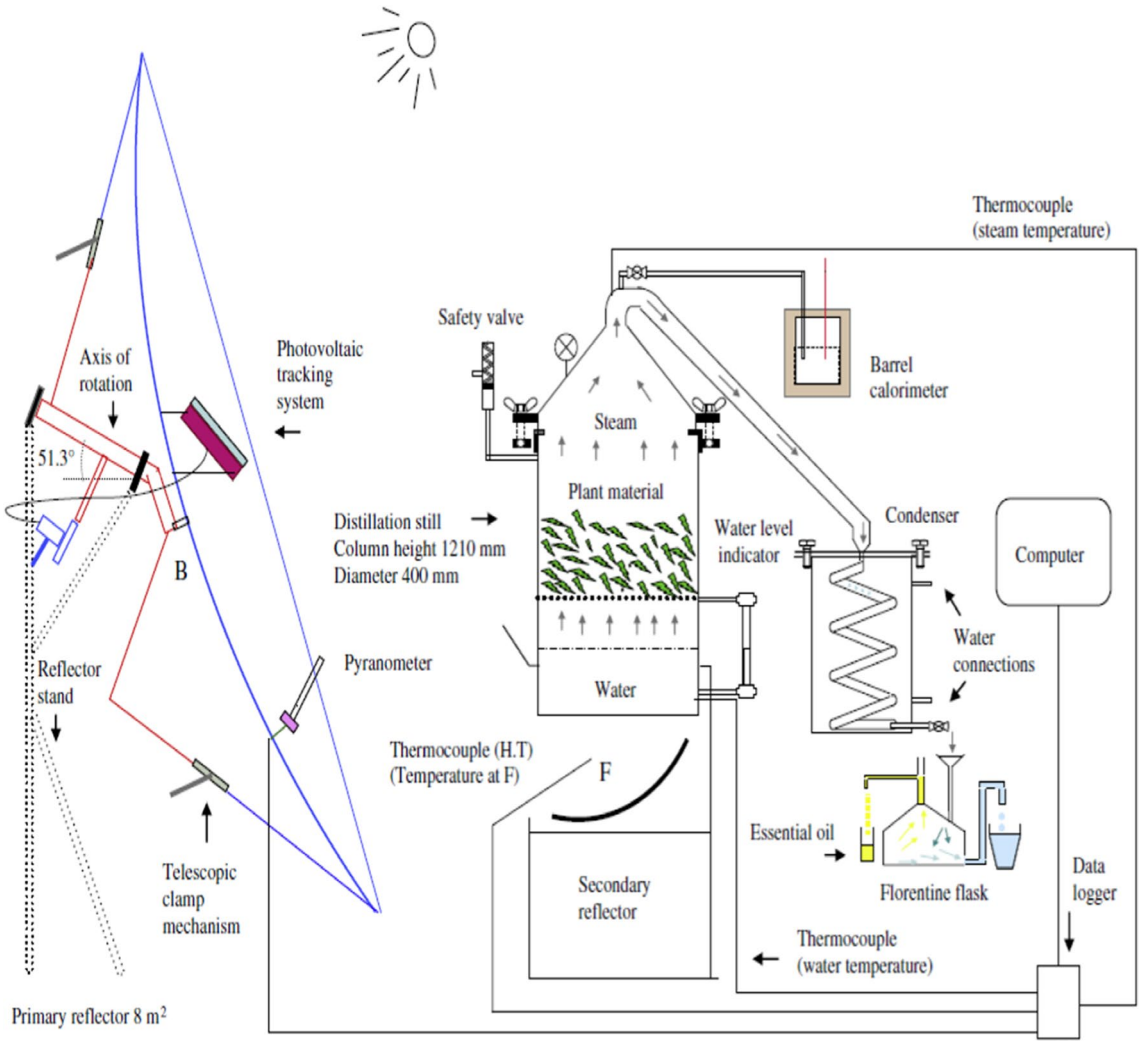



**Fig. 19** Solar distillation unit (Munir et al. 2014)

for 1 kg of oil extraction, 858 kWh of electricity is required. Therefore, the research demonstrates that the extraction of *Cymbopogon citratus* essential oil using solar energy is a viable and sustainable alternative to conventional methods (Yen and Lin 2017). Schematic and actual pictorial views of solar apparatus can be seen in Fig. 15.



**Fig. 20** Experimental setup for hydrodistillation of fresh orange peel oil (Maiti et al. 2013)

**Table 1** Summary of various studies on solar energy-based systems for essential oil extraction from aromatic and medicinal plants

| Author | Type of collector | Plant material | Process parameters | Performance parameters | | Remarks | References |
|---|---|---|---|---|---|---|---|
| | | | | Yield | Efficiency | | |
| Soni and Kalbande | Scheffler reflector (solar dish) and combustion chamber | Palmarosa grass | • Batch size (5 kg)<br>• Thermal energy requirement (10.45 and 10.60 kWh) | 29.2 ml and 31.5 ml | ___ | The maximum oil yield of 0.63% was obtained for 24-h shade-dried sample | Soni and Kalbande (2023) |
| Pise and Thorat | Parabolic trough collector (PTC), parabolic dish collector (PDC), and linear Fresnel collector (LFR) | Betel leaf | • Batch size (250 kg)<br>• Thermal energy requirement (24.9 kWh) | 1% | ___ | CSC is less expensive than briquette-based steam generation during the stipulated duration per unit generation of steam | Pise and Thorat (2023) |
| Manyako et al | Parabolic trough collector and gas powered system | • Lemon peels<br>• Orange<br>• Mandarin | • Energy requirement for PTC 2.387 kWh for orange, 2.417 kWh for lemon, 2.506 kWh for mandarin<br>• Energy requirement for gas-powered system 4.41 kWh for orange, 3.307 kWh for lemon, 4.198 kWh for mandarin | 0.67, 0.53, and 1.09% for PTC<br>0.65, 0.44, and 1.17% for gas-powered system | 54.99% | 0.53, 0.67, and 1.09% | Manyako et al. (2022) |
| Kant and Kumar | Solar dish | Peppermint | • Batch size (6 kg)<br>• Thermal energy requirement (11.81 kWh) | 6.24 ml/kWh | 41.98% | Without insulation, oil yield per unit of energy is 6.24 ml/kWh and efficiency is 41.98%. With insulation, efficiency is 46.36% | Kant and Kumar (2023b) |

**Table 1** (continued)

| Author | Type of collector | Plant material | Process parameters | Performance parameters | | Remarks | References |
|---|---|---|---|---|---|---|---|
| | | | | Yield | Efficiency | | |
| Nannaware et al | Solar PV panel | Hybrids of *Ocimum* species | • Batch size (20 kg)<br>• Thermal energy requirement (13.5 kWh) | (0.36–0.73%) | ___ | Unlike wood-based distillation apparatus, solar distillation unit (SDU) produced 0.36–0.73% of the examined plants' essential oil content. In the essential oil obtained from the SDU, the highest concentration (> 60%) of monoterpenoids and phenylpropanoids were observed | Nannaware et al. (2022) |
| Klinbun and Panyoyai | Solar panel | Lemongrass | • Batch size (40 L) | 590.84 ml | ___ | Comparison of conventional distillation with solar distillation was done and found that the total quantity of essential oil distilled when using solar collector instead of conventional method increased by 36.95% | Klinbun and Panyoyai (2022) |
| Ezzarrouqy et al | Solar dish | Rosemary leaves | • Batch size (6 kg)<br>• Thermal energy requirement (20.23 kWh) | 6.18 ml/kWh | 40.00% | Determined optical and thermal losses pertaining to each step of solar distillation system | Ezzarrouqy et al. (2022) |

**Table 1** (continued)

| Author | Type of collector | Plant material | Process parameters | Performance parameters | | Remarks | References |
|---|---|---|---|---|---|---|---|
| | | | | Yield | Efficiency | | |
| Radwan et al | Parabolic solar dish collector | Lavender | • Batch size (400, 600, and 800 g) | 0.785% (w/w) | 60.25% | The solar steam distillation system is an effective way to extract high-quality lavender volatile oil, with 7.3 ml system productivity, 60.25% system efficiency, 98.13% extraction efficiency, and 0.785% (w/w) essential oil yield | Radwan et al. (2020) |
| Hilali et al | Solar dish collector and electric resistance | Rosemary leaves | • Batch size (500 g)<br>• Thermal energy requirement (4.4 kW) | 0.82% and 0.85% | ___ | Analysis of the total amount of polyphenols and DPPH free radical scavenging activity revealed that the antioxidant and polyphenolic activity was well-preserved during the SSD procedure | Hilali et al. (2018) |

**Table 1** (continued)

| Author | Type of collector | Plant material | Process parameters | Performance parameters | | Remarks | References |
|---|---|---|---|---|---|---|---|
| | | | | Yield | Efficiency | | |
| Yen and Lin | Solar evacuated tubes and conventional Clevenger apparatus | Cymbopogon citratus | • Batch size (500 g) each for solar system and hydrodistillation | 1.30% and 1.28% | ___ | Both methods produce oil with equal yields, i.e. 1.28% in case of solar extraction and 1.30% in case of hydrodistillation and with significant components, but the oil produced using solar energy has superior antioxidant activity | Yen and Lin (2017) |
| Afzal et al | Solar dish collector | Eucalyptus and *Mentha* leaves | • Batch size (10 kg) for each<br>• Thermal energy requirement (2.32 kWh for eucalyptus) and (2.54 kWh for *Mentha* leaves) | 0.59% and 0.40% | ___ | Analysed the performance of hybrid solar system for essential oil extraction from eucalyptus and *Mentha* leaves | Afzal et al. (2017) |
| Pesimo | Direct sunlight | • Sambong<br>• Oregano<br>• Lemongrass<br>• Roses | | 0.72%, 0.36%, 0.67%, 0.32% | ___ | Four air-dried samples of various plant materials, i.e. sambong, oregano, lemongrass, and roses. The findings of study were as follows: the extraction rate for oregano sample was the fastest (3.4 ml/day) and it was slowest for rose petals (1.8 ml/day) | Pesimo (2017) |

**Table 1** (continued)

| Author | Type of collector | Plant material | Process parameters | Performance parameters | | Remarks | References |
|---|---|---|---|---|---|---|---|
| | | | | Yield | Efficiency | | |
| Kültürel and Tarhan | Compound parabolic collectors (CPCs) | Peppermint | • Batch size (5 kg) | ___ | 27.9% to 34.3% | 5 kg batch size of plant material was taken and 3 trials of experimentation were done. The combined thermal efficiency of the unit came out within a range of 27.9 to 34.3% | Kültürel and Tarhan (2016) |
| Ragula and Devanathan | Solar concentrator | Lemongrass | • Batch size (10 kg) | 2.2% | ___ | With this system, more than 8 times extraction of essential oil can be done as compared to other commercial methods of extraction and around 55 L of oil can be produced every year | Ragula et al. (2016) |
| Munir et al | Scheffler reflector (solar dish) | • Melissa<br>• Peppermint<br>• Rosemary<br>Cumin<br>• Cloves | • Batch size (10–20 kg)<br>• Thermal energy requirement (18.58 kWh) | 1.425 ml<br>28.2 ml<br>4.6 ml<br>12.4 ml<br>44 ml | 33.21% | 10–20 kg batch size of plant material of various kinds was processed. 33.21% efficiency was calculated for solar distillery | Munir et al. (2014) |
| Maiti et al | Square parabolic trough concentrator | Fresh orange peels | • Batch size (409.46 to 514.05 g) | 0.49–2.16% | 8.37% | Batch of 400–500 g of peels was distilled. The oil yield variation was noted in between 0.49 and 2.16% (w/w) | Maiti et al. (2013) |

Pesimo designed a solar-based extractor of essential oil. The study focussed on four air-dried samples of various plant materials, i.e. sambong, oregano, lemongrass, and roses. The findings of the study were as follows: the extraction rate for oregano sample was the fastest (3.4 ml/day) and it was slowest for rose petals (1.8 ml/day). The oil yield for sambong was highest (0.72%) and lowest for rose petals (0.185%) (Pesimo 2017).

Afzal et al. developed hybrid solar system for essential oil extraction from eucalyptus and *Mentha* leaves and analysed the performance for which the schematic is as shown in Fig. 16. The yield of 0.59% w/w and 0.40% w/w was obtained from eucalyptus and *Mentha* leaves, respectively (Afzal et al. 2017).

Kültürel and Tarhan developed a solar-based distillation system comprising seven compound parabolic collectors (CPCs). Five kilograms batch size of plant material was taken and three trials of experimentation were done. The combined thermal efficiency of the unit came out within a range of 27.9 to 34.3% (Kültürel and Tarhan 2016). Figure 17 depicts pictorial view of solar CPC–based assembled distillery unit.

Ragula and Devanathan tested a lab scale–based semi-continuous distillation system. They came out with findings that more than eight times extraction of essential oil can be done as compared to other commercial methods of extraction. Figure 18 shows the schematic diagram of lemongrass semi-continuous distillation system. With this system around 55 l of oil can be produced every year in Vellaramkunnu village of Kerala (India) (Ragula et al. 2016).

Munir et al. designed, developed, and experimented Scheffler reflector-based solar distillation system having an area of 8 $m^2$ as shown in the schematic diagram of solar distillation unit (Fig. 19). Mathematical explanation of designing various components of solar distillation unit is given. The flexible crossbars have been used to get the desired design of reflectors. Separation of hydrosol from essential oil was done with the help of Florentine apparatuses. With 1.548 kWh of thermal power, 33.21% efficiency was calculated for solar distillery. On a sunny day, the thermal energy obtained from the unit was 18.58 kWh. Ten to 20 kg batch size of plant material of various kinds was processed. 3.5 kWh of energy was needed for processing of 10 kg batch size. Therefore, 4–5 batches were processed notably with available energy of 18.58 kWh (Munir et al. 2014).

Maiti et al. carried out distillation of essential oil from fresh orange peels. PTC having aperture area of 1.3 $m^2$ was used. Batch of 400–500 g of peels was distilled. The oil yield variation was noted in between 0.49 and 2.16% (w/w). The peak efficiency of hydrodistillation system was found to be 8.37% (Maiti et al. 2013). Experimental setup for hydrodistillation of fresh orange peel oil can be viewed in Fig. 20.

## Performance parameters

### Essential oil yield (EOY)

EOY of developed system is explained as the ratio of mass of oil extracted to mass of plant per batch size and calculated by following Eq. (1) (Kant and Kumar 2022b).

$$EOY = \frac{m_{oil}}{m_{plant}} \tag{1}$$

In Eq. (5), $m_{oil}$ refers to the mass of oil in g or kg and $m_{plant}$ refers to the mass of plant per batch size in g or kg.

### Thermal efficiency of solar distillation system

Efficiency of the system is given by output energy to input energy. In case of solar distillation system, thermal efficiency can be written as given in Eq. (2)

$$\eta = \frac{\dot{m} \times L}{I_G \times A_S} \times 100 \tag{2}$$

## Conclusions

The present study summarised various studies performed on solar-based systems for essential oil extraction from aromatic and medicinal plants (Table 1). Moreover, various solar systems were compared based on the performance in terms of yield and efficiency. The following conclusions have been made based on the present study:

- Maximum and minimum essential oil yield have been found for lemongrass (2.2%) and rosemary (0.046%), respectively.
- For solar dish type collectors, maximum oil yield has been reported for lemongrass plant (2.2%).
- For parabolic type collectors, maximum oil yield has been reported for fresh orange peels (2.16%).
- Maximum overall efficiency has been noted for system with parabolic dish collector (60.25%). Hence, it can be said that the most effective solar energy-based system for essential oil extraction is one with parabolic dish collector having 60.25% efficiency.

The techniques discussed above can promote local economies, especially in remote or underdeveloped areas with restricted access to energy. The growing confluence of artificial intelligence (AI) technologies presents a possibility to create smart solar-based extraction systems. Such systems could be fitted with sensors to track humidity, temperature,

and energy usage in an instant, altering variables automatically to optimize the extraction procedures.

### Nomenclature

**$A_S$: Solar collector area ($m^2$); $I_G$: Global solar radiation ($W/m^2$); $L$: Latent heat of vaporisation of peppermint (kJ/kg); $\dot{m}$: Yield (kg/h); $m$: Mass (kg); $m_{oil}$: Mass of oil (g or kg); $m_{plant}$: Mass of plant (g or kg)**

### Abbreviations

***AI*: Artificial intelligence; *CSC*: Concentrating solar collectors; *CSD*: Conventional steam distillation; *EOY*: Essential oil yield; *LFR*: Linear Fresnel collector; *PTC*: Parabolic trough collector; *PDC*: Parabolic dish collector; *SDU*: Solar distillation unit; *SSD*: Solar steam distillation**

### Greek letters

**$\eta$: Efficiency**

**Acknowledgements** The first author would like to thank AICTE for awarding the fellowship for pursuing a Ph.D. in the Mechanical Engineering Department. Authors show gratitude to the Centre for Energy and Environment, Delhi Technological University, Delhi (India), for providing the facilities to conduct the present study.

**Author contribution** Gourav Dhingra: visualisation methodology, writing—original draft and investigation. Anil Kumar: conceptualisation, writing—review and editing, and supervision.

**Data availability** All data are given in either table or figure.

**Code availability (software application or custom code)** Not applicable.

## Declarations

**Ethics approval** Not applicable.

**Consent to participate** Not applicable.

**Consent for publication** Not applicable.

**Competing interests** The authors declare no competing interests.

## References

Afzal A, Munir A, Ghafoor A, Alvarado JL (2017) Development of hybrid solar distillation system for essential oil extraction. Renew Energy 113:22–29. https://doi.org/10.1016/j.renene.2017.05.027

Cherfaoui B, Chebrouk F, Hassaine R et al (2023) Hemi-synthesis of novel chiral benzodiazepine derivatives from Eucalyptus citriodora essential oil : 2D NMR experiments and differential scanning calorimetry study of diastereoisomers. Nat Prod Res 0:1–12. https://doi.org/10.1080/14786419.2023.2185888

Chouhan KBS, Souvik Mukherjee VM (2023) A Deeper Learning approach in exploring the eventualities of solvent-free microwave-based extraction of lemongrass essential oil: understanding milky emulsion, biomass temperature, first drop appearance, oil quality and biorefinery. Sustain Chem Pharm 33:101113. https://doi.org/10.1016/j.scp.2023.101113

Ezzarrouqy K, Hejjaj A, Idlimam A et al (2022) Study of the energetic, exergetic, and thermal balances of a solar distillation unit in comparison with a conventional system during the distillation of rosemary leaves. Environ Sci Pollut Res 29:25709–25722. https://doi.org/10.1007/s11356-021-17612-1

Gupta MK, Singh S, Deorankar AV et al (2024) Energy, environmental and economic evaluation of decentralized solar distillation and wood based distillation for extraction of organoleptically superior essential oils: a comparative analysis. Sustain Energy Technol Assess 69:103895. https://doi.org/10.1016/j.seta.2024.103895

Hilali S, Fabiano-Tixier AS, Elmaataoui M et al (2018) Deodorization by solar steam distillation of rosemary leaves prior to solvent extraction of rosmarinic, carnosic, and ursolic acids. ACS Sustain Chem Eng 6:10969–10979. https://doi.org/10.1021/acssuschemeng.8b02347

Kant R, Kumar A (2021) Advancements in steam distillation system for oil extraction from peppermint leaves. Mater Today Proc 47:5794–5799. https://doi.org/10.1016/j.matpr.2021.04.123

Kant R, Kumar A (2022a) Review on essential oil extraction from aromatic and medicinal plants: techniques, performance and economic analysis. Sustain Chem Pharm 30:100829. https://doi.org/10.1016/j.scp.2022.100829

Kant R, Kumar A (2022b) Environmental Effects Process optimization of conventional steam distillation system for peppermint oil extraction. Energy Sources, Part A Recover Util Environ Eff 44:3960–3980. https://doi.org/10.1080/15567036.2022.2069886

Kant R, Kumar A (2023a) Energy-economic and exergy-environment performance evaluation of solar energy integrated essential oil extraction system. Sol Energy 265:112101. https://doi.org/10.1016/j.solener.2023.112101

Kant R, Kumar A (2023b) Thermodynamic analysis of solar assisted steam distillation system for peppermint oil extraction. J Food Process Eng 46:1–11. https://doi.org/10.1111/jfpe.14228

Klinbun J, Panyoyai N (2022) Experimental study of citronella oil extraction by solar distillation using bubble pump technique. Sci Technol Asia 27:321–329. https://doi.org/10.14456/scitechasia.2022.89

Kültürel Y, Tarhan S (2016) Birleşik parabolik toplaçlı (BPT) güneş enerjili uçucu yağ damıtma sistemi. Anadolu J Agric Sci 31:72. https://doi.org/10.7161/anajas.2016.31.1.72-83

Kumar A, Kant R (2023) Empirical relation to predict essential oil yield in conventional distillation system. J Polym Compos 10(1):2023. https://doi.org/10.37591/JoPC

Lukose R, Kalbande SR, John J, Lukose R (2023) Utilisation of solar energy for essential oil extraction : an overview.https://doi.org/10.1080/01430750.2022.2163688

Maiti S, Bhatt C, Patel P, Ghosh PK (2013) Use of solar thermal energy in the hydrodistillation of essential oil 063106: https://doi.org/10.1063/1.4766819

Manyako KE, Chiyanzu I, Mulopo J, Abdulsalam J (2022) Pilot-scale evaluation of concentrating solar thermal technology for essential oil extraction and comparison with conventional heating sources for use in agro-based industrial applications. ACS Omega 7:20477–20485. https://doi.org/10.1021/acsomega.1c06879

Munir A, Hensel O, Scheffler W et al (2014) Design, development and experimental results of a solar distillery for the essential oils extraction from medicinal and aromatic plants. Sol Energy 108:548–559. https://doi.org/10.1016/j.solener.2014.07.028

Nannaware AD, Sai Kumar CM, Srivastava S et al (2022) Eco-friendly solar distillation apparatus for improving the yield of essential

oils with enhancing organoleptic characteristics. Renew Energy 191:345–356. https://doi.org/10.1016/j.renene.2022.03.147

Önem E (2022) New green solutions against bacterial resistance: palmarosa (Cymbopogon martini) essential oil and quorum sensing. Sustain Chem Pharm 25:100587. https://doi.org/10.1016/j.scp.2021.100587

Pascual-Mathey LI (2022) Elaboration of essential oil nanoemulsions of rosemary (Rosmarinus officinalis L.) and its effect on liver injury prevention. Food Bioprod Process 134:46–55. https://doi.org/10.1016/j.fbp.2022.04.006

Pesimo AR (2017) Harnessing the solar energy in extracting essential oil for community based perfumery and aromatherapy 4:1–11.https://doi.org/10.4236/oalib.1103905

Pise VH, Thorat BN (2023) Green steam for sustainable extraction of essential oils using solar steam generator: a techno-economic approach. Energy Nexus 9:1–10. https://doi.org/10.1016/j.nexus.2023.100175

Radwan MN, Morad MM, Ali MM, Wasfy KI (2020) A solar steam distillation system for extracting lavender volatile oil. Energy Rep 6:3080–3087. https://doi.org/10.1016/j.egyr.2020.11.034

Ragula UBR, Devanathan S, Mohan R (2016) Solar based lemon grass essential oil distillation for sustainability and livelihood in tribal community. GHTC 2016 - IEEE Glob Humanit Technol Conf Technol Benefit Humanit Conf Proc 738–744. https://doi.org/10.1109/GHTC.2016.7857360

Soni K, Kalbande S (2023) Bioresource Technology Reports Influence of energy sources and shade drying period on oil yield and chemical composition of the essential oil extracted from Palmarosa grass (Cymbopogon martinii) through hydro distillation method. Bioresour Technol Rep 22:101491. https://doi.org/10.1016/j.biteb.2023.101491

Teshale F, Narendiran K, Beyan SM, Srinivasan NR (2022) Extraction of essential oil from rosemary leaves : optimization by response surface methodology and mathematical modelling. Appl Food Res 2:100133. https://doi.org/10.1016/j.afres.2022.100133

Yen HY, Lin YC (2017) Green extraction of Cymbopogon citrus essential oil by solar energy. Ind Crops Prod 108:716–721. https://doi.org/10.1016/j.indcrop.2017.07.039

Yigit NO, Metin S, Didinen BI, Hakan Didinen OO (2022) Effect of lavander (Lavandula angustifolia) and laurel (Laurus nobilis) essential oils as anesthesics in rainbow trout (Oncorhynchus mykiss). Aquaculture 557:738328. https://doi.org/10.1016/j.aquaculture.2022.738328

Zhao H, Ren S, Yang H et al (2022) Biomedicine & Pharmacotherapy Peppermint essential oil : its phytochemistry, biological activity, pharmacological effect and application. Biomed Pharmacother 154:113559. https://doi.org/10.1016/j.biopha.2022.113559

**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.

ISSN 1068-3739, Russian Meteorology and Hydrology, 2024, Vol. 49, No. 10, pp. 896–903. Allerton Press, Inc., 2024.
Russian Text The Author(s), 2024, published in Meteorologiya i Gidrologiya, 2024, No. 10, pp. 78–87.

# Seasonal Variation of Aerosol Distribution and Optical Depth over the Indo-Gangetic Belt

**B. Kumar[a], B. P. Singh[b]*, P. Chandra[c], A. Yadav[b], V. Singh[b], and J. Masih[d]**

[a]*Department of Physics, Lajpat Rai College, Sahibabad, Ghaziabad (U.P.), 201005 India*
[b]*Department of Environmental Studies, Central University of Haryana, Haryana, 123031 India*
[c]*Department of Biotechnology, Delhi Technological University (Delhi College of Engineering), Shahbad Daultapur, Bawana Road Delhi, 110042 India*
[d]*Jamson Masih, Chemistry Department, Wilson College, Mumbai, 400007 India*
**e-mail: bpsingh0783@gmail.com*

Received April 2, 2023
Revised February 8, 2024
Accepted April 2, 2024

**Abstract**—Rapid industrialization, urbanization, and synoptic conditions have posed serious environmental concerns in South Asian countries, especially in the most densely populated region of the Indo-Gangetic Belt (IGB) in the last few years. In this paper, the authors are trying to show the aerosol optical depth (AOD) variation within the Indo-Gangetic Belt region from 2007 to 2012 to reveal a relationship between regional climate change and air quality. The present study is based on the online data from the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) mission. In winter, AOD is dominated by fine-mode aerosols produced by numerous anthropogenic origins. In contrast, summer months or the pre-monsoon periods are characterized by abrasive mineral dust and the dust coming mainly from the Thar Desert region. The seasonal variation of AOD and aerosol vertical distribution over IGB are determined with the help of the particle depolarization ratio, color ratio, and backscatter coefficient (at 532 nm).

**DOI:** 10.3103/S1068373924100078

*Keywords:* Aerosol backscatter, aerosol optical depth, Indo-Gangetic Belt, particle depolarization ratio, color ratio

## 1. INTRODUCTION

The Indo-Gangetic Belt (IGB) region is a home to nearly seven hundred million people, which makes it a highly populated river basin. It stretches from the west corner of India (i.e., Pakistan) and up to Bangladesh in the east [3, 29]. It is a region, where the Indus and Ganges rivers, which are downstream from the towering Himalayas in the north, are situated. The rapid growth of industrialization and urbanization, as well as the synoptic conditions have posed serious environmental concerns to South Asian countries, especially in the most densely populated region of the IGB in the last few years. Due to an increasing concentration of anthropogenic aerosols [9], the air quality and climate over IGB have been severely affected. A higher concentration of anthropogenic aerosol during the dust storms and cold dense haze in spring over northern India have affected millions of people [33, 34].

The IGB suffers from the western disturbances, moving from west to east and leading to an intense aerosol effect over the region in winter. The effect of aerosols on the prediction of climate change is controversial due to multiple factors such as their chemical and physical properties, large temporal and spatial density deviation, and complicated mechanisms of interaction with radiation and clouds [27]. More specifically, understanding vertical aerosol distribution is important for evaluating aerosol radiation impact on the stability of the atmosphere and its thermal structure [30]. During the last two decades, many researchers have tried to investigate the distribution of aerosols, their physical and chemical properties, and transport of some aerosols from the South Asian region, particularly the Indo-Gangetic region, through the use of vari-

ous methods of modeling and observation [20, 26, 29, 32, 42, 46]. Several works have indicated that aerosol is considered a major source of the changes in the climatic system because of its ability to absorb and scatter solar radiation [10, 39]. The spatial and temporal distribution of aerosol may lead to uncertainty of aerosol radiative forcing at the regional and global levels, mainly due to significant heterogeneity of its characteristic properties [15, 40]. Furthermore, various scientific studies based on ground-based measurements and satellite observations allow estimating the overall increase in aerosol optical depth (AOD) across India, particularly in the IGB region [17, 22, 28].

The main sources of aerosol in the IGB have both anthropogenic (urban and industrial aerosol) and natural (mostly dust) origin, which demonstrates a significant seasonal variability at the regional level [35]. According to [23], aerosol plays a significant role in radiative flux in the monsoon system over the Ganges basin. The authors of [13], with the use of the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) over the same region, indicated that dust was observed at higher altitudes in spring. The presence of such coarse dust particles of aerosol plays a significant role in the regulation of the radiative balance over the IGB region during the spring season.

According to the data on aerosol distribution, the IGB is one of the most polluted regions in the world [5, 16, 41]. The variations in monsoon evolution and hydroclimate are due to the presence of high aerosol loading. Surface cooling is also affected by high aerosol loading, which results from the reduction of incident solar radiation and also from the heating of the troposphere caused by absorption (direct effect). Moreover, aerosol has an ability to alter clouds and precipitation based on microphysical phenomena occurring in it (indirect effect) [19, 29]. Aerosol acts as condensation nuclei, which stimulate the processes of water evaporation via the radiation effect [28]. Aerosol pollution is a regional phenomenon, and many comprehensive approaches are required to control the anthropogenic sources of aerosols. The aerosol optical depth trends, variability and distribution of aerosols, and their percent changes over the Indo-Gangetic Plain in South Asia have been assessed for various seasons by using the modern retrospective analysis based on satellite measurements and the MERRA-2 model. Satellite data are one of efficient ways to analyze aerosol properties over a larger spatiotemporal region. The aerosol concentration over the IGB is supposed to be different from other Indian regions due to urbanized and populated centers contributing to the aerosol load over the region. The IGB also has fertile land with maximum area of fallow land during pre-monsoon, which contributes to aerosol concentration. There have been plenty of studies to show the presence of aerosol and its formation in this region using satellite data [25]. In particular, [45] highlights inter-annual and seasonal variations of AOD using MISR (the Multi-angle Imaging SpectroRadiometer). However, deep assessment and analysis of AOD at regional and seasonal scales are necessary to better understand regional climate changes. Several previous studies reported an increase in the atmospheric optical depth over various South Asian regions during the summer and winter seasons [2, 36].

This study is based on the investigations [14, 21, 31] using the CALIPSO observations, which is more advantageous than those of the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP). The CALIPSO satellite can easily deliver fundamental information on the vertical distribution of air-born particulates. Various studies on the Himalayan region [22] using aerosol profile data has been published. Several types of aerosol have been estimated by the radiative forcing in [18], whereas the dust burden during the pre-monsoon seasons of 2007 and 2008 has been characterized in [12]. The main objective of the present study is providing a characterization of the vertical variation of aerosol over the region of the Himalayan Plateau based on the CALIPSO observation from 2007 to 2012 and of the vertical distribution of AOD.

## 2. EXPERIMENTAL METHODOLOGY AND DATA ANALYSIS

The online data from the CALIPSO mission were used, especially the CALIOP level 2 Ver. 3.01 data acquired over the IGB region for the current study, which had been obtained within the National Aeronautics and Space Administration (NASA) Earth System Science Pathfinder program [44] in collaboration with Centre National d'Etudes Spatiales, the French space agency, to fill existing gaps in the knowledge on the properties of clouds and aerosols and their global distribution [43]. The CALIPSO was launched with the CNES (France) and NASA (USA) environmental satellites in April 2006 and is now flying as a part of the A-train constellation of satellites. The satellites of the A-train are located in a 705-km sun-synchronous polar orbit giving a 16-day repeat cycle, with an equator-crossing time of about 13:30 and 01:30 local solar time. A primary instrument carried on the CALIPSO satellite is the CALIOP, a near-nadir viewing two wavelengths (i.e., 532 and 1064 nm) polarization lidar, which provides global atmospheric measurements and offers new observational capabilities to the international scientific community. Therefore, the latitudinal variability of the CALIPSO functioning was used to determine the aerosol vertical profile over the Hi-

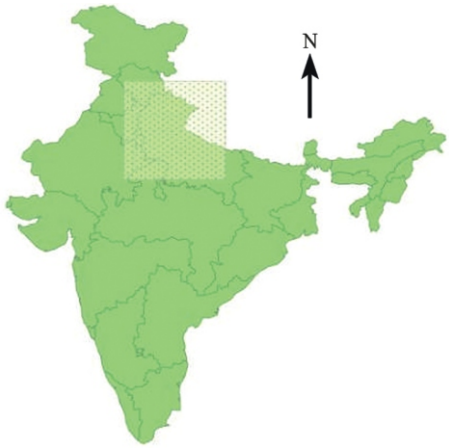



**Fig. 1.** Schematic diagram of the region over the Indo-Gangetic Belt from QGIS.

malayan region and Indo-Gangetic plains. The color ratio (CR), which is used to estimate the size of aerosol particles is determined as follows:

$$\mathrm{CR}(r) \quad {}_{1064,\,\mathrm{p}}(r)/{}_{532,\,\mathrm{p}}(r)$$

where is the backscatter coefficient at two different wavelengths, which characterizes the distance from the satellite to the observation altitude. The subscript "p" represents the atmospheric particles. The parameter for the non-spherical shape of an aerosol particle is called particle depolarization ratio (PDR), which is defined as a ratio of backscatter coefficient at 532 nm of perpendicular and parallel components of the linearly polarized laser pulse [4]:

$$\mathrm{PDR}(r) \quad {}_{532,\,\mathrm{p},\ \perp}(r)/{}_{532,\,\mathrm{p},\,\parallel}(r).$$

Figure 1 shows the area (on the map) where the study was conducted, which is located between 27 –32 N and 77 –81 E. Urbanization and enormous fossil fuel dependency in this region have a negative impact on its climate. The IGB region includes such cities as Kanpur, Lucknow, Allahabad, Varanasi, and Agra, which are extremely polluted due to their industrial enterprises.

## 3. RESULTS AND DISCUSSION

The characteristic variations of aerosol over IGB were analyzed in terms of AOD and the backscatter coefficient, as well as CR and PDR. The correlation showed a good agreement with the data derived out of two independent measurements over a specific location, taking into account the instrument uncertainties.

**Aerosol optical depth.** According to results of the present study, aerosol optical depth directly depends on the total columnar mass of aerosols. Figure 2 represents monthly averaged aerosol optical depth with error bars for four different seasons at 532 and 1064 nm from January 2007 to December 2012. The trend of variation of AOD is nearly the same every year and depends on a season. The maximum value of AOD is during the pre-monsoon (May) season, and in the winter season to the specified period, it is minimum. During this six-year period, the maximum value of AOD was 1.43 at 532 nm and 1.2 at 1064 nm in 2009, while the minimum was 0.41 at 532 nm and 0.38 at 1064 nm in 2009. As aerosol optical depth depends on the total columnar mass of aerosols, aerosol concentration was maximum in June 2009 and minimum in February 2009. This increasing value of averaged AOD during pre-monsoon is caused by dust particles that come from the deserts of northwest India and Pakistan due to westerly winds, and in the monsoon season, its value decreases due to washing out of these pollutants. In addition, westerly winds are restricted to a large extent because of monsoon winds. This variation pattern of AOD is the same for both wavelengths, and AOD values are always lower at higher wavelengths due to the Mie scattering [8]. The results obtained are consistent with the ground-based data presented in [7, 11].

**Backscatter coefficient.** The aerosol concentration over the Indo-Gangetic Belt is represented by longitudinal lines of aerosol backscatter coefficients (532 nm) for the period from January 2007 to December 2012. The maximum frequency of backscatter coefficients that is centered at 0.003–0.009 $\mathrm{km}^{-1}\ \mathrm{sr}^{-1}$

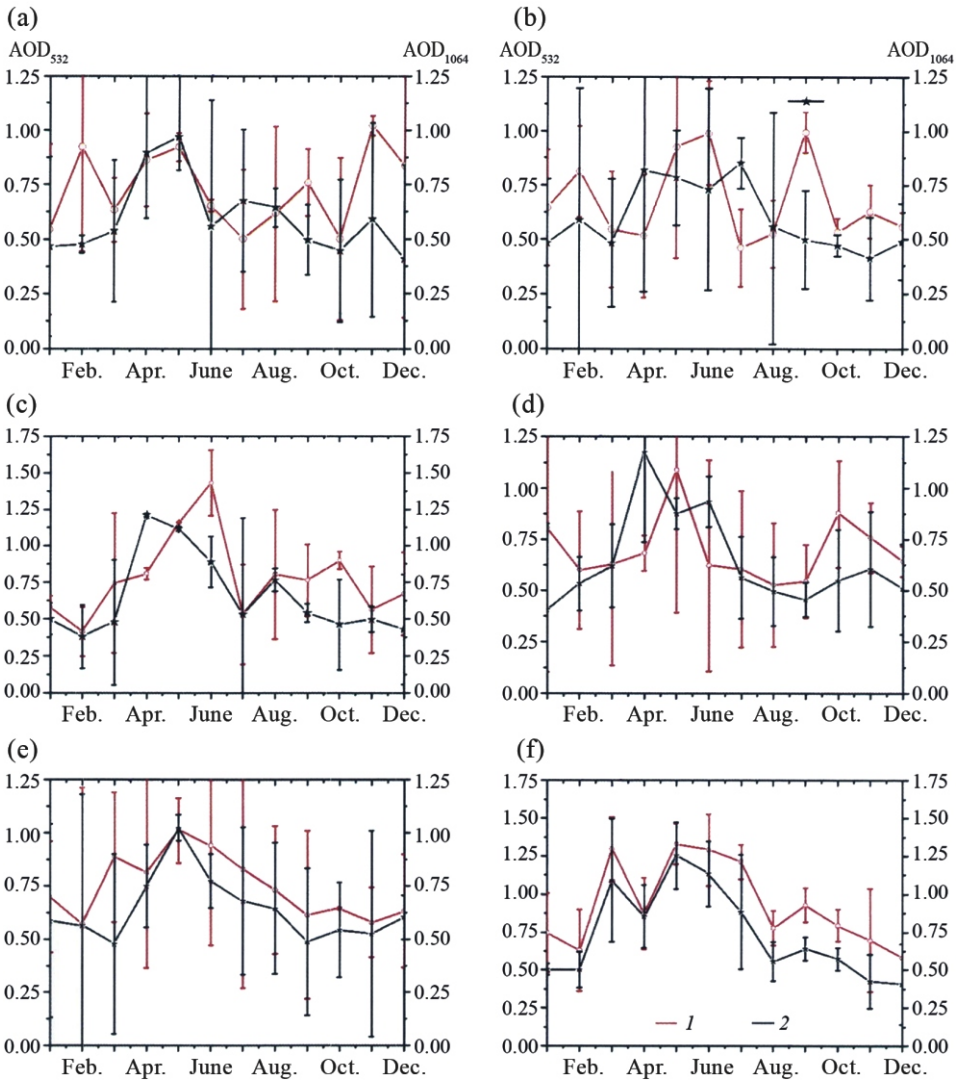



**Fig. 2.** Monthly averaged aerosol optical depth with error bars for four different seasons at (*1*) 532 and (*2*) 1064 nm from 2007 to 2012: (a) 2007, (b) 2008, (c) 2009, (d) 2010, (e) 2011, (f) 2012.

[24] shows a high concentration of different types of aerosols (dust, polluted dust, polluted continental, and smoke). The vertical profile of the aerosol backscatter coefficient in three various regions with its average value $_{532}$ at 532 nm at altitudes of 0.3–1, 1–2, and 2–5 km for six years with error bars of standard deviation has been presented in Fig. 3. The maximum value of $_{532}$ for 0.3–1 km altitude is in December and November (~5.5 $10^{-3}$ km$^{-1}$ sr$^{-1}$), while the second maxima is in May (~5.02 $10^{-3}$ km$^{-1}$ sr$^{-1}$). The results of the present study showed that the value of $_{532}$ continuously decreased from December to March, whereas the maximum peak was attained in May. Furthermore, firstly the peak was observed to increase till May and then it was found to decrease until August and again continuously increase in the post-monsoon period (September, October, and November) until November. In addition, Figure 3 shows the different seasonal patterns of $_{532}$ for heights of 1–2 and 2–5 km. A similar situation was observed in three altitudinal regions from March to July: the curve rises till May and declines after that. For a 1–2 km height from December to March and from July to November, $_{532}$ has a fixed value, which is a little bit higher than the one expected for the duration of post-monsoon period.

**Seasonal variation of CR and PDR.** Figure 4 shows the six-year averaged seasonal variation of color ratio and particle depolarization ratio with an error bar of standard deviation for three different altitude bins (0.3–1, 1–2, and 2–5 km). Figure 4a depicts the same seasonal variations of the monthly averaged color ratio for all the three altitude bins. It can be seen, that it was continuously increasing from December (~0.71)

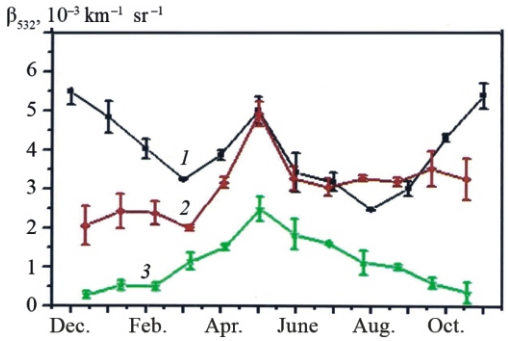



**Fig. 3.** Monthly averaged backscatter coefficient at 532 nm at altitudes of (*1*) 0.3–1 km, (*2*) 1–2 km, and (*3*) 2–5 km.

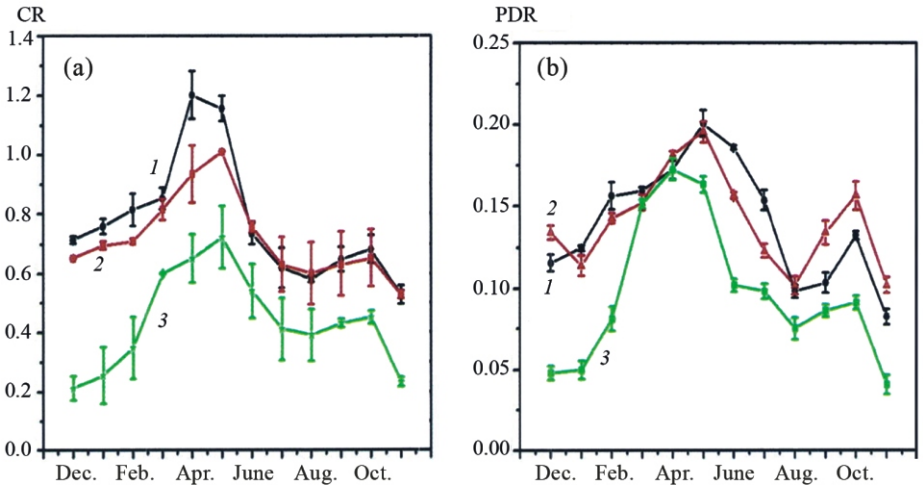



**Fig. 4.** Monthly averaged (a) CR and (b) PDR variation at altitudes of (*1*) 0.3–1 km, (*2*) 1–2 km, and (*3*) 2–5 km.

to March (~0.85) and became maximum in April (~1.2) for 0.3–1 km (the lowest altitude). After the declining phase, it shows an elevation of 0.58 to 0.68 (which is from August to October) and becomes minimum in November (~0.52). In the same way for the other two bins, the averaged CR maximum values (1.0 and 0.72) fall on May. This shows a slightly higher value of CR at 1–2 km altitude as compared to the lowest altitude bin in the monsoon season. Figure 4b represents the monthly averaged seasonal variations of particle depolarization ratio for the three altitude bins. It should be noted that the averaged PDR from August to November shows a higher value for 1–2 km bin than for 1–3 km bin. The maximum level of PDR is in May (~0.2 and ~0.19) for both altitude bins (0.3–1 and 1–2 km), while for the bin of 2–5 km, the maximum value of averaged PDR is in April (~0.17). The averaged value of PDR reflects the same seasonal variations at all the three altitude bins. The obtained values of PDR and CR give information about aerosol type in terms of shape and size. The higher values of CR (more than one) and PDR (~0.2) point to a non-spherical shape of particles and their coarse mode observed by the CALIOP during the pre-monsoon season. The concentration of these non-spherical and coarse particles increases towards 4–5 km altitude, where they can be identified as dust particles that come from north-western India, Pakistan, and desert areas of Afghanistan via long-range transport [6, 37, 38]. The average values equal to 0.5–0.7 (CR) and 0.1–0.16 (PDR) along with the minimum value of  equal to 0.0015–0.0040 km$^{-1}$ sr$^{-1}$ show the cleanness and freshness of the air during the monsoon season. In this season, these pre-existing dust particles in coarse mode are washed out, and new ones are not transported by reduced (almost non-existent) westerlies. The small values of CR and PDR refer to the fine and spherical aerosol particles during the winter and post-monsoon seasons, which are generated by industry or biomass burning [1] at the lower boundary layer of the IGB. In October, the high values of CR and PDR are due to polluted dust particles, which are a combination of dust and industrial emissions [24].

## 4. CONCLUSIONS

The examination of aerosol vertical distribution variations in different seasons was conducted through analysis of the CALIPSO data acquired over the IGB region. The continuous increase of aerosol content before the monsoon season was manifested in the rise of AOD over the IGB associated with the predominant westerly air masses. Due to prevailing westerly air mass, a significant moisture influx and high dust loading are observed over the IGB, which results in a steady and systematic increase of water vapor, CR, PDR, backscatter coefficient, and AOD. Another factor responsible for increasing the values of AOD is the mixed boundary layer formed by differential solar heating over the region. It has been also found that the pollution can easily be accumulated at low altitudes during the late winter and the post-monsoon season. According to this study, the monsoon season is characterized by clean air, which points to washing out of pollutants by rainfall. The value of AOD was the smallest in the winter season every year from 2007 to 2012. Generally, this seasonal variation analysis highlights the need for careful attention to evaluating the spatial and temporal prominent aerosol layers over the IGB to better understand their significance for the climate. Based on the present observations, the authors suggest implementing the national environmental policy across India and curbing the increase in aerosol concentration which leads to pollution of the Indo-Gangetic Plain.

## ACKNOWLEDGMENTS

The first author thanks Dr. Jyotsana Gupta for insightful discussion and valuable suggestions during the paper preparation. The authors also appreciate Ms. Chetna Gupta for proofreading and editing the manuscript.

## FUNDING

This work was supported by ongoing institutional funding. No additional grants to carry out or direct this particular research were obtained.

## CONFLICT OF INTEREST

The authors of this work declare that they have no conflicts of interest.

## REFERENCES

1. K. V. S. Badarinath, T. R. Kiran, and V. K. Prasad, “Agriculture Crop Residue Burning in the Indo-Gangetic Plains a Study Using IRS-6 AWiFS Satellite Data,” Current Science, No. 8, **91** (2006).
2. B. A. Begum and P. K. Hopke, “Ambient Air Quality in Dhaka Bangladesh over Two Decades: Impacts of Policy on Air Quality,” Aerosol Air Qual. Res., **18** (2018).
3. M. Bilal, M. Hassan, D. B. T. Tahir, M. S. Iqbal, and I. Shahid, “Understanding the Role of Atmospheric Circulations and Dispersion of Air Pollution Associated with Extreme Smog Events over South Asian Megacity,” Environ. Monit. Assess., **194** (2022).
4. F. Cairo, G. D. Donfrancesco, A. Adriani, L. Pulvirenti, and F. Fierli, “Comparison of Various Linear Depolarization Parameters Measured by Lidar,” App. Opt., **38** (1999). https://doi.org/10.1364/AO.38.004425
5. S. Dey and S. N. Tripathi, “Estimation of Aerosol Optical Properties and Radiation Effect in the Ganga Basin, Northern India,” Geophys. Res. Lett., **35** (2007), 205808. https://doi.org/10.1029/2007 G2032622
6. S. Dey, S. N. Tripathi, and R. P. Singh, “Influence of Dust Storms on the Aerosol Optical Properties over the Indo-Gangetic Basin,” J. Geophys. Res., **109** (2004), D20211. https://doi.org/10.1029/2004JD004924
7. G. K. Dimitris, R. P. Singh, R. Gautam, M. Sharma, and S. N. Tripathi, “Variability and Trends of Aerosol Properties over Kanpur, Northern India Using AERONET Data (2001–10),” Environ. Res. Lett., **7** (2012). https://doi.org/10.1088/1748-9326/7/2/024003
8. T. F. Eck, B. N. Holben, J. S. Reid, O. Dubovik, A. Smirnov, N. T. O’Neill, I. Slutsker, and S. Kinne, “Wavelength Dependence of the Optical Depth of Biomass Burning Urban and Desert Dust Aerosols,” J. Geophys. Res., **104** (1999), D24. https://doi.org/10.1029/1999JD900923
9. R. Gautam, N. C. Hsu, M. Kafatos, and S. Tsay, “Influences of Winter Haze on Fog/Low Cloud over the Indo-Gangetic Plains,” J. Geophys. Res., **112** (2007), D05207. https://doi.org/10.1029/2005JD007036
10. R. Gautam, N. C. Hsu, and K. M. Lau, “Pre-monsoon Aerosol Characterization and Radiative Effects over the Indo-Gangetic Plains: Implications for Regional Climate Warming,” J. Geophys. Res., **115** (2010), D17208.

11. R. Gautam, N. C. Hsu, S. C. Tsay, K. M. Lau, B. Holben, and S. Bell, "Accumulation of Aerosol over the Indo-Gangetic Plains and Southern Slopes of the Himalayas: Distribution, Properties and Radiative Effects during the 2009 Pre-monsoon Season," Atmos. Chem. Phys., **11** (2011). https://doi.org/10.5194/acp-11-12841-2011
12. R. Gautam, Z. Li, R. P. Singh, and N. Hsu, "Two Contrasting Dust Dominant Periods over India Observed from MODIS and CALIPSO Data," Geophys. Res. Lett., **36** (2009), L06813. https://doi.org/10.1029/2008GL036967
13. M. M. Gogoi, N. B. Lakshmi, V. S. Nair, S. K. Kompalli, K. K. Moorthy, and S. S. Babu, "Seasonal Contrast in the Vertical Profiles of Aerosol Number Concentrations and Size Distributions over India: Implications from RAWEX Aircraft Campaign," J. Earth Syst. Sci., No. 8, **128** (2019). https://doi.org/10.1007/s12040-019-1246-y
14. W. H. Hunt, M. A. Vaughan, K. A. Powell, and C. Weimer, "CALIPSO Lidar Description and Performance Assessment," J. Atmos. Oceanic Technol., No. 7, **26** (2009). https://doi.org/10.1175/2009JTECHA1223.1
15. *IPCC: Fifth Assessment Report: Climate Change* (Cambridge University Press, New York, NY, USA, 2013).
16. H. Jethva, S. K. Satheesh, and J. Srinivasan, "Seasonal Variability of Aerosols over the Indo-Gangetic Plains," J. Geophys. Res., **110D** (2005), 21204. https://doi.org/10.1029/2005JD005938
17. D. G. Kaskaoutis, S. K. Kharol, P. R. Sinha, H. D. Singh, H. D. Kambezidis, A. R. Sharma, and K. V. S. Badarinath, "Extremely Large Anthropogenic-aerosol Contribution to Total Aerosol Load over the Bay of Bengal during Winter Season," Atmos. Chem. Phys., **11** (2011).
18. J. Kuhlmann and J. Quaas, "How Can Aerosols Affect the Asian Summer Monsoon—Assessment during Three Consecutive Pre-monsoon Seasons from CALIPSO Satellite Data," Atmos. Chem. Phys., **10** (2010). https://doi.org/10.5194/acp-10-4673-2010
19. K. M. Lau, M. K. Kim, and K. M. Kim, "Aerosol Induced Anomalies in the Asian Summer Monsoon the Role of the Tibetan Plateau," Climate Dynamics, **26** (2006). https://doi.org/10.1007/s00382-006-0114-z
20. Y. Liu, Y. Zhou, and J. Lu, "Exploring the Relationship between Air Pollution and Meteorological Conditions in China under Environmental Governance," Sci. Rep., **10** (2020).
21. X. Lu, Y. Hu, A. Omar, R. Baize, M. Vaughan, S. Rodier, J. Kar, B. Getzewich, P. Lucker, C. Trepte, C. Hostetler, and D. Winker, "Global Ocean Studies from CALIOP/CALIPSO by Removing Polarization Crosstalk Effects," Remote Sens., No. 14, **13** (2021). https://doi.org/10.3390/rs13142769
22. A. K. Mishra and T. Shibata, "Climatological Aspects of Seasonal Variation of Aerosol Vertical Distribution over Central Indo-Gangetic Belt (IGB) Inferred by the Space-borne Lidar CALIOP," Atmos. Environ., **46** (2012). https://doi.org/10.1016/j.atmosenv.2011.09.052
23. D. Niogi, H. I. Chang, F. Chen, L. Gu, A. Kumar, S. Menon, A. Roger, and R. A. Sr. Pielke, "Potential Impacts of Aerosol–Land–Atmosphere Interactions on the Indian Monsoonal Rainfall Characteristics," Nat. Hazards, **42** (2007).
24. A. H. Omar, M. D. Winker, M. A. Vaughan, C. Kittaka, Z. Liu, C. R. Trepte, R. R. Rogers, R. A. Ferrare, K. Lee, R. E. Kuehn, and C. A. Hostetler, "The CALIPSO Automated Aerosol Classification and Lidar Ratio Selection Algorithm," J. Atmos. and Oceanic Technol., No. 10, **26** (2009). https://doi.org/10.1175/2009JTECHA1231.1
25. K. M. Praseed, T. Nishanth, and M. K. S. Kumar, "Spectral Variations of AOD and Its Validation Using MODIS: First Cut Results from Kannur, India," Atmos. Climate Sci., No. 02(01) (2012). https://doi.org/10.4236/acs.2012.21011
26. V. Ramanathan, C. Chung, D. Kim, T. Bettge, L. Buja, J. T. Kiehl, W. M. Washington, Q. Fu, D. R. Sikka, and M. Wild, "Atmospheric Brown Clouds: Impacts on South Asian Climate and Hydrological Cycle," Proc. Nat. Acad. Sci. USA, **102** (2005). https://doi.org/10.1073/pnas.0500656102
27. V. Ramanathan, P. J. Crutzen, J. T. Kieh, and D. Rosenfeld, "Aerosols, Climate, and the Hydrological Cycle," Science, **294** (2001). https://doi.org10.1126/science.1064034
28. V. Ramanathan, F. Li, M. V. Ramana, P. S. Praveen, D. Kim, C. E. Corrigan, H. Nguyen, E. A. Stone, J. J. Schauer, G. R. Carmichael, B. Adhikary, and S. C. Yoon, "Atmospheric Brown Clouds: Hemispherical and Regional Variations in Long-range Transport, Absorption, and Radiative Forcing," J. Geophys. Res., **112** (2007), D22S21. https://doi.org/10.1029/2006JD008124
29. A. U. Rehman, F. Chishtie, W. A. Qazi, S. Ghuffar, I. Shahid, and K. Fatima, "Evaluation of Three-Hourly TMPA Rainfall Products Using Telemetric Rain Gauge Observations at Lai Nullah Basin in Islamabad, Pakistan," Remote Sens., **10** (2018).
30. S. K. Satheesh, K. Krishna Moorthy, S. Suresh Babu, V. Vinoj, V. S. Nair, S. Naseema Beegum, D. P. Dutt Alappattu, and P. K. Kunhikrishnan, "Vertical Structure and Horizontal Gradients of Aerosol Extinction Coefficients over Coastal India Inferred from Air Borne Lidar Measurements during the Integrated Campaign for Aerosol, Gases and Radiation Budget (ICARB) Field Campaign," J. Geophys. Res., **114D** (2009), 05204. https://doi.org/10.1029/2008JD011033
31. G. Seze, J. Pelon, M. Derrien, H. le Gleau, and B. Six, "Evaluation against CALIPSO Lidar Observations of the Multi-geostationary Cloud Cover and Type Dataset Assembled in the Framework of the Megha-Tropiques Mission," Quart. J. Roy. Meteorol. Soc., **141 (688)** (2015). https://doi.org/10.1002/qj.2392
32. M. Z. Shahid, F. Chishtie, M. Bilal, and I. Shahid, "WRF-Chem Simulation for Modeling Seasonal Variations and Distributions of Aerosol Pollutants over the Middle East," Remote Sens., **13** (2021).

33. M. Z. Shahid, H. Liao, J. Li, I. Shahid, A. Lodhi, and M. Mansha, "Seasonal Variations of Aerosols in Pakistan: Contributions of Domestic Anthropogenic Emissions and Transboundary Transport," Aerosol and Air Quality Res., **15** (2015).
34. M. Z. Shahid, I. Shahid, F. Chishtie, M. I. Shahzad, and G. Bulbul, **"**Analysis of a Dense Haze Event over North-eastern Pakistan Using WRF-Chem Model and Remote Sensing," J. Atmos. Sol.-Terr. Phys., **182** (2019).
35. M. Sharma, D. G. Kaskaoutis, R. P. Singh, and S. Singh, "Seasonal Variability of Atmospheric Aerosol Parameters over Greater Noida Using Ground Sunphotometer Observations," Aerosol and Air Quality Res., **14** (2014).
36. S. Shi, T. Cheng, X. Gu, H. Guo, Y. Wu, Y. Wang, F. Bao, and X. Zuo, "Probing the Dynamic Characteristics of Aerosol Originating from South Asia Biomass Burning Using POLDER/GRASP Satellite Data with Relevant Accessory Technique Design," Environ. Int., **145** (2020), 106097.
37. R. P. Singh, S. Dey, S. N. Tripathi, V. Tare, and B. Holben, "Variability of Aerosol Parameters over Kanpur, Northern India," J. Geophys. Res., **109** (2004), D23206. https://doi.org/10.1029/2004JD004966
38. S. Singh, S. Nath, R. Kohli, and R. Singh, "Aerosols over Delhi during Pre-monsoon Months: Characterstics and Effects on Surface Radiation Forcing," Geophys. Res. Lett., **32** (2005), L13808. https://doi.org/10.1029/2005GL023062
39. A. K. Srivastava, V. Yadav, V. Pathak, S. Singh, S. Tiwari, D. S. Bish, and P. Goloub, "Variability in Radiative Properties of Major Aerosol Types: A Year-long Study over Delhi an Urban Station in Indo-Gangetic Basin," Sci. Total Environ., **473–474** (2014).
40. S. Tiwari, A. K. Srivastava, A. K. Singh, and S. Singh, "Identification Q1 of Aerosol Types over Indo-Gangetic Basin: Implications to Optical Properties and Associated Radiative Forcing," Environ. Sci. and Pollut. Res., No. 16, **22** (2015).
41. S. N. Tripathi, V. Tare, N. Chinnam, and A. K. Srivastav, et al., "Measurements of Atmospheric Parameters during Indian Space Research Organization Geosphere Biosphere Program Land Campaign II at a Typical Location in the Ganga Basin: Physical and Optical Properties, J. Geophys. Res., **111D** (2006). https://doi.org/10.1029/2006JD007279
42. S. Wagh, P. Singh, S. D. Ghude, P. Safai, T. Prabhakaran, P. P. Kumar, "Study of Ice Nucleating Particles in Fog-haze Weather at New Delhi, India: A Case of Polluted Environment," Atmos. Res., **259**, 105693.
43. D. M. Winker, J. R. Pelon, and M. P. McCormick, "The CALIPSO Mission: Space Borne Lidar for Observation of Aerosols and Clouds, SPIE 4893, Lidar Remote Sensing for Industry and Environment Monitoring, (2003). https://doi.org/10.1117/12.4665
44. D. M. Winker, M. A. Vaughan, Ali Omar, Y. Hu, and K. A. Powell, "Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms," J. Atmos. and Oceanic Technol., **26** (2009). https://doi.org/10.1175/2009JTECHA1281.1
45. X. Xia, P. Wang, Y. Wang, Z. Li, J. Xin, J. Liu, and H. Chen, "Aerosol Optical Depth over the Tibetan Plateau and Its Relation to Aerosols over the Taklimakan Desert," Geophys. Res. Lett., **35** (2008), L16804. https://doi.org/10.1029/2008GL034981
46. S. U. Zaman, M. R. S. Pavel, K. S. Joy, F. Jeba, M. S. Islam, S. Paul, M. A. Bari, and A. Salam, "A Spatial and Temporal Variation of Aerosol Optical Depths over Six Major Cities in Bangladesh," Atmos. Res., **262** (2021), 105803.

**Publisher's Note.** Allerton Press remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

AI tools may have been used in the translation or editing of this article.

2024 8th International Conference on Computational System and Information Technology for Sustainable Solutions (CSITSS) | 979-8-3315-0546-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/CSITSS64042.2024.10816775

# Signal Noise Reduction Algorithm for Object Localization and Tracking using a 2-Dimensional Ultrasonic Sensor Array

Sambhav Jain
*Dept. of Mechanical Engineering*
*Delhi Technological University*
Delhi 110042, India
sambhavjain_pe22b14_41@dtu.ac.in

Kuldeep Singh
*Combat Vehicles Research and Development Establishment*
*Defence Research and Development Organization*
Chennai 60054, India
kuldeep.cvrde@gov.in

***Abstract*— Ultrasonic sensors return time of flight (TOF) readings; by producing inaudible sound waves and calculating the time they take to return to the sensors. When arranged in an array (containing multiple sensors) they can have various applications, both independently and in conjunction with other transducers and cameras. One of the most common applications that has been developed in the last couple of decades has been object detection, localization and tracking. This research paper presents a novel algorithm that addresses one of the biggest challenges with using ultrasonic sensor arrays for object localization – signal noise. The proposed system also includes a laser module mounted on a pan-tilt motor mechanism to track objects, utilizing the signal noise reduction and angles calculation algorithm. The improvement in signal quality has been quantified with signal-to-noise ratio (SNR) and mean squared error (MSE). The presented algorithm will significantly improve the application of ultrasonic sensors, in autonomous vehicles, robotics and industrial automation.**

***Keywords*— Noise Reduction Algorithm, Pan-Tilt Mechanism, Sensor Array, Signal Noise Reduction, Ultrasonic Sensors***

## I. Introduction

Ultrasonic sensors are one of the cheapest and most commonly used transducers in both academia and industry. They have been widely used to detect, classify, target and mitigate objects. Ultrasonic sensors operate using sound waves [1]. They emit a burst of sound waves at a frequency higher than what humans can hear. These sound waves reflect off the nearest object and return to the sensor's receiver. The sensor then measures the time of flight (TOF), which is the time interval between when the sound wave is emitted and when it is received by the sensor [2][3].

A 2-Dimensional sensor array is a spatial arrangement of sensors used to collect data from multiple nodes (a node is a single sensor) at the same time. The data collected from the array of sensors can be put to multiple uses including localizing objects for tracking. Triangulation has been a popular approach for the same, that involves using a minimum of two sensors [4], [5], [6]. This paper presents an alternative approach that utilizes a novel signal noise reduction and angle calculation algorithm for controlling pan-tilt motor (PTM) mechanism to track objects with a laser module. A PTM mechanism is a combination of two motors which allows rotation in both horizontal and vertical direction, similar to a turret. A twin servo-based assembly employing two SG90 motors was used to construct the pan-tilt mechanism [7], [8].

Noise in sensor readings can occur due to various factors like cross talk between sensors, failure of complete wave reflection around the edges of an object and sensing other objects in vicinity. A great deal of noise is observed especially in the case of non-stationary objects. A method to reduce noise that uses breadth first search (BFS) algorithm and statistics has

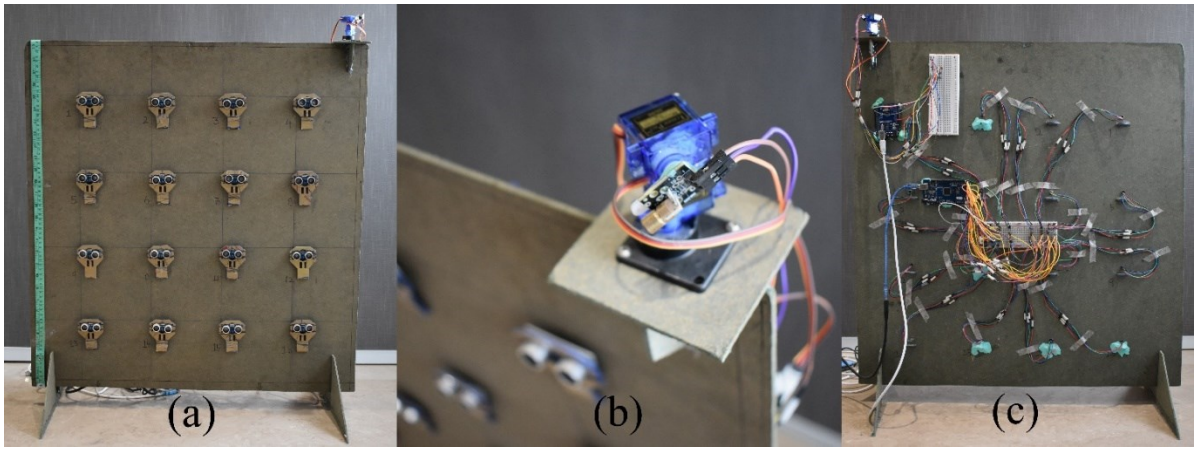



Fig. 1 (a) Front view of sensor array board (b) Pan-Tilt mechanism with laser module (c) Back view of sensor array board

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 06,2025 at 10:51:50 UTC from IEEE Xplore. Restrictions apply.

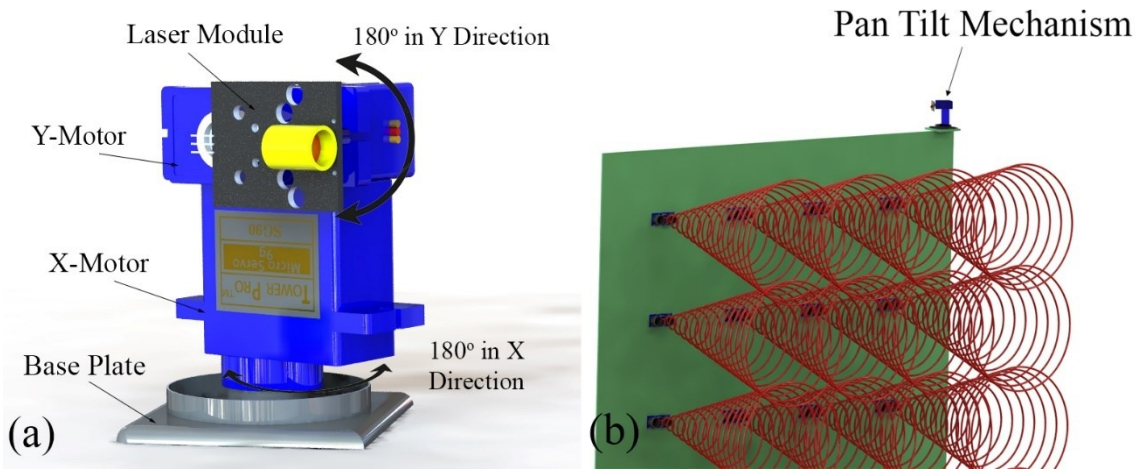



Fig. 2 (a) Pan-Tilt servo motor mechanism (b) Placement of Pan-Tilt motor mechanism on the top-right of the sensor array board

been used before calculating the angles at which the object lies in front of the sensor array board. This is a preliminary approach that is relatively computationally inexpensive compared to some other alternatives like the Kalman filter and the extended Kalman filter which have been previously demonstrated to reduce noise in sensor readings [9], [10], [11]

The presented algorithm significantly solidifies the applicability of ultrasonic sensors in cases where object localization has proven to be difficult due to signal noise. Aid technologies for the blind [12], [13], [14], parking assistance [15], autonomous robots and drones [16] are a few of the many applications of ultrasonic sensors where the performance accuracy of these sensors can be improved by implementing the general methods and algorithms discussed in this paper.

## II. Methods And Materials

### A. Sensor Array

HC-SR04 ultrasonic sensors are spatially arranged in a square to form a 4x4 sensor array network. A total of 16 sensors have been utilized to make the sensor array, which collect the TOF data simultaneously for an object present in front of them. The sensor array has been mounted on an MDF board of dimensions 750 mm × 700 mm × 4 mm. The sensor array board stands on two trapeziums made of the same MDF sheet. The complete sensor array is shown in Fig. 1 (a). Each sensor consumes around 15 mA of current (Hence, a total of 240 mA) and require 5 V to operate. The sensors are ready to use and do not require any initialization, they output TOF data which can be converted to distance value simultaneously.

### B. Pan-Tilt Servo Motor Mechanism

A turret-like mechanism which can rotate in both X and Y directions was required to mount the laser. Two SG90 servo motors were used. A base plate with appropriate gear teeth carvings was selected to couple the driving gear of the motor rotating in X direction. The plain side of the X-motor at the top was attached with the Y-motor with an adhesive. The driving gear of the Y-motor was attached to the laser module. Instead of using a commercial pan and tilt bracket for the SG90 motors, the motors were mounted directly on top of each other. The PTM mechanism is shown in Fig. 1 (b). SG90 is a micro servo motor that can rotate up to a maximum angle of 180 degrees. The Pan-Tilt Mechanism (PTM) was attached at the top of the rectangular sensor array base, allowing the laser to point at any object that lies in front of the array board. The construction of the PTM is shown in Fig. 2 (a) and its placement on the sensor array is shown in Fig. 2 (b). The motors require the most amount of power to function. They can consume a maximum of 500 mA current [17] (hence, a maximum of 1 A current for two servos).

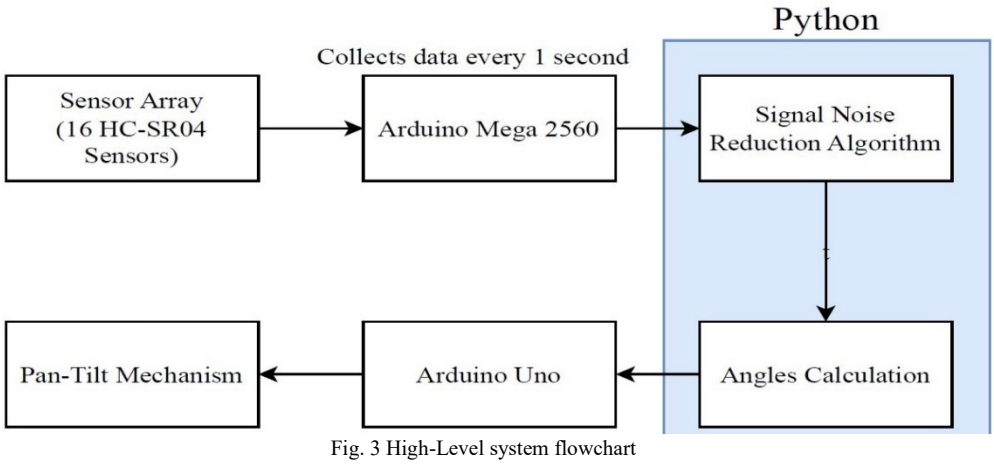



Fig. 3 High-Level system flowchart

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 06,2025 at 10:51:50 UTC from IEEE Xplore. Restrictions apply.

The maximum current consumed by the laser module is around 40 mA, and practically it consumes no current when idle.

### C. Control System

A control system is required to collect TOF data from the sensor array, process the data using the noise reduction and angles calculation algorithm, and finally control the PTM mechanism to point at the object lying in front of the sensor array board.

Arduino is the microcontroller used to conduct this study due to its compatibility with the motors, sensors and the laser module involved. A single Arduino would not have sufficed to control both the mechanisms due to power output considerations. Hence, two Arduinos were used to create a distributed control system. The complete control system is shown in Fig. 3.

An Arduino Mega 2560, with its 54 digital input/output pins was well-suited to interface with 16 ultrasonic sensors, each requiring two connections (Trig and Echo). It was deployed to gather data from the sensor array at 1-second intervals. The calculation of distance using Time-of-Flight (TOF) was performed on the Arduino Mega before the values were transmitted further. The Mega consumes around 40 mA of current, and requires an input voltage in the range of 7 -12 V to provide an output voltage of 5 V.

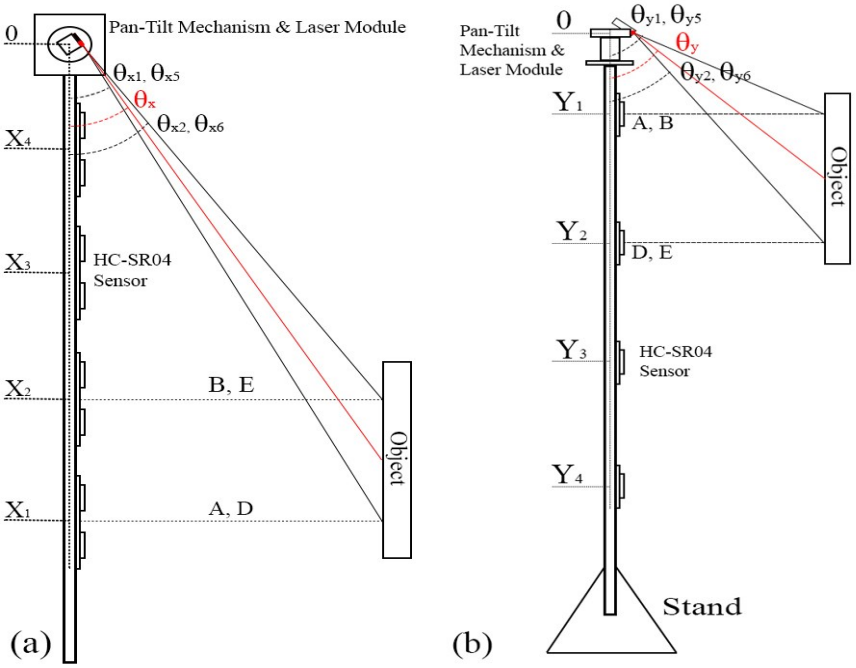



Fig. 4 (a) Top view of the sensor array for $\theta_X$ calculation (b) side view of the sensor array for $\theta_Y$ calculation

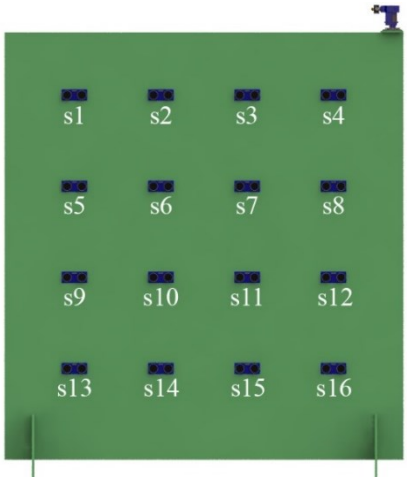



Fig. 5 Ultrasonic Sensor nodes naming

Connected to Python running on a Windows operating system, the Arduino Mega facilitated data collection through serial communication. This communication was facilitated by the PySerial library in Python. Upon receiving the 16 TOF data values in the Python interpreter, the noise reduction algorithm was applied to minimize signal noise, resulting in more accurate angle calculations. Two angles $\theta_X$ and $\theta_Y$ were calculated using matrix operations and trigonometry based on the clean set of values. At this point, Python exported these angles to the Arduino Uno using serial communication.

Arduino Uno used these angles to control the PTM to target at the threatful object in front of it. The laser module mounted on the PTM was also connected to the Uno and was always set to HIGH.

Once the PTM was targeting the object, the system terminated and waited for the next set of values. In case of stationary objects, the PTM remained stationary once it had started pointing at the object. In case the object was non-stationary (assuming it remained in front of the array), the PTM started tracking the object, rotating every 1 second.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 06,2025 at 10:51:50 UTC from IEEE Xplore. Restrictions apply.

The advantage of using this method was the reduction in the time it took for the sensor array readings to be recorded and for the PTM to point at the object. If only one Arduino were used, the next set of sensor readings would have been collected only after the current set had been fully processed and the loop had completed. By employing two Arduinos, the data collection from the sensor array by the Mega was decoupled from the processing and PTM operations. This allowed the incoming data to be exported to Python independent of the loop's completion, thus improving efficiency and making the system more responsive. The control system wiring is shown in Fig. 1 (c).

## III. Signal Noise Reduction And Angle Calculation Algorithm

As discussed previously, the HC-SR04 sensor is prone to a lot of noise. In 2018, Li et al. [10] presented a study using the Kalman filter to reduce noise in the ultrasonic sensor array's readings. The Kalman filter is a recursive algorithm that combines the noisy measurement with a prediction of the state. Hence, an estimate of where the object may lie is created. The prediction is refined with each new set of readings [11]. The Kalman filter assumes that the system dynamics are linear in nature. The process and measurement noise are assumed to be normally distributed. Some domain knowledge is also required to set its parameters. Apart from that, when scaled to larger models, the process of matrix inversion becomes computationally expensive.

Considering these limitations, a different approach to reduce noise in sensor array's readings is suggested in this paper. The method has primarily 2 steps:

- Creation of Proximity Clusters: Using Breadth first search (BFS) algorithm, clusters of distance measurements with respect to the physical proximity of sensors to each other were created.
- Range Constraint: The clusters created were then constrained by taking a multiple of the standard deviation around the mean of the readings.

BFS is a graph traversal algorithm. To analyze the distance measurements from the sensor array, BFS is used to systematically search the nodes and edges in the graph representing the sensor network. Naming of all sensor nodes is shown in Fig. 5. It begins its search from a starting node or root. The algorithm then considers all the adjacent nodes separated by an edge. If it finds a value greater than 0, it is considered a part of a cluster. Once all the adjacent nodes of the root node have been considered, the algorithm proceeds to consider the adjacent nodes of those newly discovered nodes. Consider a set of initial readings from the array to be I, as input to the BFS algorithm.

$$I = \begin{bmatrix} A & B & C & 0 \\ D & E & 0 & 0 \\ 0 & 0 & 0 & F \\ 0 & 0 & 0 & G \end{bmatrix} \quad (1)$$

Here, $\{A, B, C, D, E, F, G\} \geq 0$

The object is most likely in front of the sensors forming the biggest cluster. Two clusters will be created in the example 4x4 array I, {A, B, C, D, E} and {F, G}. Among these two, length of the first cluster is greater than that of the second cluster, therefore readings of the second cluster are considered to be noise. Hence, F and G are set to 0.

$$I_1 = \begin{bmatrix} A & B & C & 0 \\ D & E & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (2)$$

There can be more than two clusters. Only the cluster with the greatest length is permitted to be processed further and the rest of the matrix is set to 0. In case there are two or more clusters with equal length, only the cluster encompassing the maximum number of central values i.e. s6, s7, s10 and s11 is permitted to move to the next step.

It is highly likely that the sensors report values when the object is not directly in front of them due to the sound waves being propagated in a conical manner. As mentioned before, the sensors in front of the edge of an object often tend to misbehave, giving values much greater or smaller than their adjacent sensors. Hence a constraint on the range of the values is implemented. The mean (μ) and standard deviation (σ) of the non-zero values in array $I_1$ are calculated. The range is defined as:

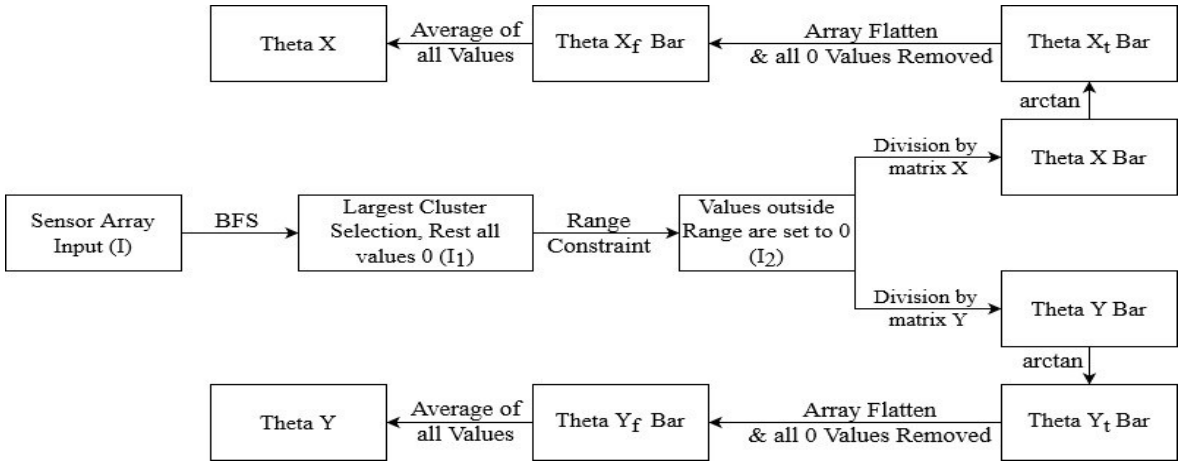



Fig. 6 Noise reduction and angles calculation algorithm flowchart

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 06,2025 at 10:51:50 UTC from IEEE Xplore. Restrictions apply.

$$Range = \{(\mu - n\sigma, \mu + n\sigma): \mu, \sigma, n \in Q\} \quad (3)$$

Here n is a hyperparameter that can be fine-tuned. In the study conducted in this paper, the value of n is set to 1. Any non-zero value lying outside the range will be set to 0. Assuming that, $C > \mu + n\sigma$ or $C < \mu - n\sigma$; Hence, C will be set to 0.

$$I_2 = \begin{bmatrix} A & B & 0 & 0 \\ D & E & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (4)$$

Array $I_2$ is noise free and the angles $\theta_X$ and $\theta_Y$ are calculated from this array. Values {A, B, D, E} are the distance of the object in front of the sensors. These values are approximately the same. The PTM is placed on the top right corner of the array. The angles $\theta_X$ and $\theta_Y$ are values in degrees on which the X-Motor and Y-Motor rotate respectively in order to target the object in front of them with the laser module. Matrix $X$ and $Y$ consists of distances between the sensor and the PTM in X and Y axes. Each row in the X matrix contains a single, distinct value being repeated across the row. Similarly, each column in the Y matrix contains a single, distinct value. Both $X$ and $Y$ contain constant values in centimeters. Two new matrices $\overline{\theta_X}$ and $\overline{\theta_Y}$ are calculated by dividing $I_2$ with $X$ and $Y$ respectively.

**Algorithm: Angle Calculation From Cleaned Array**

**Input:** Noise-free array $I_2$

**Output:** Angles $\theta_X$ and $\theta_Y$

$\overline{\theta_{X[m\times n]}} \leftarrow \frac{I_{2[m\times n]}}{Y_{[m\times n]}}$ // m and n are rows and columns

$\overline{\theta_{Y[m\times n]}} \leftarrow \frac{I_{2[m\times n]}}{Y_{[m\times n]}}$

$\overline{\theta_{Xt}} \leftarrow$ arctan of $\overline{\theta_X}$ // to calculate angle of rotation

$\overline{\theta_{Yt}} \leftarrow$ arctan of $\overline{\theta_Y}$

$\overline{\theta_{Xf}} \leftarrow$ Non-zero values of $\overline{\theta_X}$ flattened to 1-D

$\overline{\theta_{Yf}} \leftarrow$ Non-zero values of $\overline{\theta_Y}$ flattened to 1-D

$\theta_X \leftarrow average(\overline{\theta_{Xf}})$

$\theta_Y \leftarrow average(\overline{\theta_{Yf}})$

**End**

$\theta_X$ and $\theta_Y$ are the final angles as shown in Fig. 4 (a) and Fig. 4 (b). Noise Reduction and Angle Calculation algorithm has been depicted in the Fig. 6. After both the angles are calculated as shown in, the angles are finally actuated by the PTM.

## IV. Results

The objects that were tracked to test the working of the system was done across three shapes – circle, triangle and rectangle. These objects were cut out from thin cardboard sheets. There were three circles in the range of 20 – 40 cm of diameter, four triangles having side lengths in 15 – 40 cm range, and four rectangles having side lengths in 10 – 35 cm range.

The sensors and actuators can draw a maximum total of 1.28 A of current from both the microcontrollers. The Arduino boards can themselves consume 40 – 80 mA of current when idle and not supplying any power to the circuit components.

The experimental setup is highly portable overall due to the detachable trapezium-shaped legs and a weight of approximately 2.1 kg.

The ultrasonic sound waves tend to get absorbed by the object they are trying to reflect from. The degree of reflectivity of the sound waves depend on the material. As cardboard sheets were used to form the objects for tracking, the HC-SR04 sensor starts providing steady distance values after 15 cm [18]. Also, the sensor works accurately till a distance of 70 cm [19]. These results were also established during the experimental work for this study. Thus, the most accurate tracking performance is observed in this range, primarily due to hardware limitation.

The environment in which the testing of algorithm took place was indoors. In highly dynamic environments, it is likely that the tracking would be less accurate, but it could be improved by having a larger sensor array. More sensors would rule out smaller objects better and provide more accurate data for estimation of larger objects. There is a limitation to the size of the objects that the sensor array could gauge. The object cannot be bigger than the array itself in the defined distance range. A larger than array object can be detected if more accurate ultrasonic sensors with respect to distance are used.

The standard metric to deduce the effectiveness of the noise reduced in a signal is signal to noise ratio (SNR), which is calculated in decibels.

$$SNR(dB) = 10\log_{10}\left(\frac{Power\ of\ Signal}{Power\ of\ Noise}\right) \quad (5)$$

A dataset of 100 rows was created from three different object-sensor array orientations, containing distance values. This data was then processed with the noise reduction algorithm creating a new dataset with reduced noise. A reference array is a set of values which are considered to be ideal and true. It is the signal which would be measured in a completely noise-free environment. Three such true arrays were measured manually from the sensors in the front of the object, one for each shape. These three are considered to be the base value for both the noisy readings and the cleaned readings from the algorithm. Arithmetic mean of the square of each value in a single row is the power. SNR was calculated for both the noisy values and the cleaned values from the sensor against the three true reference arrays. 79 out of the 100 SNR values for the noisy readings were less than 0. This means that there was more noise than signal in those 79 sets of distance readings. The average SNR value for 100 readings in the noisy dataset was -5.73 dB. For the cleaned data, only 18 sets of distance values were negative. The average *SNR* value was 9.59 dB. The improvement in the SNR can be said to be the difference of these two values. $\Delta SNR = 15.32\ dB$. Hence, the signal was improved by a SNR of 15.32 dB on average.

Mean squared error (MSE) is another important metric that will quantify the improvement in signal quality. MSE is the average squared difference between two values. In this case, MSE will be the clean and noisy data set with their corresponding true value arrays. The signal noise reduction algorithm significantly reduced the average MSE from noisy to clean data. Average MSE for noise data was 387.35 and MSE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 06,2025 at 10:51:50 UTC from IEEE Xplore. Restrictions apply.

for clean data was 50.31. This resulted in an overall reduction of 87% of noise from the signal. The majority of sensors showed substantial improvement after data cleaning. Notably, sensors s9 and s16 achieved perfect error reduction, with MSE decreasing to zero. Other sensors with significant improvements include s4 (99.36% reduction), s8 (97.89% reduction), and s13 (92.34% reduction).

Some sensors (s1, s6, s7) showed an increase in MSE after cleaning. This unexpected result suggests that the cleaning algorithm may need further refinement for sensors present at a certain location in the sensor array in respect to other sensors.

The cleaning process successfully reduced the maximum observed MSE from 4884.61 to 2116.00, demonstrating its effectiveness in mitigating extreme errors. The minimum MSE remained at 0.00 for both noisy and clean data, indicating that perfect readings were possible in both datasets.

Kalman filter was also applied on the noisy data to compare the SNR values with the algorithm presented in this paper. The Kalman-filtered output, optimized with hyperparameters Q=0.01 (Process Noise Covariance) and R=10 (Measurement Noise Covariance), refined these readings, leading to an average improvement of $\Delta SNR = 8.04\ dB$. The true sensor readings, based on manual measurements were considered for comparison, providing a benchmark to assess the accuracy and effectiveness of the noise reduction technique.

## V. Conclusion

A novel method to reduce signal noise collected from ultrasonic sensors arranged in a 2-Dimensional square array was introduced in this paper. The sensor array was integrated with a pan-tilt motor mechanism used to target the localized object with a laser module. An open-loop control system was used to control the complete mechatronic system. The noise reduction algorithm is capable to reduce the noise in signal by an average of SNR 15.32 dB and reduce MSE by 87% also at an average. The SNR improvement shown by this algorithm outperformed Kalman filter. This algorithm can be extensively used in object localization from TOF data collected from any transducer array. While the cleaning process significantly improved overall data quality and reduced errors for most sensors, the varied results across different sensors highlight the complexity of ultrasonic sensor data cleaning. The algorithm can have many applications, including but not limited to – military, blind aid, automotive parking assistance, drones and industrial robots.

## References


[1] M. Javaid, A. Haleem, S. Rab, R. P. Singh, and R. Suman, "Sensors for daily life: A review," 2021, doi: 10.1016/j.sintl.2021.100121.

[2] V. Parihar, Y. Rohilla, and K. Kumari, "Ultrasonic Sensor based Smart Cap as Electronic Travel Aid for Blind People," in *Proceedings of the Third International Conference on Smart Systems and Inventive Technology (ICSSIT 2020)*, 2020.

[3] K. A. Raza and W. Monnet, *Moving objects detection and direction-finding with HC-SR04 ultrasonic linear array*. 2019.

[4] K. Blümel, F. Tagliaferri, and M. Kuhl, "Algorithm for calculating distance and sensor-object angle from raw data of ultra-low power, long-range ultrasonic time-of-flight range sensors," *Procedia CIRP*, vol. 118, pp. 1061–1065, Jan. 2023, doi: 10.1016/J.PROCIR.2023.06.182.

[5] L. Chassagne, O. Bruneau, A. Bialek, C. Falguière, E. Broussard, and O. Barrois, "Ultrasonic Sensor Triangulation for Accurate 3D Relative Positioning of Humanoid Robot Feet," *IEEE Sens J*, vol. 15, no. 5, pp. 2856–2865, May 2015, doi: 10.1109/JSEN.2014.2382435.

[6] O. Tekdas and V. Isler, "Sensor placement for triangulation-based localization," *IEEE Transactions on Automation Science and Engineering*, vol. 7, no. 3, pp. 681–685, Jul. 2010, doi: 10.1109/TASE.2009.2037135.

[7] A. Bist and C. Singhal, "Efficient immersive surveillance of inaccessible regions using UAV network," *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications Workshops, INFOCOM WKSHPS 2021*, May 2021, doi: 10.1109/INFOCOMWKSHPS51825.2021.9484539.

[8] N. Ammar, "Design of Real-Time Vision-Based Face Verification and Tracking Via Pan Tilt Camera," *2023 3rd International Conference on Computing and Information Technology, ICCIT 2023*, pp. 314–318, 2023, doi: 10.1109/ICCIT58132.2023.10273925.

[9] A. Rafi, A. Tahtawi, and K. Filter, "Kalman Filter Algorithm Design for HC-SR04 Ultrasonic Sensor Data Acquisition System," *IJITEE (International Journal of Information Technology and Electrical Engineering)*, vol. 2, no. 1, pp. 15–19, Jul. 2018, doi: 10.22146/IJITEE.36646.

[10] S. E. Li *et al.*, "Kalman filter-based tracking of moving objects using linear ultrasonic sensor array for road vehicles," *Mech Syst Signal Process*, vol. 98, pp. 173–189, Jan. 2018, doi: 10.1016/j.ymssp.2017.04.041.

[11] C. Urrea and R. Agramonte, "Kalman Filter: Historical Overview and Review of Its Use in Robotics 60 Years after Its Creation," *J Sens*, vol. 2021, 2021, doi: 10.1155/2021/9674015.

[12] S. Karmakar, S. K. Kundu, S. Sengupta, G. Priya, and A. Roy, "Third Eye: Obstacle Avoidance Device for Visually Impaired Persons," *Interdisciplinary Research in Technology and Management*, pp. 619–626, May 2024, doi: 10.1201/9781003430469-73.

[13] K. Goel, M. Nagar, M. Tiwari, A. R. Mishra, and S. S. Chauhan, "Enhanced Ultrasonic Cane for Visually Impaired People," *2024 2nd International Conference on Disruptive Technologies, ICDT 2024*, pp. 500–505, 2024, doi: 10.1109/ICDT61202.2024.10489694.

[14] C. Ramisetti *et al.*, "An Ultrasonic Sensor-based blind stick analysis with instant accident alert for Blind People," *2022 International Conference on Computer Communication and Informatics, ICCCI 2022*, 2022, doi: 10.1109/ICCCI54379.2022.9740786.

[15] T. F. Wu, P. S. Tsai, N. T. Hu, and J. Y. Chen, "Research and implementation of auto parking system based on ultrasonic sensors," *Proceedings of the IEEE International Conference on Advanced Materials for Science and Engineering: Innovation, Science and Engineering, IEEE-ICAMSE 2016*, pp. 643–645, Feb. 2017, doi: 10.1109/ICAMSE.2016.7840267.

[16] J. N. Yasin, S. A. S. Mohamed, M. H. Haghbayan, J. Heikkonen, H. Tenhunen, and J. Plosila, "Low-cost ultrasonic based object detection and collision avoidance method for autonomous robots," *International Journal of Information Technology (Singapore)*, vol. 13, no. 1, pp. 97–107, Feb. 2021, doi: 10.1007/S41870-020-00513-W/FIGURES/13.

[17] D. M. Harris and S. L. Harris, "Memory and I/O Systems," *Digital Design and Computer Architecture*, pp. 474–580, 2013, doi: 10.1016/B978-0-12-394424-5.00008-2.

[18] S. Adarsh, S. M. Kaleemuddin, D. Bose, and K. I. Ramachandran, "Performance comparison of Infrared and Ultrasonic sensors for obstacles of different materials in vehicle/ robot navigation applications," *IOP Conf Ser Mater Sci Eng*, vol. 149, no. 1, p. 012141, Sep. 2016, doi: 10.1088/1757-899X/149/1/012141.

[19] S. Mexias, V. Fotopoulos, and K. Giannakopoulos, "Comparison of HC-SR04 and TF-LC02 Distance Sensors for Indoor Mapping Applications," in *2024 Panhellenic Conference on Electronics and Telecommunications, PACET 2024 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/PACET60398.2024.10497028.


Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 06,2025 at 10:51:50 UTC from IEEE Xplore. Restrictions apply.

AIP Conference Proceedings

RESEARCH ARTICLE | JANUARY 16 2025

# Small kernelled deep capsule networks for handwritten Ge'ez numerals recognition

Ruchika Malhotra; Maru Tesfaye Addis

*AIP Conf. Proc.* 3233, 020014 (2025)
https://doi.org/10.1063/5.0234108

View Online Export Citation

Articles You May Be Interested In

An efficient method to enhance the readability of historical handwritten artifacts

*AIP Conference Proceedings* (March 2016)

Horizontal feature extraction of handwritten character by statistical and structural sorts compared with diagonal direction

*AIP Conf. Proc.* (May 2024)

Artificial intelligence based handwritten text recognition system

*AIP Conference Proceedings* (May 2022)

27 January 2025 05:55:07

AIP Publishing

# Small Kernelled Deep Capsule Networks for Handwritten Ge'ez Numerals Recognition

Ruchika Malhotra[1], Maru Tesfaye Addis[2, a)]

[1]*Department of Software Engineering, Delhi Technological University, Delhi 110042, India*
[2]*Department of Computer Science & Engineering, Delhi Technological University, Delhi 110042, India*

[a)] *Corresponding author E-mail: marutcomp@gmail.com*

**ABSTRACT:** Handwritten numeral recognition poses a significant challenge in various scripts, including Ge'ez, an ancient script predominantly utilized in Ethiopia and Eritrea for producing historical documents and religious books. We propose a deep learning approach for recognizing handwritten Ge'ez numerals, which applies a Capsule Network (CapsNet) with small kernel architecture, consisting of the convolutional, primary capsule, and digit capsule layers. To train and evaluate the model, we used a dataset of 51,952 handwritten Ge'ez numeral images from 524 individuals. The experimental results demonstrate exceptional performance, with an impressive testing accuracy of 93.32% and a low testing loss of 11.62%. These results surpass previous studies in handwritten Ge'ez numeral recognition, highlighting the effectiveness of our approach.

**Keywords:** Handwritten recognition, Handwritten digit recognition, Ge'ez numerals, Deep learning, Capsule networks, Character recognition.

## Introduction

The Ge'ez language, used in Ethiopian churches and for historical and religious texts, employs the Ethiopic script. It follows a left-to-right reading and writing system, similar to the Latin script [1]. The Ge'ez numeral system comprises 20 distinct digit symbols, as depicted in Table I. Notably, no specific symbol exists for the digit 0. Each digit from 1 to 9 has its corresponding representation, while numbers such as 10, 20, 30, and so on, as well as 100 and 10,000, possess their unique symbols. Additional digital representations are formed through the combination of these main symbols, with each digit featuring two horizontal lines above and below it.

Table I. Ge'ez numbering and English equivalent numeric symbols

| Digits | English symbol | Numbers | English symbol |
|---|---|---|---|
| ፩ | 1 | ፲ | 10 |
| ፪ | 2 | ፳ | 20 |
| ፫ | 3 | ፴ | 30 |
| ፬ | 4 | ፵ | 40 |
| ፭ | 5 | ፶ | 50 |
| ፮ | 6 | ፷ | 60 |
| ፯ | 7 | ፸ | 70 |
| ፰ | 8 | ፹ | 80 |
| ፱ | 9 | ፺ | 90 |
| | | ፻ | 100 |
| | | ፼ | 10,000 |

Handwritten digit recognition is a crucial task in various applications, including automated mail sorting, bank check reading, and touchscreen text recognition [2] [3]. In the domain of optical character recognition (OCR), the study of recognizing handwritten digits is a prominent area of research [4] . However, it poses challenges such as handwriting variability, incomplete or overlapping strokes, noise, distortion, large dataset sizes, and generalization to new data [3] [5] [6] . To address these challenges, researchers have explored techniques like pre-processing, feature extraction, and deep learning algorithms to develop accurate and robust models for handwritten digit recognition [7]. While many ancient languages have been extensively studied for OCR, the Ge'ez script has received less attention, with most previous works focused on printed texts [8] [9] [10] [11]. The recognition of handwritten Ge'ez scripts

*Recent Advancements in Artificial Intelligence and its Applications in Engineering*
AIP Conf. Proc. 3233, 020014-1–020014-7; https://doi.org/10.1063/5.0234108
Published under an exclusive license by AIP Publishing. 978-0-7354-5098-1/$30.00



27 January 2025 05:55:07

remains largely unexplored due to the scarcity of datasets [1]. Despite the data requirements of deep learning, the Ge'ez digit dataset is limited to the one prepared by Nur et al.[1], making Ge'ez digit recognition particularly challenging.

Deep learning approaches, known for their superior performance in pattern recognition tasks, have significantly advanced the field [2]. Convolutional neural networks (CNNs) have played a crucial role in image processing-based pattern recognition. A notable advancement in deep learning is the introduction of Capsule Networks (CapsNets) by Geoffrey E. Hinton and colleagues, which overcomes limitations of traditional CNNs [12]. CapsNets have shown great potential across various domains, and their application to Ge'ez digit recognition has not been explored until now. This study contributes in the following ways:

Recognizing Ge'ez handwritten digits using a CapsNet-based model, which addresses the lack of attention given to Ge'ez language recognition.

Exploring the Capsule Network algorithm by applying it to a local dataset and providing extensive experimental results and analysis.

Utilizing multiple convolutional layers with small kernels in the proposed network to reduce trainable parameters and enhance network non-linearity through activation functions.

## *I. Related works*

Numeral recognition has been drastically studied for the reason that 1980s, using numerous techniques starting from conventional system mastering to deep studying strategies [13]. Traditional feature extraction methods for digit reputation have proven excessive accuracy [7]. Boukharouba and Bennia proposed a novel function extraction method the use of a Freeman chain code with vertical and horizontal information for spotting Persian handwritten digits [14]. Their technique removed the want for normalization and accomplished 98.55% accuracy the usage of a assist vector system (SVM) on a dataset of 80,000 Persian numerals. Nur et al. Utilized convolutional neural networks (CNNs) to recognize handwritten Ge'ez digits[1]. Agegnehu et al. Hired deep studying to understand Amharic punctuation marks and offline handwritten digits [15]. Their CNN structure consisted of three convolutional layers with a 3x3 kernel length, 3 max-pooling layers with a 2x2 kernel, and one completely linked layer with a density of 500. Using a dataset of 5,800 snap shots from 100 handwriting samples, they done a checking out accuracy of 70.04%. Ahamed et al. Proposed a CNN architecture to decorate Arabic handwritten numeral reputation [7]. Their version comprised two convolutional networks with 32 filters and 5x5 kernels, two max-pooling networks with a 2x2 kernel, dropouts, and absolutely related layers. By augmenting their preliminary dataset from three,000 to seventy two,000 photographs, they performed 98.95% accuracy for the progressed model and ninety nine.Ninety six% for the newly proposed model. Yu et al. Supplied an improved LeNet5 digit recognition algorithm with the aid of changing its final layers with an SVM classifier [16]. LeNet5 served as the feature extractor, and SVM performed the classification. Yoa et al. introduced the "FOD_DCNet," a CapsNet-based model for recognizing and separating fully overlapping digits [17]. They made three key modifications to the original capsule network: applying a small convolution kernel, expanding the capsule's dimension to avoid loss of features, and employing dual dynamic routing. Their proposed model achieved 93.53% accuracy, which is 5.43% higher than the traditional CapsNet. Table II shows the literature review.

Table II. Literature Review

| Authors | Title | Dataset size | Techniques | Results |
|---|---|---|---|---|
| Boukharouba and Bennia [14] | “Novel feature extraction technique for the recognition of handwritten digits” | 80,000 images | FreeMan Code + SVM | 98.55% |
| Nur et al [1] | “Handwritten Geez Digit Recognition Using Deep Learning” | 51,912 images | CNN | 96.21% |
| Agegnehu et al[15] | “Offline Handwritten Amharic Digit and Punctuation Mark Script Recognition using Deep learning” | 5,800 | CNN | 70.04% |

27 January 2025 05:55:07

| | | | | |
|---|---|---|---|---|
| Ahamed et al[7] | "Handwritten Arabic numerals recognition using convolutional neural network" | 72,000 images | CNN | 99.96% |
| Yoa et al[17] | "Deep capsule network for recognition and separation of fully overlapping handwritten digits" | 60,000 images | CapsNet | 93.5% |

## Methodology

The dataset utilized in this study comprises 51,952 handwritten Geez-digit images, which were provided by Nur et al. [1]. These images were collected from a diverse group of 524 individuals. The complete dataset can be accessed on Google Drive using the following link: https://drive.google.com/file/d/1KeXQhW8UbkmHZQqu4y-2PMnPk8pmYUN1/view?usp=sharing. For the purpose of experimentation, the dataset has been pre-processed and resized to 32x32 grayscale images. Minor adjustments have been made to the dataset as it was already prepared for training.

CapsNet was developed by Geoffrey Hinton and his colleagues in 2017 as a novel alternative to traditional Convolutional Neural Networks (CNNs) [12]. Unlike CNNs, which use max pooling that may discard important features, CapsNet introduces capsules as grouped neurons. These capsules capture various attributes of the detected entities in an image, such as position, size, orientation, and texture. Each capsule generates a probability vector representing these attributes[12]. Capsule networks are still an active area of research, and their practical applications and limitations are being explored. The architecture of a capsule network typically consists of six main layers: the input layer, the convolutional layer, the primary capsule layer, the routing layer by agreement, the digit capsule layer, and optionally, the decoder layer [12]. These layers work together to process and transform the input data, allowing the network to recognize and represent the hierarchical relationships among different entities in the image.

*A. The proposed capsule network architecture*

The proposed recognition model architecture (Figure 1) comprises multiple layers. The first layer is the input layer, taking in 32x32 handwritten grayscale Ge'ez digit images. The second layer is a convolutional layer with 64 filters, 5 kernels, ReLU activation, no padding, and a stride of 1, followed by batch normalization and max-pooling. The third and fourth layers are similar convolutional layers with 128 and 256 filters, respectively, incorporating batch normalization, ReLU activation, no padding, and a stride of 1, followed by max-pooling. Following these, there are dense layers to capture abstract features, dropout layers for regularization, and a learning rate scheduler for better convergence. The fifth layer is the primary capsule layer, utilizing the squash activation function, 32 channels, 8-dimensional vectors, a kernel size of 9, a stride of 2, and no padding. Lastly, the model includes the digit capsule layer, which takes parameters such as the number of capsules, dimensional vectors, and the number of routing iterations. Data augmentation techniques are applied during training, and the code is well-commented for clarity.

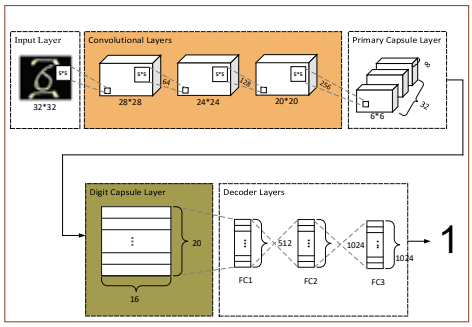



Figure 1 Architectures of proposed Capsule Networks

27 January 2025 05:55:07

## Experimental Setup and Results

### *A. Model Configuration and Experimental Setup*

Three capsule network models, namely Model1, Model2, and Model3, were tested with varying configurations. The models consist of convolutional layers, primary capsule layers, digit capsule layers, and decoding layers. The primary capsule, digit capsule, and decoder layers have the same configuration values in all models, while the number of convolutional layers and their corresponding input and output filters differ. Model1 is the simplest, with a unmarried convolutional layer, 1 input channel, thirteen kernels, ReLU activation characteristic, and 256 output channels. It offers a straightforward technique to feature extraction. Model2 contains two convolutional layers, with the primary layer having 1 input channel, 5 kernels, ReLU activation function, and 128 output channels. The 2d layer takes the 128 channels as enter, applies nine kernels with ReLU activation, and outputs 256 channels. This model increases complexity by introducing additional layers. Model3 is the most intricate, comprising three convolutional layers. The initial layer has 1 input channel, 5 kernels, and ReLU activation, producing 64 output channels. The subsequent layers sequentially increase channel size (64, 128, 256), using 5 kernels and ReLU activation each time. Model3's deep architecture offers enhanced capacity for learning and feature extraction. The choice of model depends on the task, dataset, and desired complexity, with Model1 suitable for simpler tasks and limited resources, while Model3 provides more capacity for complex tasks requiring high-level feature extraction. Models with small-sized kernels, like Model3, can capture fine-grained features. In this model architecture we modify the convolutional layer model from 1 two three and the kernel size from 9 to 5 from the original study of the capsule network [12].

### *B. Results Discussion*

The datasets of handwritten Ge'ez digits were divided into training, validation, and testing sets. Since the dataset is average in size we took 10% validation set was extracted from the training set. Three models were developed, varying in the number of convolutional layers specifically their kernel values. The models were trained with a batch size of 32, 30 epochs, and 3 routing iterations. Testing results for the three model configurations are presented in Table III and the comparison of the three models testing accuracy are presented in Figure 2. Among the models, Model3 exhibited superior performance in accuracy and loss tests, surpassing Model1 and Model2. Model3 achieved an accuracy rate of 93.32% and a loss of 11.62%, making it the most suitable configuration for this study. For a detailed reference, refer to the model architecture depicted in Figure 1. Early stop-control mechanisms were employed to prevent overfitting. Despite setting the number of epochs to 30, significant improvements in validation results were observed at earlier stages for all three model configurations, resulting in early stops at epochs 17, 20, and 19 for Model1, Model2, and Model3, respectively.

Table III. The performance of the CapsNet for three different model configurations.

| Models | Numbers of Routings | Numbers of epochs | Batch Size | Parameters | Loss (%) | Accuracy (%) |
|---|---|---|---|---|---|---|
| Model1 | 3 | 17 | 32 | 7,091,456 | 25.70 | 84.72 |
| Model2 | 3 | 18 | 32 | 7,712,576 | 14.59 | 91.51 |
| Model3 | 3 | 22 | 32 | 8,073,984 | 11.62 | 93.32 |

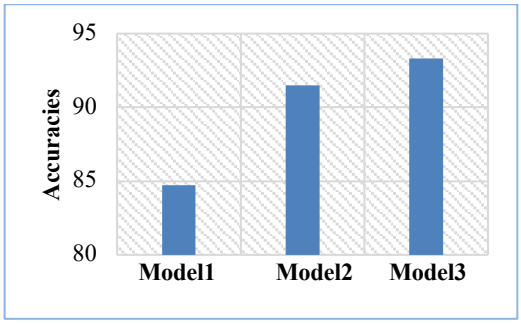



Figure 2 The comparison chart for three model's testing accuracy results.

27 January 2025 05:55:07

Figure 3 illustrates the training and validation accuracy and losses obtained from the experiments. The Model 3 configuration has better results. As we saw here, the model configuration with multiple convolutional layers and a small kernel size performs better in terms of recognition accuracy. This is likely due to the fact that a multi-layer convolutional neural network is able to capture more complex features and patterns in the data [18], which in turn leads to better recognition accuracy. The small kernel size is also beneficial because it allows for more layers to be added to the network, increasing the model's capacity to capture more features.

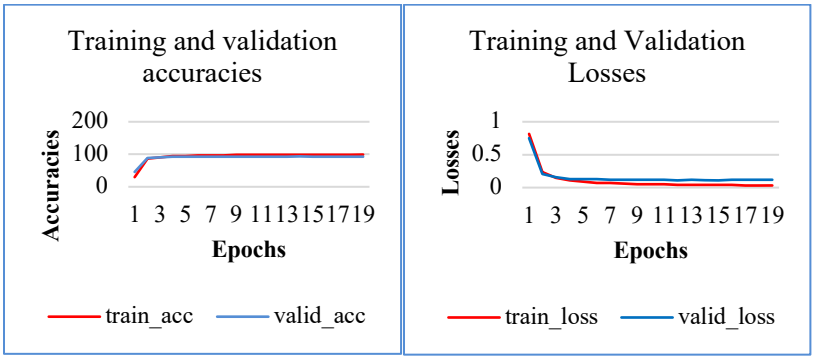



Figure 3 presents two distinct plots: (a) illustrates the training and validation accuracies, while (b) depicts the training and validation losses across epochs for the proposed CapsNet recognition model.

A confusion matrix was constructed for the proposed model, as illustrated in Figure 4, to provide a visual representation of correctly and incorrectly predicted Ge'ez digits. The diagonal elements marked in blue represent the accurately predicted digits, while the remaining values correspond to incorrect predictions. Notably, some digits were misidentified by the Ge'ez digit recognition model being evaluated. Among the 20 digits examined, the digit **%** shares a similarity with other digits such as ,8, and 0and, leading to high confusion with those digits. It is worth mentioning that in such scenarios, the erroneously recognized samples tend to possess shapes very similar to those of the predicted digits. Therefore, it is common for closely related characters to generate comparable features, resulting in confusion during the recognition process.

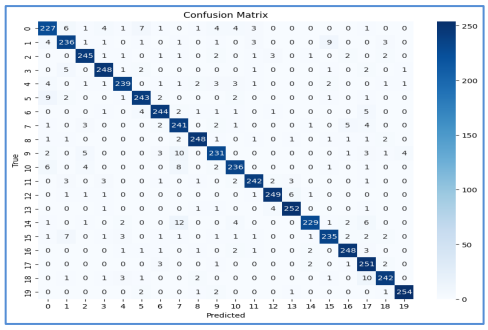



Figure 4 The confusion matrix of the proposed CapsNet recognition model

### *C. Result in Comparison*

In comparison to reference [19] , which achieved an accuracy of 89.88% in Ge'ez handwritten digit recognition, our proposed model "Model3" outperformed it by achieving a higher accuracy of 93.32%. On the other hand, when considering the testing loss, our study [1] reported a minimum testing loss of 20.13%. In comparison, our proposed model demonstrated superior performance in terms of testing loss values 11.62%, indicating better accuracy in predicting the Ge'ez handwritten digits. Therefore, our proposed model not only surpasses the accuracy achieved in reference [19] but also outperforms the testing loss reported in reference [1].

27 January 2025 05:55:07

# Conclusion

This research paper proposes a deep learning approach using a Capsule Network model for recognizing handwritten Ge'ez digits. The proposed CapsNet architecture consists of three convolutional layers, one primary capsule layer, and one DigitsCap layer, tailored to capture Ge'ez digit characteristics. Extensive configuration tuning optimizes the model architecture, with a focus on smaller kernel values. Promising results in recognition accuracy are achieved, but further investigation is needed to enhance Ge'ez digit recognition. Future work can explore dataset variation by including diverse professional groups' handwriting samples. Additionally, the application of the capsule network algorithm to develop a multi-digit recognition model is recommended. Furthermore, improving recognition accuracy, especially in handling shape similarity and complex features, should be a focal point. These research directions aim to advance Ge'ez handwritten digit recognition and enhance the performance of the Capsule Network model.

# References


1. Nur, M. A., Abebe, M., & Rajendran, R. S. (2022). "Handwritten Geez Digit Recognition Using Deep Learning." Applied Computational Intelligence and Soft Computing, Open Access, vol. 2022, pp. 1-12.
2. Aly, S., & Mohamed, A. (2019). "Unknown-Length Handwritten Numeral String Recognition Using Cascade of PCA-SVMNet Classifiers." IEEE Access, vol. 7, pp. 52024-52034.
3. Fateh, A., Fateh, M., & Abolghasemi, V. (2021). "Multilingual handwritten numeral recognition using a robust deep network joint with transfer learning." Information Sciences, vol. 581, pp. 479-494.
4. RAHMAN, A. B. M. A., et al. (2022). "Two Decades of Bengali Handwritten Digit Recognition: A Survey." IEEE Access, vol. 10, p. 36.
5. Goel, P., & Garantra, A. (2023). "Handwritten Gujarati Numerals Classification Based on Deep Convolution Neural Networks Using Transfer Learning Scenarios." IEEE Access, vol. 11, pp. 20202-20215.
6. Qiao, J., Wang, G., Li, W. W., & Chen, M. (2018). "An adaptive deep Q-learning strategy for handwritten digit recognition." Neural Networks, pp. 61–71.
7. Ahamed, P., Kundu, S., Khan, T., Bhateja, V., Sarkar, R., & Mollah, A. F. (2020). "Handwritten Arabic numerals recognition using convolutional neural network." Journal of Ambient Intelligence and Humanized Computing, vol. 11, pp. 5445–5457.
8. Meshesha, M., & Jawahar, C. V. (2005). "Recognition of printed Amharic documents." Presented at the Eighth International Conference on Document Analysis and Recognition (ICDAR'05), Seoul, Korea. [Available here](https://ieeexplore.ieee.org/ielx5/10526/33307/01575652.pdf?tp=&arnumber=1575652&isnumber=33307&ref=)
9. Belay, B. H., Habtegebrial, T. A., & Stricker, D. (2018). "Amharic Character Image Recognition." Presented at the 18th IEEE International Conference on Communication Technology, Chongqing, China.
10. Malhotra, R., & Addis, M. T. (2021). "Ethiopic Base Characters Image Recognition using LSTM." Presented at the 2021 2nd International Conference on Computational Methods in Science & Technology (ICCMST), Mohali, India.
11. Addis, D., Liu, C. D., & Ta, V. D. (2018). "Printed Ethiopic Script Recognition by Using LSTM Networks." Presented at the 2018 International Conference on System Science and Engineering (ICSSE), New Taipei, Taiwan. [Available here](https://ieeexplore.ieee.org/ielx7/8500054/8519965/08519972.pdf?tp=&arnumber=8519972&isnumber=8519965&ref=)
12. Sabour, S., Frosst, N., & Hinton, G. E. (2017). "Dynamic Routing Between Capsules." Presented at the NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems.
13. Nidhal, A. (2023). "Handwritten digits recognition using transfer learning." Computers and Electrical Engineering, vol. 106, p. 108604.
14. Boukharouba, A., & Bennia, A. (2017). "Novel feature extraction technique for the recognition of handwritten digits." Applied Computing and Informatics, vol. 13(1), pp. 19-26.


27 January 2025 05:55:07

15. Agegnehu, M., Tigistu, G., & Samuel, M. (2021). "Offline Handwritten Amharic Digit and Punctuation Mark Script Recognition using Deep learning." Presented at the 2nd Deep Learning Indaba-X Ethiopia Conference 2021, Adama, Ethiopia.
16. Yu, N., Jiao, P., & Zheng, Y. (2015). "Handwritten Digits Recognition Based on Improved LeNet5." Presented at The 27th Chinese Control and Decision Conference (2015 CCDC), Qingdao, China.
17. Yao, H., Tan, T., Xu, C., Yu, J., & Bai, X. (2021). "Deep capsule network for recognition and separation of fully overlapping handwritten digits." Computers & Electrical Engineering, vol. 91, pp. 1-12.
18. Abidi, S. F. H., Krishnan, A. R., Sumukhi, T., Clement, A., & Sasi, D. S. (2021). "A Novel Deep Convolutional Neural Network for Tuberculosis Detection." International Journal for Research in Applied Science & Engineering Technology, vol. 9(VI), pp. 1135-1139.
19. Beyene, E. G. (2019). "Handwritten and Machine printed OCR for Geez Numbers Using Artificial Neural Network." arXiv, vol. 1.

27 January 2025 05:55:07

Neural Computing and Applications
https://doi.org/10.1007/s00521-024-10902-y

S.I.: HYBRID APPROACHES TO NATURE-INSPIRED OPTIMIZATION ALGORITHMS AND THEIR APPLICATIONS

# Software defect prediction based on multi-filter wrapper feature selection and deep neural network with attention mechanism

Ruchika Malhotra[1] · Sonali Chawla[1] · Anjali Sharma[2]

Received: 10 May 2023 / Accepted: 12 December 2024
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2025

**Abstract**
Software defect prediction (SDP) models rely on various software metrics and defect data to identify potential defects in new software modules. However, the performance of these predictive models can be negatively impacted by irrelevant, redundant metrics and the imbalanced nature of defect datasets. Additionally, the previous studies mainly use conventional machine learning (ML) techniques, but their predictive performance is not superior enough. Addressing these issues is crucial to improve the accuracy and effectiveness of SDP models. This study presents a novel approach to SDP using a multi-filter wrapper feature selection technique (MFWFS). To identify a subset of relevant and informative features, we leverage the combination of filter techniques—Information gain (IG), Chi-square (CS), and Relief-F (RF) method, and a wrapper technique—Opposition-Based Whale Optimization Algorithm (OBWOA). One-dimensional-Convolutional Neural Network (CNN) with an attention mechanism is employed to enhance the classification performance of the predictive model by efficiently integrating the selected characteristics into abstract deep semantic features. We undertake experiments on seventeen open-source software datasets on four performance measures—AUC, G-mean, F-measure, and MCC and compare the obtained results with existing state-of-the-art ML and hybrid algorithms. The experimental findings demonstrate the greater efficiency of our approach, highlighting the usefulness of the multi-filter wrapper feature selection technique and 1D-CNN with attention to SDP.

**Keywords** Software defects · Filter techniques · Wrapper techniques · Opposition-based whale optimization algorithm · CNN · Attention mechanism

## 1 Introduction

SDP plays a significant role as a software quality assurance technique. It is a process of developing an efficient predictive model by leveraging data mining or ML algorithms using the software metrics and historical defect data obtained from different software repositories [1]. The early estimation of software defects during the development phase helps minimize the effort and time required to complete a project, where developers and managers can efficiently find potential defects by allocating the testing resources judiciously toward defect-prone modules.

Researchers have identified metrics that strongly correlate with defects in software defect datasets based on factors such as code complexity and development process analysis. These metrics can be considered attributes or

✉ Anjali Sharma
anjali@nplindia.org

Ruchika Malhotra
ruchikamalhotra@dtu.ac.in

Sonali Chawla
sonalichawla_phd2k19@dtu.ac.in

[1] Department of Software Engineering, Delhi Technological University, Bawana Road, Delhi 110042, India

[2] Department of Planning Monitoring and Evaluation, CSIR-National Physical Laboratory, Dr. KS Krishnan Marg, Pusa, New Delhi 110012, India

Published online: 09 January 2025

features. However, not all the attributes are desirable for creating defect prediction models. The amount of time needed to train a model and its prediction's quality may be negatively impacted by redundant and irrelevant attributes, commonly known as the curse of dimensionality. The previous works [2–5] have proven that feature selection techniques help diminish the big dimensionality problem in software defect datasets.

Feature selection identifies the most representative feature subset from the initial set by assessing each feature's significance while preserving the model's high classification performance. The three categories of feature selection techniques that help select the informative features are—filter, wrapper, and hybrid approach. A filter approach relies on the data's characteristics to assess the feature's relevance rather than any classification algorithm to evaluate the feature subset [6, 7]. Wrapper approaches outperform filter methods by employing the greedy approach to evaluate all potential optimal feature subsets with a classification algorithm [8, 9]. Though filter methods are computationally less expensive and faster, the wrapper methods produce a better subset of features. Hybrid methods merge the complementary strengths of both filter and wrapper techniques [10].

Metaheuristic algorithms are frequently applied in a variety of optimization problems in recent years, and feature selection is another vital study area where these methods have been studied. Unfortunately, there is no assurance that the feature selection will provide the best combination of attributes due to the stochastic nature of metaheuristic algorithms. Several metaheuristic algorithms have been applied in SDP, such as genetic algorithms [11], particle swarm optimization [12], artificial bee colony [13, 14], ant colony [15], moth flame optimization [2], salp swarm optimization [16], etc. Whale optimization algorithm (WOA), a recently developed metaheuristic algorithm, mimics the intelligent search strategy used by humpback whales to find food by the method known as bubble-net feeding. Although WOA has superior search ability compared to other existing algorithms, it has limitations [17] such as getting stuck in local optima. WOA starts with random positions for each agent in the group and switches locations with either a randomly assigned search agent or the current best solution available throughout the search. The WOA is combined with the idea of an opposition-based learning concept to form OBWOA that improves its base version [18]. By determining if the opposite position for each agent is better, the opposition-based idea can assist WOA in finding accurate solutions.

Instead of employing the filter and wrapper techniques independently, we combine the features produced from both techniques to take maximum advantage of both techniques. This helps to avoid missing any relevant information and enhances the prediction model's overall search performance. This research work uses multiple filter methods such as Information Gain, Chi-square, and Relief-F methods as per the previous research findings [19, 20]. The features obtained are then integrated with the feature subset obtained from the wrapper method—opposition-based whale optimization to maximize the classification performance (AUC) of the predictive model. Filter methods extract the information from features, and the wrapper uses the learning algorithm for judgment. The computational time and complexity are reduced compared to the pure wrapper method [10]. We believe that this multi-filter wrapper technique has never been applied earlier in the SDP domain.

A predictive model's performance mainly depends on two important factors: feature representation and classification algorithm. Researchers have been applying ML and deep learning (DL) algorithms to efficiently construct models for predicting software defects [3, 21–27]. ML algorithms need manually extracted software metrics for building classifiers to predict defect or non-defect-prone areas. While DL algorithms automatically extract high-level features and learn from high-dimensional and more complex data [28]. As a result, a lot of researchers are now concentrating on creating SDP models using DL algorithms. CNN has gained popularity in many fields, including speech recognition [29], semantic search [30], object detection [31], etc., recently because of its capability to extract semantic features with powerful discriminating capacity compared to ML algorithms. They are useful for developing SDP models, but they seem complicated and have insufficient low accuracy values. This may be because CNNs were originally built using 2D structures based on 2D data such as images and videos. Recently, there has been a development of a new modified variant of 2D-CNN called 1D-CNN [32–34]. They are preferable over 2D-CNN due to reasons such as—the complexity of the 1D variant (O(NK)) is lower than that of the 2D variant ($O(N^2K^2)$) [32]. One-dimensional variant implementation is feasible on a standard computer and comparatively faster due to fewer hidden layers and fewer parameters involved. Its exceptional performance in real-time electrocardiogram (ECG) monitoring [34], structural damage detection [35], high-power engine fault monitoring [36], and automatic speech recognition [37] have all contributed to its increased popularity. Inspired by the successful application of 1D-CNN in these areas, 1D-CNN with an attention mechanism has been employed as the classification algorithm for creating a defect prediction model.

The following are the main objectives of the paper:

- A novel hybrid model based on the **multiple filter wrapper feature selection (MFWFS) techniques** is

proposed. This algorithm combines three filter methods—**Information Gain**, **Chi-square**, and **Relief-F** methods with a swarm-based **oppositional whale optimization algorithm (OBWOA)** that eliminates the redundant and irrelevant features and selects the optimal feature subset to improve the overall model performance.

- A powerful unified defect predictor based on **CNN using an attention mechanism** is created that integrates the automatic features generated from the output of CNN and features obtained from the MFWFS technique.
- The hybrid model is evaluated on 17 software defect datasets obtained from NASA, AEEEM, and Promise repositories. We have computed the performance metrics—**AUC, *G*-mean, *F*-measure, and MCC** to validate and analyze the classification performance of the proposed model.
- The predictive capability of the proposed model is compared to other state-of-the-art ML and hybrid algorithms present in the literature to depict that our model has produced better results across all the datasets. We employed the **Wilcoxon test** to establish the statistical significance of our results.

The above objectives were achieved by addressing the research questions outlined below:

RQ1. How efficient is the SDP model developed using 1D-CNN with and without multi-filter wrapper feature selection?
RQ2. How efficient is the SDP model developed using MFWFS-1D-CNN with and without attention?
RQ3. How effective is the SDP model developed using the proposed hybrid model MFWFS-1D-CNN-Attention in comparison with the models created with individual filters and wrapper-based CNN model?
RQ4. How effective is the SDP model developed using the proposed hybrid classifier (MFWFS-1D-CNN-Attention) in comparison with other state-of-the-art ML and hybrid techniques?

The remainder of the paper is organized as follows. Section 2 analyzes the relevant works in the field of SDP, including DL models and other hybrid models that use feature selection techniques as a part of pre-processing. Section 3 gives a background on the feature selection algorithms used in this study. In Sect. 4, an in-depth explanation of the proposed methodology is presented. Section 5 describes the experimental layout of the paper. Section 6 presents the obtained results and analyzes them in the form of answers to the research questions. Section 7 mentions the validity threats posed in this study, and Sect. 8 gives the conclusion and future work to our study.

## 2 Related work

Many researchers in the past have explored the use of feature selection techniques with DL-based classification methods in SDP due to their significant performance characteristics.

Zhu et al. [3] proposed an improved feature selection algorithm combining two metaheuristic algorithms named WOA and a simulated annealing algorithm called EMWS. A defect predictor named WSHCKE was constructed as a hybrid deep neural network combining CNN and KELM. The experiment was conducted on 20 different software projects using four performance evaluation measures—F1 score, MCC, G-measure, and AUC and concluded that EMWS and WSHCKE performed superior to baseline feature selection and classification algorithms.

Arar and Ayan [14] proposed a hybrid classifier where an artificial bee colony algorithm (ABC) optimizes the connection weights of the neural network (ANN). The cost-sensitive feature was introduced to the ANN by implementing the error function that considers the misclassification cost of each class and tries to minimize the total misclassification cost. The suggested method was evaluated using five publicly accessible NASA software datasets, taking into account accuracy, probability of detection, probability of false alarm, balance, and AUC measures. The obtained results show that the proposed approach outperformed other ML algorithms.

Kassaymeh et al. [16] optimized the parameters of the BPNN algorithm using the metaheuristic technique of salp swarm optimization aiming to enhance the accuracy of the SDP model. The evaluation of results was conducted on 22 software datasets using accuracy, AUC, specificity, sensitivity, and error rate as the performance measures. In comparison with state-of-the-art algorithms, the SSA-BPNN method demonstrated superior performance, outperforming all other methods.

Jin et al. [38] developed a model that combines ANN and quantum particle swarm optimization (QPSO). ANN acts as a classifier for predicting defects and non-defects, whereas QPSO is used for feature selection. The results were evaluated on four NASA software projects using the AUC measure. The performance of this hybrid model was compared to other baseline ML classifiers which showed the superiority of the QPSO + ANN model.

Turabieh et al. [39] employed binary variants of three wrapper-based feature selection methods—particle swarm optimization (PSO), ant colony optimization (ACO), and genetic algorithm (GA), iteratively, to improve the performance of a layered recurrent neural network (LRNN) for the SFP problem. The results were compared with and without 5 cross-validation, and statistical analysis proved

that cross-validation with feature selection was better than other approaches using AUC values.

Li et al. [21] presented a framework named Defect Prediction via CNN (DP-CNN) that automatically produces syntactic and semantic features obtained from source code using abstract syntax trees and integrates them with traditional features for SDP. According to the results, DP-CNN advances the existing techniques achieving an average 12% increase in F-measure.

Hassouneh et al. [40] presented an efficient feature selection approach for SDP using WOA augmented with five natural selection operators. Four different classifiers were used—KNN, DT, LDA, and SVM. The tournament-based WOA with DT classifier outperformed other variants of WOA in terms of AUC measure over 17 public software datasets.

Balogun et al. [41] introduced a new hybrid multi-filter wrapper feature selection technique that uses a rank aggregation method to choose irrelevant and redundant features in software defect datasets. Twenty-five datasets were selected from various software repositories to check the feasibility of the novel technique on Naïve Bayes and Decision Tree classifier to evaluate its performance. The evaluation metrics used were F-measure, AUC, and accuracy. The results indicate this feature selection method effectively addresses issues such as filter rank selection and local optima stagnation.

Zain et al. [42] proposed a novel 1D-CNN structure for SDP that is efficient in comparison with other 1D-CNN models varying in filter size, kernel size, dropout layers, and a number of convolutional and max-pool layers. The models were tuned with different hyperparameters and evaluated based on F-measure, accuracy, training, and testing time. The comparison with other ML algorithms showcased the exceptional performance of the proposed model.

Fan et al. [43] presented a DL-based approach named defect prediction via an attention-based RNN (DP-ARNN). This model makes use of abstract syntax trees that help to learn syntactic and semantic features from the source code of seven open-source Java software projects. F-measure and AUC were chosen as the performance metrics. DP-ARNN, on average, improved the F-measure and AUC scores by 14% and 7% in comparison with state-of-the-art techniques.

Akay and Akay [44] utilized the ABC algorithm to train the ANN by optimizing its weights and bias for predicting the number of software defects in five object-oriented software datasets. They compared the results with three other training algorithms, and the performance metrics—recall, precision, accuracy, and F1 scores were computed. In SDP, the ABC algorithm proved to be a superior training algorithm for ANN.

However, the hybrid approaches applied in existing literature have very limited use of feature selection methods combined with deep neural network architectures for SDP. The main focus of this study is to propose a unique hybrid feature selection technique that integrates the existing filter methods—IG, CS, RF, and the modified wrapper method—OBWOA. The OBWOA is chosen because of its significant competitiveness over other metaheuristic algorithms in solving complex neural network problems. The features obtained from the above unification, along with superior feature extraction capabilities of 1D-CNN combined with the coordinate attention mechanism result in a powerful unified defect predictor that has superior classification performance compared to other existing ML algorithms. Further, the Wilcoxon test is employed for statistical analysis to validate the results and assess the robustness of the model's performance compared to other algorithms present in the literature.

## 3 Background

### 3.1 Filter methods

Filter methods rely on univariate statistics to capture the inherent characteristics of features, independent of any learning algorithm. These approaches are quicker and less computationally intensive compared to wrapper methods [7]. In this work, we have used the information gain, chi-square, and relief-F methods described below.

- **Information gain**

  It is a technique that determines the relevance of a given attribute for the class label prediction. It is an entropy-based concept that measures the uncertainty in the class label after the attribute value is observed [45]. It can be found using Eq. (1).

$$\begin{aligned} IG(y) = &-\sum_{i}^{m} p(x_i)\log p(x_i) \\ &+ p(y)\sum_{i}^{m} p(x_i/y)\log p(x_i/y) \\ &+ p(y')\sum_{i}^{m} p(x_i/y')\log p(x_i/y') \end{aligned} \tag{1}$$

  where $x_i$ denotes the ith class—defective/non-defective, $p(x_i)$ denotes the probability of the ith class, $p(y)$ and $p(y')$ signify the probabilities of the presence/absence of the attribute $y$, and $p(x_i/y)$ and $p(x_i/y')$ denote conditional probabilities given the presence/absence of the attribute $y$, respectively.

- **Chi-Square**

  It is the goodness of the fit test which assesses the

independence of two variables [46]. It determines the most relevant features by considering the highest values for the chi-squared statistic of the class label computed using Eq. (2)

$$\chi 2 = \sum \frac{(\text{Obs}_i - \text{Exp}_i)^2}{\text{Exp}_i} \tag{2}$$

where $\text{Obs}_i$ and $\text{Exp}_i$ are the observed and expected values. Higher values of chi-squared statistic indicate a stronger dependence of the feature on the output variable.

- **Relief-F method**

  This is an instance-based method that relies on the concept of nearest neighbors to estimate the quality of the attributes [47]. For each instance in a training set, two nearest neighbors are determined—one belonging to the same class (HIT) and the other of a different class (MISS). The importance of the attribute is then calculated using Eq. (3).

$$I(Y) = P(y/\text{nearestinstance}_{\text{different}}\text{class}) - P(y/\text{nearest instance}_{\text{same}}\text{class}) \tag{3}$$

  where $P$ is the probability, and $y$ denotes the value of an attribute. An important attribute should be able to effectively differentiate between instances of different classes while maintaining similar values for instances within the same class [48].

### 3.2 Wrapper method

Wrapper methods assess the quality of various feature subsets utilizing a particular machine learning algorithm. In contrast with filter methods, that rely on intrinsic properties, wrappers measure the usefulness of the features by training and evaluating a classifier based on the chosen algorithm [49]. In this research, we apply the swarm-based whale optimization algorithm as a wrapper method for finding the optimal feature subset.

- Whale Optimization Algorithm

  It is a stochastic population-based metaheuristic algorithm recently invented in 2016 [17]. This algorithm mimics the intelligent search strategy adopted by humpback whales known as bubble-net feeding for searching their food. They create bubbles in a spiral motion around their food, i.e., small fishes, and then slowly swim toward the surface.

  There are two stages in a WOA algorithm—the first involves encircling the prey and using a spiral-bubble-net attacking mechanism (**the exploitation stage**), and the second involves randomly searching for prey (**the exploration stage**) [50].

- **Exploitation phase** *Shrinking Encircling Mechanism*: The humpback whales swim in a spiral-like pattern as they approach their prey in a shrinking encircling mechanism. They update their positions and move toward the best solution according to the current optimal solution, using Eqs. (4) and (5):

$$\overrightarrow{D} = \left|\overrightarrow{C}.\overrightarrow{Z_{\text{best}\,t}} - \overrightarrow{Z_t}\right| \tag{4}$$

$$\overrightarrow{Z}_{t+1} = \overrightarrow{Z_{\text{best}\,t}} - \overrightarrow{A}.\overrightarrow{D} \tag{5}$$

  where $\overrightarrow{Z_{\text{best}\,t}}$ denotes the position of the best solution thus far. The coefficients $\overrightarrow{A}$ and $\overrightarrow{C}$ can be computed using Eqs. (6) and (7),

$$\overrightarrow{A} = 2\overrightarrow{a}.\overrightarrow{r} - \overrightarrow{a} \tag{6}$$

$$C = 2\overrightarrow{r} \tag{7}$$

  $\vec{a} = 2 - t\left(\frac{2}{M}\right)$ is linearly decreased from 2 to 0, $M$ refers to the maximum number of iterations, and $\overrightarrow{r}$ refers to any random vector in [0, 1].

- *Spiral path updating position*:

  The bubble-net feeding mechanism calculates the distance between the current and the best solution thus far using the spiral Eq. (8), as illustrated below.

$$\overrightarrow{Z}_{t+1} = \overrightarrow{D\prime}.e^{bl}.\cos(2\pi l) + \overrightarrow{Z_{\text{best}\,t}} \tag{8}$$

  where $b$ defines the constant that dictates the shape of the logarithmic spiral function, and $l$ represents a random value in [-1,1].

  According to Eq. (9), humpback whales have a 50% chance of randomly selecting either a spiral-shaped path model or a shrinking encircling mechanism to approach their prey.

$$\vec{Z}_t = \left\{ \begin{array}{ll} \overrightarrow{Z_{\text{best}\,t}} - \vec{A}.\vec{D}, & p < 0.5 \\ \overrightarrow{D\prime}.e^{bl}.\cos(2\pi l) + \overrightarrow{Z_{\text{best}\,t}} & p \geq 0.5, \end{array} p\varepsilon[0,1] \right\} \tag{9}$$

  where $p$ is some random number $\varepsilon$ [0,1].

- **Exploration phase**

  This phase is controlled by the variable $\overrightarrow{A}$ that lies between [-1,1]. If $\left|\overrightarrow{A}\right| \geq 1$, the position of each whale is modified according to the position of the randomly chosen whale, and if $\left|\overrightarrow{A}\right| < 1$, the position of each whale is modified according to the current optimal solution. This global search capability of WOA is shown in Eqs. (10) and (11):

$$\overrightarrow{D} = \left|\overrightarrow{C}.\overrightarrow{Z_{\text{rand}\,t}} - \overrightarrow{Z}_t\right| \tag{10}$$

$$\overrightarrow{Z}_{t+1} = \overrightarrow{Z_{\text{rand}_t}} - \overrightarrow{A}.\overrightarrow{D} \tag{11}$$

where $\overrightarrow{Z_{\text{rand}}}$ is a position of the whale selected randomly from the current population.

### 3.3 Opposition-based learning

This technology was proposed by Hamid R. Tizhoosh [51, 52], and it has been proven to greatly enhance the performance of optimization algorithms such as PSO [53], DE [54], GWO [55], and ACO [56], by evaluating the current and opposite solutions simultaneously. It increases the exploitation of search space and ultimately helps us to reach a globally optimal solution more quickly. We can use the below equations to find the opposite values.

*Opposite number*—Assume $z \in [a, b]$ is a real number, where $a$ and $b$ represent lower and upper limits. The opposite number is defined using Eq. (12).

$$\tilde{z} = a + b - z \tag{12}$$

*Opposite point*—Assume $Z(z_1, z_2, z_{3,\ldots}z_n)$ be a point where $z_1, z_2, z_3 \ldots\ldots z_d$ are all real numbers and $z_i \in [a_i, b_i]$, $i = 1,2,3,\ldots n$. The opposite of $Z$ represented as $(\tilde{Z}(\tilde{z_i}, \tilde{z_2}, \tilde{z_3}, \ldots \tilde{z_n}))$ and defined using Eq. (13).

$$\tilde{z} = a_i + b_i - z_i \tag{13}$$

*Opposition-based optimization*—Let $Z(z_1, z_2, z_{3,\ldots}, z_n)$ is a candidate solution in the search space. Its objective function is given by $f(Z)$. According to Eq. (14), $\tilde{Z}(\tilde{z_i}, \tilde{z_2}, \tilde{z_3}, \ldots \tilde{z_n})$ is the opposite of $Z(z_1, z_2, z_{3,\ldots}z_n)$ and its fitness function is given by $f(\tilde{Z})$. While evaluating the fitness function of both these points simultaneously, if $f(\tilde{Z})$ is found to be better than $f(Z)$, then the point $\tilde{Z}$ is considered a better solution and replaced by $Z$, or we retain $Z$ as the better solution.

The algorithm for opposition-based whale optimizations is shown below.

**OBWOA**

1. Initialize the whale population randomly $Z_i$ ($i=1,2,3\ldots\ldots,n$).
2. Compute the opposite for each $Z_i$ i.e. $Z_i$*=a+b- $Z_i$ using equation(13)
3. Compute the objective function (error rate) for each search agent
4. Compute the objective function for the opposite value of each search agent.
5. $\hat{Z}$= Best solution obtained
6. While (k<M)
    a. For each solution
        i. Update the coefficients – a, l, p, A, C
        ii. If(p<0.5)
            If ($|A| < 1$)
                Update the current solution's position using equation (5)
            Else if ($|A| \geq 1$)
                Select a solution randomly, and update the current solution's position by equation (12)
        iii. Else if (p>=0.5)
            Update the current solution's position using the upward spiral path equation (9)
    b. Compute the objective function for each search agent.
    c. Compute the objective function for the opposite value of each search agent.
    d. If there is any better solution than the original and opposite fitness value, update $\hat{Z}$.
    e. k=k+1
7. End while
8. Return $\hat{Z}$

## 4 Proposed methodology

This research paper presents a hybrid of the multi-filter wrapper feature selection technique with the attention-based 1D-CNN model. The proposed methodology is explained in Fig. 1.

### 4.1 Feature selection based on multi-filter wrapper technique

The proposed MFWFS technique aims to analyze and combine the computational strengths of various independent filter and wrapper-based feature selection methods. The filter methods (IG, CS, and RF) evaluate the features on the basis of their usefulness toward better prediction performance to construct an optimal subset of features with excellent prediction characteristics. The swarm-based oppositional whale optimization algorithm helps accelerate the convergence to the global optimum by finding solutions in opposite directions, increasing the chances of finding better regions effectively [57]. The MFWFS is carried out using the algorithm described below.

$$\text{Fitness} = \alpha.(1-p) + (1-\alpha).\frac{|S|}{|N|} \tag{14}$$

where $\alpha$ represents user-defined hyperparameter $\in$ [1, 0] that strikes a balance between both objectives. $p$ denotes AUC measure, $|S|$ represents feature subset length, and $|N|$ indicates total no. of features. $\alpha = 0.99$ as suggested in the literature [59, 60]. We apply the OBWOA on the original training input dataset to extract the subset of features with high AUC values, using KNN ($K = 5$) as the classifier. To provide the best possible set of features, feature subsets from filter and wrapper approaches are combined. The MFWFS is used only for feature selection for every separate software project. For classification purpose, we use the attention-based 1D-CNN defect predictor.

The primary contribution of the above feature selection technique is to explore the complementary nature of filter with wrapper feature selection with the aim to reduce the feature subset of software defect datasets and enhance the classification performance without compromising on the computational cost. The selection process of features for filter methods is fast, as there is no learning algorithm

Input: Dataset D
Output: Feature Subset
a. IG_Set = Apply Information Gain FS method
b. CS_Set = Apply Chi-square FS method
c. RF_Set = Apply Relief-F method
d. Filter_Set = $\log_2 n$(IG_Set) ∪ $\log_2 n$(CS_Set) ∪ $\log_2 n$(RF_Set)
e. Wrapper_Set = Apply Opposition-Based Whale Optimization algorithm
Reduced_Feature_Set = **Filter_Set** ∪ **Wrapper_Set**

The input software defect datasets are split into training and test datasets in the ratio of 70:30. To determine the ranks of all the features for each filter method, we independently apply the information gain, chi-square, and relief-F methods to the training set. The top-ranked features are selected from the generated rank list using $\log_2 n$, where '$n$' stands for the total number of features in every dataset. Based on the previous empirical studies, the choice of $\log_2 n$ is made [45, 58]. As a result, software defect datasets with fewer optimal features will be produced. In OBWOA, each solution refers to a feature subset for every project that is the 1D binary vector with '$n$' elements, each of which represents '$n$' features and has a value of either '0' or '1.' Two optimization goals should be considered—reducing the number of features and maximizing the classification performance. We combine the above two goals into an objective (fitness) function in Eq. (14) [40].

required, whereas, in the wrapper method, the selection process is slow due to the involvement of the learning algorithm. The integration of multiple filter methods IG, CS, and RF with the OBWOA wrapper-based method aims to leverage the strengths of both, resulting in a novel, optimized, and robust hybrid feature selection process that is unprecedented in the domain of SDP.

### 4.2 Defect prediction based on attention-based CNN

CNN is a deep neural network that efficiently handles image data. It expands the traditional artificial neural network by adding more layers, such as a convolutional layer, pooling layer, flattening layer, etc. A classic CNN applies to 2D data such as images and videos; therefore, it is commonly referred to as 2D-CNN. Another recent development is 1D-CNN, a modified form of 2D-CNN [32–34].

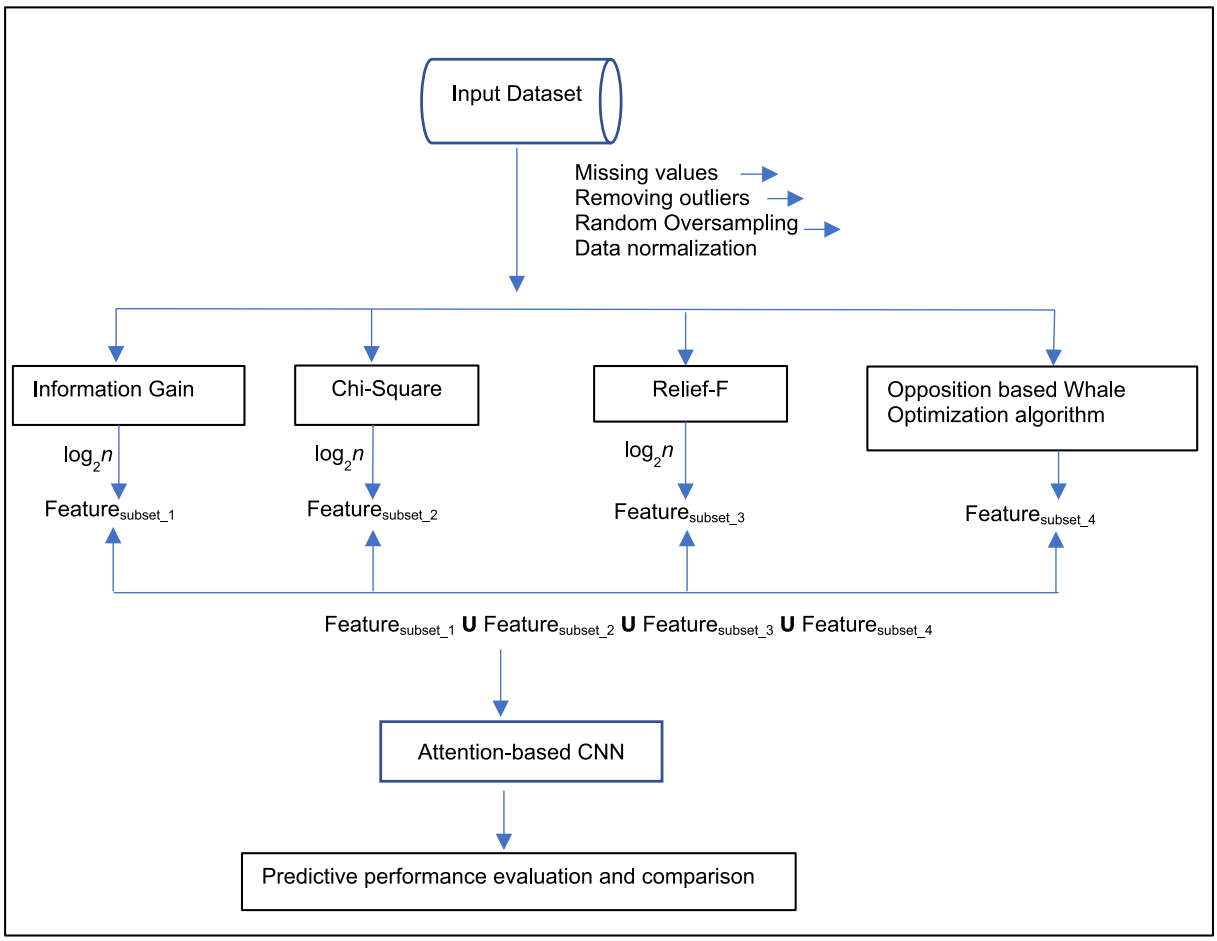



**Fig. 1** Proposed methodology

One-dimensional variant implementation is feasible on a standard computer and comparatively faster due to fewer hidden layers and fewer parameters involved.

Figure 2 illustrates the structure of the proposed 1D-CNN model with attention comprising the following important layers: -

a) **Input layer**—The 1D-input vector $K[n]$ where $n = 0,1, 2,\ldots\ldots N\text{-}1$ is given to the first layer of the CNN architecture. 'n' represents the software metrics present in the dataset. E.g.; the JM1 dataset contains 21 software metrics, so the input layer will accept 21 metrics.

b) **Convolutional layer**—It is the foundational layer of 1D-CNN. A series of convolutions are performed using filters or kernels to acquire local features from the input data. The filters help detect deep semantic features from the low-level feature representation in the defect instances. We have used two convolutional layers with 64 and 32 filters and a kernel size of 1 for each filter.

c) **ReLU layer—**The output from the preceding layer (feature map) is subjected to a nonlinear activation function by the ReLU layer. It gives the model nonlinearity and enhances its capacity for expression.

d) **Flatten layer—**The feature maps produced as a result of 2nd convolutional layer are transformed into a single long continuous linear vector through the process of flattening.

e) **Fully connected (FC) layer**—The output from the flattening layer is linearly transformed into a new representation through a set of weights and biases. To classify inputs accurately, we employed two FC layers. The first FC layer used ReLU activation, while the second FC layer used softmax activation generating a probability ranging from 0 to 1.

f) **Dropout layer—** A regularization approach randomly removes a particular proportion of the neurons in the FC layer during training. Hence, the model's generalization performance is enhanced, and overfitting is prevented.

g) **Attention mechanism—** It is a technique commonly used in DL models to focus on the most important

features or regions of an input, enabling the model to focus on relevant information and improve its performance. This mechanism has been effectively utilized in several computer vision tasks, such as object recognition [61], segmentation [62], and detection [63].

In the context of SDP, attention mechanisms can be used in a CNN to learn discriminative features from software artifacts, such as source code or execution traces. This facilitates the identification of potential defects in a software system before they turn into failures or errors. We use the coordinate attention as described by [64]. The coordinate attention captures direction-aware and position-sensitive data that aid models in precisely locating and identifying the defects in software. As a computing unit, a coordinate attention block tries to increase the expressive power of the learned features. Any intermediate feature tensor, $X = [x_1, x_2,..., x_C] \in \mathbb{R}^{C*H*W}$, may be used as input; transformed vector with $Y = [y_1, y_2,..., y_C]$ where X and Y are of the same size is the output.

The channel attention algorithm modifies the global pooling operation by factorizing the pooling into two 1D feature encoding processes, one along the horizontal and one along the vertical coordinates [64]. The input X undergoes encoding through two pooling kernels—(H, 1) and (1, W), capturing information along both horizontal and vertical coordinates for each channel. Therefore, Eq. (15) can be used to express the output of the c-th channel at height 'h.'

$$o_c^h(h) = \frac{1}{W}\sum_{i=0}^{W} x_c(h,i) \tag{15}$$

Likewise, c-th channel's output at width '$w$' can be written as

$$o_c^w(w) = \frac{1}{W}\sum_{j=0}^{H} x_c(j,i) \tag{16}$$

This allows the attention block to capture long-range spatial interactions while preserving precise positional information. This modification can be used in SDP to help 1D-CNN accurately locate defects of interest and improve the overall structure to capture deep semantic features.

The integration of CNNs with attention mechanisms aims to leverage the powerful feature extraction capabilities of deep neural networks while enhancing focus on discriminative regions within the input data. This enhances the overall performance and interpretability of the model, making it more effective in tasks that require discerning important features from complex input data.

## 5 Experimental setup

### 5.1 Dataset description

We perform experiments on the 17 open-source datasets obtained from—NASA [65], PROMISE [66], and AEEEM [67] which are publically available benchmark datasets for the SDP problem to ensure that the research results can be verified, replicated, and refuted [3, 16, 68, 69]. We have selected six datasets from NASA (JM1, KC1, MC1, PC3, PC4, and PC5), which contain traditional software metrics such as McCabe, Halstead, Line Count, etc. Five datasets are selected from AEEEM (Equinox, Eclipse JDT Core, Eclipse PDE, Lucene, and Mylyn) which contain 61 measures, including 17 source code metrics, 17 entropy of source code metrics, 17 churn of source code metrics, 5 previous-defect metrics, and 5 entropy of change metrics. The PROMISE datasets (ant-1.7, camel-1.6, jEdit-4.3, poi-3.0, xalan-2.7, and xerces-1.4) contain 20 Object oriented metrics, Mc Cabe's metrics, and CK metrics.

Table 1 provides basic statistical information for 17 projects across three datasets that include the number of instances, attributes, defects, defective ratio, and imbalance ratio. All these datasets vary in both size and complexity.

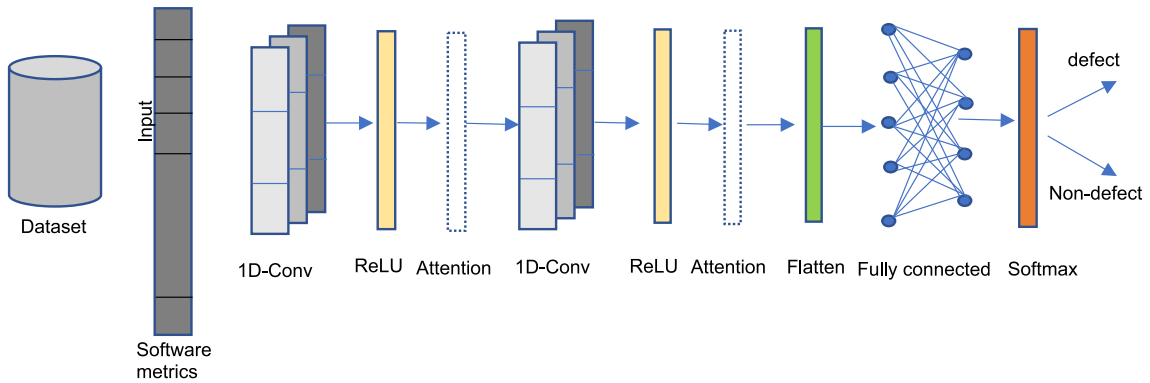



**Fig. 2** Architecture of attention-based 1D-CNN

Also, the imbalance ratio in all the datasets is quite high. Therefore, there is a need to address the imbalance issue in these software projects. The independent variables in these datasets denote the features/attributes, and the dependent variable is the class label ('1'—defective/'0'—non-defective). In some datasets, the dependent variable shows the count of defects within a particular module; in that case, we represent the class label as defective if there is more than one defect present; otherwise, we label it non-defective.

### 5.2 Data pre-processing

- **Missing values and Removal of Outliers**—This stage involved searching the datasets for any outliers and missing values that might have been present. The data points are analyzed in each dataset, and those that vary extremely from the remaining data points are considered outliers. These outliers are critical and need to be handled. In our study, we have used median imputation to treat the outliers since deleting the outliers would lead to degradation of the dataset and a reduction in the available data volume.
- **Class Imbalance processing**—The datasets used to solve SDP problems are usually imbalanced. There is a relatively small percentage of defective modules (minority class) as compared to non-defective modules (majority class) [70]. If an imbalanced dataset is put directly to train a model, the results may be poor and biased in favor of the majority class. Oversampling and undersampling are common approaches to resolve the class imbalance problem [71]. Oversampling includes generating duplicates of minority class instances, while undersampling entails eliminating instances from the majority class. A part of the training dataset may be partially lost in undersampling, resulting in underfitting. We use the random oversampling method to create a class-balanced training set that replicates the instances from the minority class randomly [72].
- **Data Standardization—**It is crucial to first normalize the input dataset before initiating the training process because each dimension in an input space may have a distinct range of values. The original dimension values must be standardized in order to produce reliable results, and the processed dimension values must correspond to the original distribution. In a training process, it is a common practice to normalize the data, ensuring that input values fall within the range [0, 1] [73]. This is typically done to boost the neural network training process and solve model learning issues. Equation (18) is used to normalize the dataset.

$$X_{\text{normalized}} = \frac{X - \min(X)}{\max(X) - \min(X)} \tag{18}$$

In the above equation, the minimum and maximum values of software metric X are denoted by min(X) and max(X), respectively.

### 5.3 Performance measures

It is possible to assess the effectiveness of predictive models for datasets with imbalanced class distributions by utilizing appropriate performance measures. In our study, we have considered AUC, F-measure, MCC, and G-mean, as these metrics are widely utilized in the literature [3, 39, 74–76]. The most traditional and commonly used performance metric to evaluate a classifier is accuracy. However, it gives biased results when handling imbalanced datasets [77]. Studies have advocated the use of AUC (Area under the Receiver's Operating Curve), a robust measure particularly effective for imbalanced and noisy datasets [68, 78]. Thus, in our research, we have employed AUC to examine the relationship between true-positive rate (recall or sensitivity) and false-positive rate [5, 79] defined by Eq. (19).

$$\text{TPR} = \frac{\text{True}_{\text{pos}}}{\text{True}_{\text{pos}}+\text{False}_{\text{neg}}} \tag{19}$$

Precision represents a measure that determines correctly classified defective instances out of all the defective instances. It can be calculated using Eq. (20):

$$\text{Precision} = \frac{\text{True}_{\text{pos}}}{\text{True}_{\text{pos}}+\text{False}_{\text{pos}}} \tag{20}$$

F-measure combines recall and precision into a single measure and is a harmonic mean of the two, as shown in Eq. (21).

$$F-\text{measure} = 2\left(\frac{\text{Precision. Recall}}{\text{Precision + Recall}}\right) \tag{21}$$

Matthews Correlation Coefficient (MCC) measures the correlation between predicted and actual instances. Equation (22) is used to calculate its value, which falls between [-1,1]

$$\text{MCC} = \frac{\text{True}_{\text{pos}} * \text{True}_{\text{neg}} - \text{False}_{\text{pos}} * \text{False}_{\text{neg}}}{\sqrt{(\text{True}_{\text{pos}} + \text{False}_{\text{pos}}) * (\text{True}_{\text{pos}} + \text{False}_{\text{neg}}) * (\text{True}_{\text{neg}} + \text{False}_{\text{pos}}) * (\text{True}_{\text{neg}} + \text{False}_{\text{neg}})}} \tag{22}$$

**Table 1** Statistical information of software projects

| Datasets | No. of instances | No. of attributes | No. of defective instances | No. of non-defective instances | Defective ratio | Imbalance ratio |
|---|---|---|---|---|---|---|
| JM1 | 7783 | 22 | 1672 | 6111 | 21.48 | 3.65 |
| MC1 | 1988 | 38 | 46 | 1942 | 2.31 | 42.22 |
| KC1 | 2109 | 22 | 326 | 1783 | 15.46 | 5.47 |
| PC3 | 1077 | 37 | 134 | 943 | 12.44 | 7.04 |
| PC4 | 1458 | 37 | 178 | 1280 | 12.21 | 7.19 |
| PC5 | 1711 | 38 | 471 | 1240 | 27.53 | 2.63 |
| Eclipse JDT Core | 997 | 61 | 206 | 791 | 20.66 | 3.84 |
| Eclipse PDE | 1492 | 61 | 209 | 1283 | 14.01 | 6.14 |
| Equinox | 324 | 61 | 129 | 195 | 39.81 | 1.51 |
| Lucene | 691 | 61 | 64 | 627 | 9.26 | 9.80 |
| Mylyn | 1862 | 61 | 245 | 1617 | 13.16 | 6.60 |
| ant-1.7 | 745 | 20 | 166 | 579 | 22.28 | 3.49 |
| camel-1.6 | 965 | 20 | 188 | 777 | 19.48 | 4.13 |
| jedit-4.3 | 492 | 20 | 11 | 481 | 2.24 | 43.73 |
| poi-3.0 | 442 | 20 | 281 | 161 | 63.57 | 0.57 |
| Xalan-2.7 | 909 | 20 | 898 | 11 | 98.79 | 0.01 |
| Xerces-1.4 | 588 | 20 | 437 | 151 | 74.32 | 0.35 |

G-mean represents a balanced measure that is computed using the geometric mean of sensitivity and specificity, as shown in the equation below (23).

$$G - \text{mean} = \sqrt{\text{Sensitivity}} * \sqrt{\text{Specificity}} \tag{23}$$

### 5.4 Parameter settings

For performing feature selection, only the training dataset is used in this paper. The model is evaluated using the same features that were selected from the training dataset. The filter methods—IG, CS, and RF are implemented using the Weka Tool (default settings), and top $\log_2 n$ features are selected. The wrapper-based, oppositional whale optimization algorithm and 1D-CNN model are implemented in Python 3.9 on Spyder 5.2.2. We have divided the data into 70% training and 30% testing before executing the CNN algorithm. Tables 2 and 3 show the parameter settings for OBWOA and 1D-CNN.

### 5.5 Statistical analysis

The study employed statistical tests, mainly the Wilcoxon test, to support and reinforce the findings and conclusions. Demšar [80] has suggested using non-parametric tests because assumptions regarding data distribution, such as normality and homogeneity, are relaxed.

We applied the Wilcoxon signed-rank test [81] to detect significant differences in the techniques pair. With the help of this test, we may determine whether the null hypothesis, suggesting no difference in the performance of techniques can be rejected or not. We compute the p-value for every technique pair and make a decision on either to accept or reject the null hypothesis, with a chosen significance level = 0.05.

## 6 Results and discussion

RQ1. How efficient is the SDP model developed using 1D-CNN with and without multi-filter wrapper feature selection?

The study utilized 17 software defect datasets to implement the 1D-CNN model with and without MFWFS. The model was designed with parameter settings specified in Table 3. The datasets are initially balanced using the random oversampling technique before training the model. The top $\log_2 n$ features are selected from IG, CS, and Relief-F methods applied to the training set. The OBWOA uses KNN (K = 5) as the internal classifier with parameter settings specified in Table 2 Parameter settings for OBWOA. To handle the stochastic nature of OBWOA, we performed 20 runs of this technique to find the average results. Table 4 shows the original number of features in all

datasets, the average number of selected features after performing OBWOA, the average selection ratio, the best fitness value, and number of features in the MFWFS technique. For instance, the JM1 dataset has originally 21 features, we select top $\log_2 n$ features, i.e., four features each from IG, CS, and RF methods, nine features are selected from OBWOA, and after their union—we get 12 features in MFWFS. We can observe from Table 4 that the attribute reduction rate varies from 20 to 77% across all seventeen datasets, resulting in a total average reduction of 50% attributes using the above MFWFS technique. This implies that the new MFWFS technique generates feature subsets with substantially fewer features that produce higher-quality solutions.

The classification performance of the SDP model developed using 1D-CNN defect predictor without and with MFWFS is evaluated across 17 software projects measuring performance using F-measure, AUC, G-mean, and MCC. The results are displayed in Tables 5, 6, 7, and 8. Results show that CNN with feature selection depicts the best F-measure, AUC, G-mean, and MCC values in nine, fifteen, ten, and ten out of seventeen datasets. There is an improvement in values of F-measure (1–8%), AUC (0.1–30%), G-mean (0.2–6%), and MCC (1–23.5%) for all the datasets except for the PDE and camel-1.6 dataset. The performance values of the PDE and camel-1.6 dataset after feature selection did not improve; hence, the feature selection was not helpful for that particular dataset.

To assess how the number of features affects classifier performance, the average F-measure, AUC, G-mean, and MCC values were compared using the Wilcoxon signed-rank test with a significance threshold of 0.05. The test results are presented in Table 9, showing the p-values for each comparison. A p-value below 0.05 suggests a statistically significant difference between the two techniques. The performance difference between the two techniques is denoted as either significantly different (S +) or not significantly different (NS), respectively. The analysis revealed a significant difference between the two methods in terms of AUC (p-value = 0.015) but not for F-measure, G-mean, or MCC.

**Table 2** Parameter settings for OBWOA

| Parameters | Values |
|---|---|
| No. of iterations | 100 |
| No. of search agents (whales) | 50 |
| Spiral factor b | 1 |
| Convergence constant 'a' | 1 |
| No. of runs | 20 |

This clearly implies that the impact of MFWFS on the 1D-CNN classifier is significant with respect to AUC values. Although 1D-CNN extracts high-level deep semantic features using various layers, feature selection using a hybrid filter wrapper approach is applied prior to training the model, removing all the defect features that are irrelevant, redundant, and cause noise in the dataset and overfitting the model. As a result, this integration of static defect feature subsets with abstract deep semantic features in CNN improves the model's efficiency and builds a stronger ability to distinguish between defective/non-defective classes.

RQ2. How efficient is the SDP model developed using MFWFS-1D-CNN with and without attention?

The study evaluated the performance of SDP models created with MFWFS-1D-CNN, both with and without an attention layer, by measuring their F-measure, AUC, G-mean, and MCC values. The results are presented in Tables 10, 11, 12, and 13. Across all datasets, the F-measure, AUC, G-mean, and MCC values range from 0.613 to 0.901, 0.788 to 0.974, 0.761 to 0.958, and 0.381 to 0.897, respectively. Adding the attention layer improves the predictive model's performance in all the datasets except for one, i.e., xalan. We can clearly observe that there is an improvement of about 0.1–11% in the values of F-measure, 0.5–11% in the values of AUC, 0.3–12% in the values of G-mean, and 0.12–35% in the values of MCC measures obtained after applying the attention layer in the 1D-CNN model. The improvement in the values is due to the application of the attention mechanism in the 1D-CNN model, which can detect the direction and position of the data, thus enabling the precise identification of the defects in the software.

**Table 3** Parameter settings for 1D-CNN

| Parameters | Values |
|---|---|
| No. of convolutional layers | 2 |
| No. of dense layers | 2 |
| Activation function | ReLU + Softmax |
| Number of attention layers | 2 |
| Filter | 64,32 |
| Kernel | 1,1 |
| Optimizer | Adam |
| Learning rate | 0.01 |
| Loss function | Binary cross entropy |
| No. of epochs | 250 |
| Dropout | 0.2 |

**Table 4** Average selected features obtained by MFWFS

| Datasets | Original #features | #features (OBWOA) | Selection ratio | Best fitness value (OBWOA) | #features(IG + CS + RF + OBWOA) |
|---|---|---|---|---|---|
| JM1 | 21 | 9 | 0.4286 | 0.410967 | 12 |
| MC1 | 38 | 5 | 0.1316 | 0.40894 | 13 |
| KC1 | 21 | 8 | 0.381 | 0.338598 | 15 |
| PC3 | 37 | 7 | 0.1892 | 0.359309 | 16 |
| PC4 | 37 | 6 | 0.1622 | 0.251322 | 16 |
| PC5 | 38 | 12 | 0.3158 | 0.31292 | 21 |
| EQ | 61 | 4 | 0.0656 | 0.178979 | 14 |
| JDT | 61 | 15 | 0.2459 | 0.22763 | 24 |
| PDE | 61 | 9 | 0.1475 | 0.36215 | 22 |
| Lucene | 61 | 3 | 0.0492 | 0.321484 | 17 |
| Mylyn | 61 | 9 | 0.1475 | 0.342688 | 17 |
| ant-1.7 | 20 | 6 | 0.3 | 0.267295 | 12 |
| camel-1.6 | 20 | 11 | 0.55 | 0.3794 | 16 |
| jedit-4.3 | 20 | 6 | 0.3 | 0.410779 | 12 |
| poi-3.0 | 20 | 9 | 0.45 | 0.15855 | 14 |
| Xalan-2.7 | 20 | 3 | 0.15 | 0.158033 | 12 |
| xerces-1.4 | 20 | 5 | 0.25 | 0.086125 | 12 |

To statistically validate our results, we performed the Wilcoxon test on all four performance measures. Table 14 represents the p-values obtained when the MFWFS-1D-CNN-Attention model is compared with MFWFS-1D-CNN. The p-values obtained with respect to F-measure = 0.03, AUC = 0.041, and MCC = 0.012 show a significant difference in the two approaches but not significantly with respect to the G-mean measure. These results suggest that the MFWFS-1D-CNN-Attention model outperforms the MFWFS-1D-CNN model without attention, signifying the importance of the attention layer in the neural network.

RQ3. How effective are the SDP models developed using the proposed hybrid model MFWFS-1D-CNN-Attention in comparison with the models created with individual filters and wrapper-based CNN model?

Answer: To assess whether the proposed MFWFS-1D-CNN-Attention model is able to effectively classify the software defects, we made a comparison of this model with the sole wrapper method (OBWOA) and combinations of filter (IG, CS, and RF) and wrapper method. Table 15 reports the results using the four performance measures—F-measure, AUC, G-mean, and MCC.

In terms of all the evaluation measures, MFWFS (IG + CS + RF + OBWOA) on CNN attention model produced the highest values in JM1, MC1, KC1, PC1, PC3, PC4, PC5, EQ, JDT, Lucene, Mylyn, ant-1.7, camel-1.6, and poi-3.0 datasets. However, for the jedit-4.3 dataset, the performance of the IG + CS + RF + OBWOA method was better in terms of AUC and G-mean, while the OBWOA wrapper method alone was better in terms of F-measure and MCC scores. For Xalan-2.7, IG + OBWOA provided the best F-measure and MCC value, while OBWOA gave the best performance in terms of AUC and G-mean. For the Xerces-1.4 dataset, IG + OBWOA performed the best, with the highest values for all evaluation measures.

We also present the box plots to visually compare the difference in the prediction performances between OBWOA, IG + OBWOA, CS + OBWOA, RF + OBWOA, and IG + CS + RF + OBWOA over seventeen datasets, with respect to four performance measures in Fig. 3, respectively. The x-axis denotes the different methodologies being compared, and the y-axis denotes the four evaluation measures. We observe that for each of the four evaluation metrics, the median value achieved by IG + CS + RF + OBWOA is higher than that obtained by other approaches. The average predicted F-measure value by IG + CS + RF + OBWOA is 0.794 (for all the datasets), that is, an improvement between 7% (IG + OBWOA) and 10.41% (OBWOA), the average AUC (0.896) by IG + CS + RF + OBWOA obtains improvement between 4.4% (RF + OBWOA) and 7% (OBWOA), the average G-mean (0.875) by IG + CS + RF + OBWOA obtains improvement between 5.4% (IG +

**Table 5** Comparative results of 1D-CNN with and without MFWFS based on *F*-measure

| | JM1 | MC1 | KC1 | PC3 | PC4 | PC5 | EQ | JDT | PDE | Lucene | Mylyn | Ant-1.7 | Camel-1.6 | Jedit-4.3 | Poi-3.0 | Xalan-2.7 | Xerces-1.4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Without FS | 0.614 | 0.66 | 0.726 | 0.764 | 0.782 | 0.731 | 0.854 | 0.859 | 0.846 | 0.825 | 0.732 | 0.79 | 0.746 | 0.69 | 0.802 | 0.831 | 0.909 |
| With FS | 0.651 | 0.653 | 0.721 | 0.776 | 0.841 | 0.739 | 0.898 | 0.901 | 0.839 | 0.776 | 0.753 | 0.789 | 0.708 | 0.706 | 0.796 | 0.677 | 0.9 |

**Table 6** Comparative results of 1D-CNN with and without MFWFS based on AUC

| | JM1 | MC1 | KC1 | PC3 | PC4 | PC5 | EQ | JDT | PDE | Lucene | Mylyn | Ant-1.7 | Camel-1.6 | Jedit-4.3 | Poi-3.0 | Xalan-2.7 | Xerces-1.4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Without FS | 0.712 | 0.878 | 0.847 | 0.915 | 0.916 | 0.806 | 0.871 | 0.902 | 0.912 | 0.847 | 0.833 | 0.805 | 0.826 | 0.935 | 0.825 | 0.754 | 0.895 |
| With FS | 0.752 | 0.888 | 0.848 | 0.919 | 0.94 | 0.813 | 0.932 | 0.926 | 0.89 | 0.861 | 0.875 | 0.833 | 0.826 | 0.88 | 0.833 | 0.98 | 0.942 |

**Table 7** Comparative results of 1D-CNN with and without MFWFS based on G-mean

| | JM1 | MC1 | KC1 | PC3 | PC4 | PC5 | EQ | JDT | PDE | Lucene | Mylyn | Ant-1.7 | Camel-1.6 | Jedit-4.3 | Poi-3.0 | Xalan-2.7 | Xerces-1.4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Without FS | 0.662 | 0.865 | 0.821 | 0.91 | 0.89 | 0.771 | 0.864 | 0.891 | 0.904 | 0.848 | 0.827 | 0.813 | 0.802 | 0.888 | 0.828 | 0.393 | 0.899 |
| With FS | 0.701 | 0.858 | 0.804 | 0.906 | 0.914 | 0.775 | 0.914 | 0.914 | 0.883 | 0.863 | 0.853 | 0.815 | 0.789 | 0.855 | 0.824 | 0.969 | 0.917 |

**Table 8** Comparative results of 1D-CNN with and without MFWFS based on MCC

| | JM1 | MC1 | KC1 | PC3 | PC4 | PC5 | EQ | JDT | PDE | Lucene | Mylyn | Ant-1.7 | Camel-1.6 | Jedit-4.3 | Poi-3.0 | Xalan-2.7 | Xerces-1.4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Without FS | 0.281 | 0.406 | 0.511 | 0.61 | 0.623 | 0.496 | 0.717 | 0.732 | 0.713 | 0.668 | 0.515 | 0.59 | 0.53 | 0.466 | 0.637 | 0.227 | 0.822 |
| With FS | 0.347 | 0.475 | 0.491 | 0.624 | 0.709 | 0.506 | 0.815 | 0.806 | 0.671 | 0.586 | 0.56 | 0.588 | 0.479 | 0.471 | 0.627 | 0.467 | 0.807 |

**Table 9** Wilcoxon test result based on average *F*-measure, AUC, *G*-mean, and MCC

| Comparison approach | *p*-value (*F*-measure) | *p*-value (AUC) | *p*-value (*G*-mean) | *p*-value (MCC) |
|---|---|---|---|---|
| CNN without feature selection/ CNN with MFWFS | NS (0.463) | S + (0.015) | NS (0.193) | NS (0.201) |

OBWOA) and 8.65% (OBWOA), and the average MCC (0.632) by IG + CS + RF + OBWOA obtains improvement between 18% (IG + OBWOA) and 28.78% (OBWOA).

We also compared the performance of the proposed feature selection method IG + CS + RF + OBWOA with transformer-based model BERT [82] used in a previous literature study to validate the performance and robustness of the proposed method against the original 1D-CNN-Attention-based model. The F-measure, AUC, G-mean, and MCC values are reported in Table 16. The results indicate that the **IG + CS + RF + OBWOA with 1D-CNN** model consistently outperforms the **IG + CS + RF + OBWOA with BERT** approach across all datasets in terms of ***F*-measure**, **AUC**, ***G*-mean**, and **MCC** values. Higher F-measure values for 1D-CNN indicate better balance between precision and recall, which is essential for accurate defect prediction. AUC values are also consistently higher for 1D-CNN, indicating stronger discriminatory power and better classification ability. Similarly, G-mean values show that 1D-CNN is better at controlling the trade-off between sensitivity and specificity, resulting in more balanced classification results. Additionally, MCC values are notably higher for CNN which provide insight into the overall quality of predictions taking into account both positive and negative classes. This indicates a stronger correlation between predicted and actual classes compared to BERT. Although BERT like transformers generally perform well in handling complex data patterns and have successful results in existing literature studies, in this specific case, IG + CS + RF + OBWOA with 1D-CNN-Attention combination leads to enhanced performance across these critical metrics, making it a more accurate model for predicting software defects.

In conclusion, we can say that the IG + CS + RF + OBWOA-based CNN attention model behaves consistently over all the considered datasets, mainly performing better than other combinations of the filter and wrapper methods. Since all of the measures—F-measure, AUC, G-mean, and MCC—show significant improvement in the results, we can justify the use of this combination of feature selection techniques for SDP.

**Table 10** Comparative results of MFWFS-1D-CNN with and without attention based on *F*-measure

| | JM1 | MC1 | KC1 | PC3 | PC4 | PC5 | EQ | JDT | PDE | Lucene | Mylyn | Ant-1.7 | Camel-1.6 | Jedit-4.3 | Poi-3.0 | Xalan-2.7 | Xerces-1.4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Without attention | 0.651 | 0.653 | 0.721 | 0.776 | 0.841 | 0.739 | 0.898 | 0.901 | 0.839 | 0.776 | 0.753 | 0.789 | 0.708 | 0.706 | 0.796 | 0.677 | 0.9 |
| With attention | 0.722 | 0.712 | 0.753 | 0.803 | 0.854 | 0.75 | 0.889 | 0.862 | 0.879 | 0.831 | 0.811 | 0.809 | 0.761 | 0.706 | 0.818 | 0.613 | 0.901 |

**Table 11** Comparative results of MFWFS-1D-CNN with and without attention based on AUC

| | JM1 | MC1 | KC1 | PC3 | PC4 | PC5 | EQ | JDT | PDE | Lucene | Mylyn | Ant-1.7 | Camel-1.6 | Jedit-4.3 | Poi-3.0 | Xalan-2.7 | Xerces-1.4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Without attention | 0.752 | 0.888 | 0.848 | 0.919 | 0.94 | 0.813 | 0.932 | 0.926 | 0.89 | 0.861 | 0.875 | 0.833 | 0.826 | 0.88 | 0.833 | 0.98 | 0.942 |
| With attention | 0.788 | 0.933 | 0.868 | 0.924 | 0.938 | 0.835 | 0.93 | 0.926 | 0.937 | 0.9 | 0.901 | 0.837 | 0.85 | 0.974 | 0.838 | 0.948 | 0.891 |

**Table 12** Comparative results of MFWFS-1D-CNN with and without attention based on *G*-mean

| | JM1 | MC1 | KC1 | PC3 | PC4 | PC5 | EQ | JDT | PDE | Lucene | Mylyn | Ant-1.7 | Camel-1.6 | Jedit-4.3 | Poi-3.0 | Xalan-2.7 | Xerces-1.4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Without attention | 0.701 | 0.858 | 0.804 | 0.906 | 0.914 | 0.775 | 0.914 | 0.914 | 0.883 | 0.863 | 0.853 | 0.815 | 0.789 | 0.855 | 0.824 | 0.969 | 0.917 |
| With attention | 0.761 | 0.921 | 0.834 | 0.911 | 0.917 | 0.79 | 0.897 | 0.896 | 0.927 | 0.881 | 0.876 | 0.822 | 0.816 | 0.958 | 0.835 | 0.935 | 0.89 |

**Table 13** Comparative results of MFWFS-1D-CNN with and without attention based on MCC

| | JM1 | MC1 | KC1 | PC3 | PC4 | PC5 | EQ | JDT | PDE | Lucene | Mylyn | Ant-1.7 | Camel-1.6 | Jedit-4.3 | Poi-3.0 | Xalan-2.7 | Xerces-1.4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Without attention | 0.347 | 0.475 | 0.491 | 0.624 | 0.709 | 0.506 | 0.815 | 0.806 | 0.671 | 0.586 | 0.56 | 0.588 | 0.479 | 0.471 | 0.627 | 0.467 | 0.807 |
| With attention | 0.469 | 0.509 | 0.551 | 0.667 | 0.73 | 0.534 | 0.897 | 0.735 | 0.773 | 0.678 | 0.646 | 0.623 | 0.555 | 0.517 | 0.649 | 0.381 | 0.808 |

RQ4. How effective is the SDP model developed using the proposed hybrid classifier (MFWFS-1D-CNN-Attention) in comparison with other state-of-the-art ML and hybrid techniques?

In order to assess the effectiveness of the proposed approach (MFWFS-1D-CNN-Attention), we conducted a comparison of its results with those of other state-of-the-art techniques that are currently available in the literature. We report the results of only those techniques evaluated on the same datasets used in our study and common performance measure (AUC) since it is commonly available in the literature studies. Table 17 provides a list of the abbreviated comparison algorithms along with their particular implementation/parameter values utilized in evaluating the study.

Table 18 compares the proposed algorithm MFWFS-1D-CNN-Attention and other algorithms listed in Table 17, with respect to AUC performance measure. The best values are highlighted in bold. We observe that the proposed approach achieves the highest values of AUC measure in seven datasets (KC1, PC3, camel-1.6, jedit-4.3, poi-3.0, xalan-2.7, and xerces-1.4) out of the ten datasets used for comparison with eighteen algorithms and second best in other three datasets (JM1, PC4, and ant-1.7). For the JM1 dataset, PCA-ANN obtained the highest AUC value of 0.81. For the PC4 dataset, RF and PCA-ANN obtained the best AUC value of 0.97. For the ant-1.7 dataset, ANN obtained the best AUC value of 0.847.

We have also compared MFWFS-1D-CNN-Attention with state-of-the-art models including BERT [82] and CodeBERT [83] transformer models used in the previous studies for software defect prediction in Table 19. The average performance values are recorded across all three datasets—NASA, AEEEM, and PROMISE. Our proposed approach consistently outperformed all other models in terms of F-measure, AUC, G-mean, and MCC, demonstrating its superiority. In case of AEEEM datasets, an impressive F-measure of 0.855 is achieved which is substantially higher than the values obtained by CNN-WSHCKE (0.691), BERT (0.540), and CodeBERT (0.32). Similarly, in case of PROMISE datasets, a remarkable F-measure of 0.768 is obtained, outperforming the F-measure values of CNN-WSHCKE (0.485), BERT (0.652), and CodeBERT (0.306). Furthermore, our approach exhibits notable improvements in AUC, G-mean, and MCC metrics across all datasets compared to the other models.

In a nutshell, we can say that the MFWFS-1D-CNN-based attention model has a powerful ability to address the SDP issue. In order to obtain the best feature subset for each software project, the unique properties of wrapper method OBWOA are integrated with IG + CS + RF filter

**Table 14** Wilcoxon test result based on *F*-measure, AUC, G-mean, and MCC

| Comparison approach | p-value (*F*-measure) | *p*-value (AUC) | *p*-value (*G*-mean) | *p*-value (MCC) |
|---|---|---|---|---|
| MFWFS-1D-CNN without attention/MFWFS-1D-CNN with attention | S + (0.030) | S + (0.041) | NS (0.061) | S + (0.012) |

**Table 15** Performance of different feature selection methods on 1D-CNN

| Datasets | Methodology | F-measure | AUC | G-mean | MCC |
|---|---|---|---|---|---|
| JM1 | OBWOA | 0.632 | 0.734 | 0.682 | 0.315 |
| | IG + OBWOA | 0.632 | 0.73 | 0.68 | 0.314 |
| | CS + OBWOA | 0.616 | 0.712 | 0.659 | 0.279 |
| | RF + OBWOA | 0.638 | 0.73 | 0.688 | 0.325 |
| | IG + CS + RF + OBWOA | **0.722** | **0.788** | **0.761** | **0.469** |
| MC1 | OBWOA | 0.581 | 0.836 | 0.815 | 0.289 |
| | IG + OBWOA | 0.585 | 0.872 | 0.86 | 0.313 |
| | CS + OBWOA | 0.594 | 0.852 | 0.842 | 0.315 |
| | RF + OBWOA | 0.592 | 0.887 | 0.885 | 0.329 |
| | IG + CS + RF + OBWOA | **0.712** | **0.933** | **0.921** | **0.509** |
| KC1 | OBWOA | 0.709 | 0.834 | 0.78 | 0.463 |
| | IG + OBWOA | 0.701 | 0.832 | 0.767 | 0.444 |
| | CS + OBWOA | 0.66 | 0.845 | 0.757 | 0.42 |
| | RF + OBWOA | 0.672 | 0.821 | 0.76 | 0.412 |
| | IG + CS + RF + OBWOA | **0.753** | **0.868** | **0.834** | **0.551** |
| PC3 | OBWOA | 0.691 | 0.827 | 0.824 | 0.469 |
| | IG + OBWOA | 0.697 | 0.856 | 0.849 | 0.512 |
| | CS + OBWOA | 0.721 | 0.85 | 0.854 | 0.523 |
| | RF + OBWOA | 0.732 | 0.89 | 0.868 | 0.547 |
| | IG + CS + RF + OBWOA | **0.803** | **0.924** | **0.911** | **0.667** |
| PC4 | OBWOA | 0.745 | 0.891 | 0.861 | 0.556 |
| | IG + OBWOA | 0.794 | 0.923 | 0.901 | 0.641 |
| | CS + OBWOA | 0.788 | 0.864 | 0.849 | 0.598 |
| | RF + OBWOA | 0.811 | 0.925 | 0.906 | 0.664 |
| | IG + CS + RF + OBWOA | **0.854** | **0.938** | **0.917** | **0.73** |
| PC5 | OBWOA | 0.692 | 0.796 | 0.73 | 0.422 |
| | IG + OBWOA | 0.68 | 0.778 | 0.719 | 0.402 |
| | CS + OBWOA | 0.696 | 0.789 | 0.735 | 0.431 |
| | RF + OBWOA | 0.702 | 0.798 | 0.727 | 0.427 |
| | IG + CS + RF + OBWOA | **0.75** | **0.835** | **0.79** | **0.534** |
| EQ | OBWOA | 0.78 | 0.816 | 0.796 | 0.583 |
| | IG + OBWOA | 0.806 | 0.862 | 0.833 | 0.658 |
| | CS + OBWOA | 0.797 | 0.845 | 0.815 | 0.624 |
| | RF + OBWOA | 0.84 | 0.879 | 0.856 | 0.701 |
| | IG + CS + RF + OBWOA | **0.889** | **0.93** | **0.897** | **0.897** |
| JDT | OBWOA | 0.824 | 0.881 | 0.866 | 0.672 |
| | IG + OBWOA | 0.861 | 0.897 | 0.892 | 0.733 |
| | CS + OBWOA | 0.841 | 0.882 | 0.874 | 0.697 |
| | RF + OBWOA | 0.855 | 0.897 | 0.887 | 0.724 |
| | IG + CS + RF + OBWOA | **0.862** | **0.926** | **0.896** | **0.735** |

**Table 15** (continued)

| Datasets | Methodology | F-measure | AUC | G-mean | MCC |
|---|---|---|---|---|---|
| PDE | OBWOA | 0.734 | 0.827 | 0.805 | 0.505 |
| | IG + OBWOA | 0.805 | 0.846 | 0.848 | 0.624 |
| | CS + OBWOA | 0.799 | 0.861 | 0.853 | 0.619 |
| | RF + OBWOA | 0.778 | 0.841 | 0.843 | 0.583 |
| | IG + CS + RF + OBWOA | **0.879** | **0.937** | **0.927** | **0.773** |
| Lucene | OBWOA | 0.613 | 0.732 | 0.72 | 0.305 |
| | IG + OBWOA | 0.704 | 0.864 | 0.828 | 0.475 |
| | CS + OBWOA | 0.673 | 0.771 | 0.747 | 0.385 |
| | RF + OBWOA | 0.708 | 0.835 | 0.828 | 0.48 |
| | IG + CS + RF + OBWOA | **0.831** | **0.9** | **0.881** | **0.678** |
| Mylyn | OBWOA | 0.773 | 0.864 | 0.849 | 0.578 |
| | IG + OBWOA | 0.757 | 0.863 | 0.829 | 0.546 |
| | CS + OBWOA | 0.751 | 0.851 | 0.835 | 0.54 |
| | RF + OBWOA | 0.738 | 0.837 | 0.821 | 0.517 |
| | IG + CS + RF + OBWOA | **0.811** | **0.901** | **0.876** | **0.646** |
| ant-1.7 | OBWOA | 0.771 | 0.837 | 0.798 | 0.555 |
| | IG + OBWOA | 0.757 | 0.831 | 0.784 | 0.527 |
| | CS + OBWOA | 0.786 | 0.839 | 0.813 | 0.585 |
| | RF + OBWOA | 0.794 | 0.821 | 0.811 | 0.601 |
| | IG + CS + RF + OBWOA | **0.809** | **0.837** | **0.822** | **0.623** |
| camel-1.6 | OBWOA | 0.578 | 0.739 | 0.668 | 0.271 |
| | IG + OBWOA | 0.698 | 0.81 | 0.773 | 0.46 |
| | CS + OBWOA | 0.681 | 0.809 | 0.754 | 0.424 |
| | RF + OBWOA | 0.689 | 0.79 | 0.75 | 0.424 |
| | IG + CS + RF + OBWOA | **0.761** | **0.85** | **0.816** | **0.555** |
| jedit-4.3 | OBWOA | **0.761** | 0.97 | 0.869 | **0.558** |
| | IG + OBWOA | 0.669 | 0.958 | 0.942 | 0.46 |
| | CS + OBWOA | 0.741 | 0.942 | 0.871 | 0.525 |
| | RF + OBWOA | 0.652 | 0.96 | 0.945 | 0.446 |
| | IG + CS + RF + OBWOA | 0.722 | **0.982** | **0.964** | 0.544 |
| poi-3.0 | OBWOA | 0.774 | 0.797 | 0.794 | 0.574 |
| | IG + OBWOA | 0.795 | 0.822 | 0.82 | 0.619 |
| | CS + OBWOA | 0.8 | 0.821 | 0.818 | 0.641 |
| | RF + OBWOA | 0.807 | 0.833 | 0.832 | 0.64 |
| | IG + CS + RF + OBWOA | **0.818** | **0.838** | **0.835** | **0.649** |
| Xalan-2.7 | OBWOA | 0.687 | **0.989** | **0.97** | 0.482 |
| | IG + OBWOA | **0.755** | 0.879 | 0.873 | **0.554** |
| | CS + OBWOA | 0.687 | 0.974 | 0.8 | 0.403 |
| | RF + OBWOA | 0.612 | 0.97 | 0.815 | 0.323 |
| | IG + CS + RF + OBWOA | 0.613 | 0.948 | 0.935 | 0.381 |
| xerces-1.4 | OBWOA | 0.873 | 0.873 | 0.863 | 0.75 |
| | IG + OBWOA | **0.912** | **0.917** | **0.907** | **0.827** |
| | CS + OBWOA | 0.873 | 0.874 | 0.872 | 0.748 |
| | RF + OBWOA | 0.873 | 0.872 | 0.87 | 0.749 |
| | IG + CS + RF + OBWOA | 0.901 | 0.891 | 0.89 | 0.808 |

Bold values represent the highest values among all the techniques for each dataset

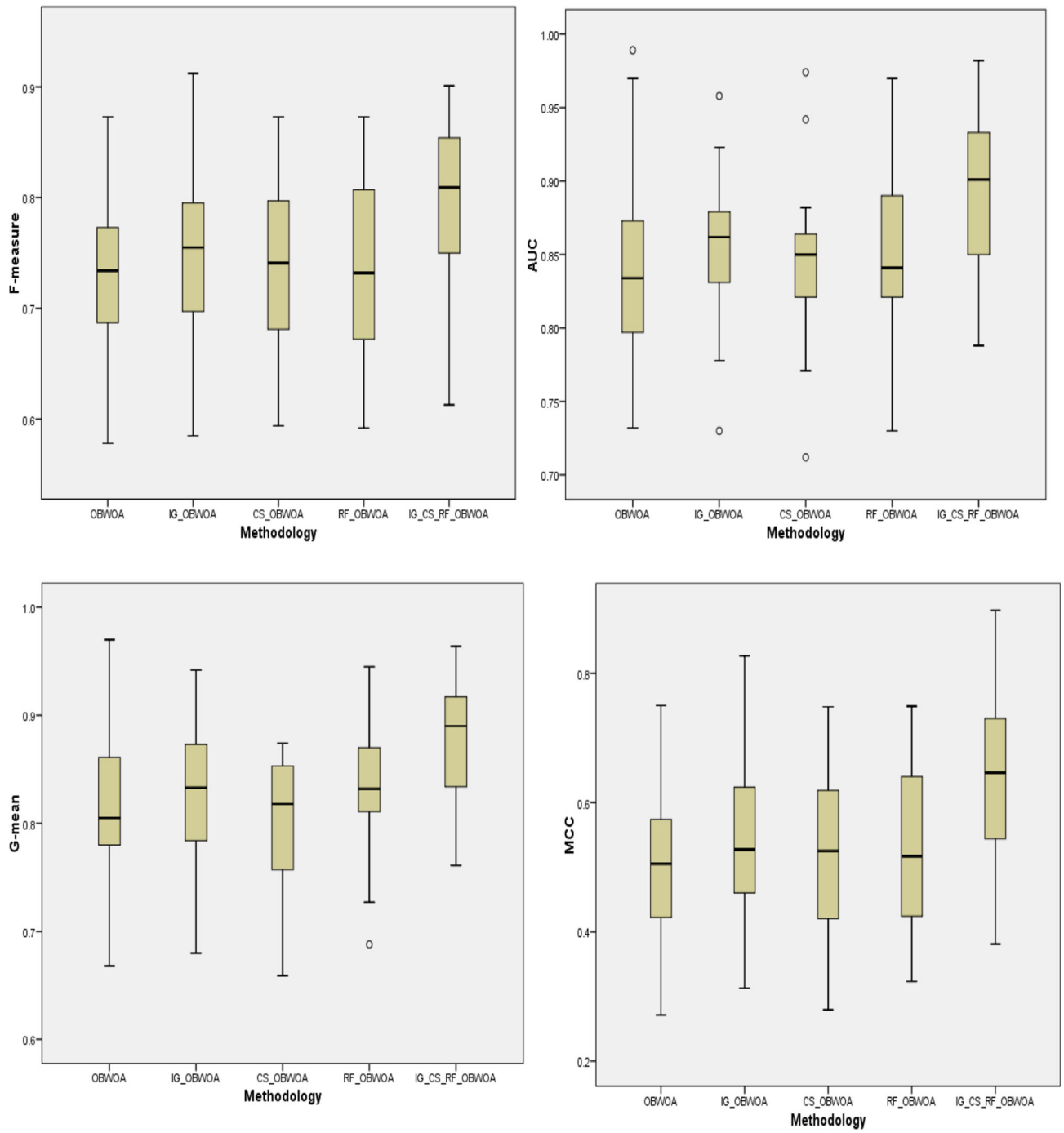



**Fig. 3** Box plot for F-measure, AUC, G-mean, and MCC values for different feature selection methodologies

**Table 16** Comparison of performance of proposed feature selection with 1D-CNN and BERT transformer

| Datasets | Methodology | F-measure | AUC | G-mean | MCC |
|---|---|---|---|---|---|
| JM1 | IG + CS + RF + OBWOA with CNN | **0.722** | **0.788** | **0.761** | **0.469** |
| | IG + CS + RF + OBWOA with BERT | 0.516 | 0.571 | 0.503 | 0.287 |
| MC1 | IG + CS + RF + OBWOA with CNN | **0.712** | **0.933** | **0.921** | **0.509** |
| | IG + CS + RF + OBWOA with BERT | 0.579 | 0.5 | 0.235 | 0.117 |
| KC1 | IG + CS + RF + OBWOA with CNN | **0.753** | **0.868** | **0.834** | **0.551** |
| | IG + CS + RF + OBWOA with BERT | 0.611 | 0.818 | 0.521 | 0.475 |
| PC3 | IG + CS + RF + OBWOA with CNN | **0.803** | **0.924** | **0.911** | **0.667** |
| | IG + CS + RF + OBWOA with BERT | 0.47 | 0.58 | 0.569 | 0.18 |
| PC4 | IG + CS + RF + OBWOA with CNN | **0.854** | **0.938** | **0.917** | **0.73** |
| | IG + CS + RF + OBWOA with BERT | 0.381 | 0.589 | 0.53 | 0.153 |
| PC5 | IG + CS + RF + OBWOA with CNN | **0.75** | **0.835** | **0.79** | **0.534** |
| | IG + CS + RF + OBWOA with BERT | 0.463 | 0.562 | 0.534 | 0.173 |
| EQ | IG + CS + RF + OBWOA with CNN | **0.889** | **0.93** | **0.897** | **0.897** |
| | IG + CS + RF + OBWOA with BERT | 0.677 | 0.792 | 0.673 | 0.509 |
| JDT | IG + CS + RF + OBWOA with CNN | **0.862** | **0.926** | **0.896** | **0.735** |
| | IG + CS + RF + OBWOA with BERT | 0.861 | 0.914 | 0.805 | 0.729 |
| PDE | IG + CS + RF + OBWOA with CNN | **0.879** | **0.937** | **0.927** | **0.773** |
| | IG + CS + RF + OBWOA with BERT | 0.388 | 0.557 | 0.463 | 0.119 |
| Lucene | IG + CS + RF + OBWOA with CNN | **0.831** | **0.9** | **0.881** | **0.678** |
| | IG + CS + RF + OBWOA with BERT | 0.456 | 0.763 | 0.617 | 0.272 |
| Mylyn | IG + CS + RF + OBWOA with CNN | **0.811** | **0.901** | **0.876** | **0.646** |
| | IG + CS + RF + OBWOA with BERT | 0.393 | 0.615 | 0.554 | 0.155 |
| Ant-1.7 | IG + CS + RF + OBWOA with CNN | **0.809** | **0.837** | **0.822** | **0.623** |
| | IG + CS + RF + OBWOA with BERT | 0.452 | 0.752 | 0.621 | 0.304 |
| Camel-1.6 | IG + CS + RF + OBWOA with CNN | **0.761** | **0.85** | **0.816** | **0.555** |
| | IG + CS + RF + OBWOA with BERT | 0.314 | 0.684 | 0.563 | 0.218 |
| Jedit-4.3 | IG + CS + RF + OBWOA with CNN | **0.722** | **0.982** | **0.964** | **0.544** |
| | IG + CS + RF + OBWOA with BERT | 0.65 | 0.98 | 0.703 | 0.393 |
| Poi-3.0 | IG + CS + RF + OBWOA with CNN | 0.818 | 0.838 | 0.835 | 0.649 |
| | IG + CS + RF + OBWOA with BERT | **0.851** | **0.912** | **0.874** | **0.732** |
| Xalan-2.7 | IG + CS + RF + OBWOA with CNN | 0.613 | **0.948** | **0.935** | 0.381 |
| | IG + CS + RF + OBWOA with BERT | **0.643** | 0.874 | 0.723 | **0.548** |
| Xerces-1.4 | IG + CS + RF + OBWOA with CNN | **0.901** | **0.891** | **0.89** | **0.808** |
| | IG + CS + RF + OBWOA with BERT | 0.854 | 0.832 | 0.819 | 0.747 |

Bold values represent the highest values among all the techniques for each dataset

methods significantly outperformed than using them separately. Further, applying a deep learning-based CNN model with an attention mechanism results in the identification of crucial features that improve the performance of SDP.

## 7 Threats to validity

This section presents a comprehensive list of potential threats that could impact the results of the study. Examining these threats is essential to enhance the reliability of the obtained results.

Internal validity threats may arise from unimpeded factors impacting experimental outcomes, such as experimental errors, parameter settings, and implementation errors. Although the experiment process was carried out very carefully, some errors may have gone unnoticed [84].

External validity threat is associated with the generalization and replication of our results. This threat is reduced by applying the proposed model to 17 open-source datasets from PROMISE, NASA, and AEEEM repositories that are publicly available and used in earlier SDP studies. The projects belong to different programming languages, sizes, and application domains. The parameter settings and fitness

**Table 17** Abbreviations of the algorithms used for comparison with the proposed model

| Abbr. | Algorithm | References | Implementation/parameter values |
|---|---|---|---|
| k-NN | k-nearest neighbor | [68, 69] | The number of neighbors was varied between 1, 3, 5, up to 15 |
| SVM | Support vector machine | [68, 69] | Kernel used—RBF, A multilevel grid search was employed to tune kernel width and regularization parameter CC, ranging from log(C) = [− 6, − 5,…,20] |
| RF | Random forest | [68] | Number of trees tuned between 10, 50, 100, 250, 500, and 1000, and the number of attributes selected per tree as [0.5,1,2] × $\sqrt{M}$, where $M$ is the number of attributes in the dataset |
| L-SVM | Lagrangian support vector machine | [68, 69] | No kernel function used, A range from log.C = [6; 5;... 20] has been evaluated |
| LS-SVM | Least squares support vector machine | [68, 69] | Kernel—RBF, A multilevel grid search was employed to tune kernel width and regularization parameter CC, ranging from log(C) = [− 6, − 5,…,20] |
| NB | Naïve Bayes | [69] | Default (Gaussian NB) |
| LDA | Linear discriminant analysis | [68, 69] | Solver: SVD, shrinkage:None |
| C4.5 | Decision tree | [68] | Pruning strategies varied confidence levels from 0.05 to 0.7, with and without Laplacian smoothing and subtree raising |
| ANN | Artificial neural network | [38] | Three-layer network, hidden neurons (10), no. of epochs (300), activation function (sigmoid), learning rate (0.1) |
| CNN-WSHCKE | CNN-whale optimization-simulated annealing-based kernel extreme machine learning | [3] | Number of neurons in convolutional layers: 32, 64, 16, kernel sizes: 3 × 3, 3 × 3, and 4 × 4, activation function: ReLU, KELM hyperparameters: The Gaussian kernel's bandwidth σ |
| SSA-BPNN | Salp swarm algorithm-backpropagation neural network | [16] | SSA: Population size—30, max iterations—300<br>BPNN: three-layer network, No. of hidden layer:1, number of hidden neurons: 2n + 1 |
| PCA-ANN | Principal component analysis-ANN | [69] | PCA uses maximum likelihood estimation,<br>ANN- no. of hidden neurons: 2n + 1,<br>Sigmoid activation function, minimizes MSE during training |
| LRNN | Layered recurrent neural network | [39] | No. of iterations: 1000, No. of input layer neurons: no. of features, No. of hidden layer neurons: number of features/2. Number of neurons in output layer: 1, threshold value: 0.5 |
| BGA-BPSO-BACO-LRNN | Binary genetic algorithm-binary particle swarm optimizaton-binary ant colony optimization-layered recurrent neural network | [39] | BGA—Number of Iterations: 300, Population Size: 40, Crossover Rate: 0.7, Mutation Rate: 0.1, Selection Type: Roulette Wheel Selection, Crossover Type: Single, double, or uniform (randomly selected each iteration)<br>BPSO—Number of Iterations: 300, Swarm Size: 40, Degree of Influence (c1, c2): Both set to 1.5, Inertia Weight (w): 0.8, Maximum and Minimum Velocity (v_max, v_min): 1 and 0, respectively<br>BACO—Number of Iterations: 300,Number of Ants (Agents): 20, Initial Pheromone Level: 1, Pheromone Exponential Weight (α): 0.8, Heuristic Exponential Weight (β): 0.8, Evaporation Rate: 0.6 |
| TBWOA-DT | Tournament selection method with binary Whale optimization algorith-decision tree | [40] | Population Size: 10, Number of Iterations: 100<br>Tournament Size (T): 3 |
| EBMFOV3 | Enhanced binary moth flame optimization | [2] | Population size: 20–50, Maximum Iterations: 100, Transfer functions used: S-shaped and V-shaped |

**Table 17** (continued)

| Abbr. | Algorithm | References | Implementation/parameter values |
|---|---|---|---|
| DP-ARNN | Defect prediction via attention-based recurrent neural network | [43] | Embedding dimension: 30, AST vector length: 2000<br>Bi-LSTM units: 40 per layer, First hidden layer nodes: 16, Second hidden layer nodes: 24, Batch size: 32, Epochs: 20, Activation functions: tanh in the first layer, linear in the second layer, sigmoid in the output layer, Loss function: Binary Cross-Entropy<br>Optimizer: RMSprop |
| CNN | CNN | [43] | Number of filters: 10, Filter length: 5, Fully connected layer nodes: 100 |
| RNN | Recurrent neural network | [43] | Same as for DP-ARNN [43] |

**Table 18** Results of the proposed approach against eighteen other algorithms in respect of AUC

| Techniques | JM1 | KC1 | PC3 | PC4 | ant-1.7 | camel-1.6 | jedit-4.3 | poi-3.0 | xalan-2.7 | xerces-1.4 |
|---|---|---|---|---|---|---|---|---|---|---|
| KNN | 0.71 | 0.7 | 0.77 | 0.87 | – | – | – | – | – | – |
| SVM | 0.72 | 0.76 | 0.77 | 0.92 | – | – | – | – | – | – |
| RF | 0.76 | 0.78 | 0.82 | **0.97** | – | 0.677 | 0.797 | 0.636 | 0.674 | 0.576 |
| L-SVM | 0.73 | 0.76 | 0.84 | 0.92 | – | – | – | – | – | – |
| LS-SVM | 0.74 | 0.77 | 0.83 | 0.94 | – | – | – | – | – | – |
| NB | 0.69 | 0.76 | 0.81 | 0.84 | – | – | – | – | – | – |
| LDA | 0.73 | 0.78 | 0.82 | 0.88 | – | – | – | – | – | – |
| C4.5 | 0.72 | 0.71 | 0.78 | 0.93 | – | – | – | – | – | – |
| ANN | 0.713 | 0.735 | – | – | **0.847** | 0.681 | 0.461 | – | 0.817 | – |
| SSA-BPNN | 0.7 | 0.79 | 0.87 | 0.9 | 0.79 | – | 0.97 | – | – | |
| PCA-ANN | **0.81** | 0.79 | 0.89 | **0.97** | – | – | – | – | – | – |
| LRNN | – | – | - | - | 0.605 | 0.542 | 0.696 | – | 0.443 | – |
| BGA-BPSO-BACO-LRNN | | | | | 0.523 | 0.505 | 0.684 | – | 0.42 | – |
| TBWOA-DT | – | – | – | – | 0.737 | 0.677 | 0.847 | – | 0.796 | – |
| EBMFOV3 | – | – | – | – | 0.762 | 0.658 | 0.75 | – | 0.846 | – |
| DP-ARNN | – | – | – | – | | 0.79 | 0.82 | 0.796 | 0.674 | 0.761 |
| CNN | | | | | | 0.732 | 0.841 | 0.745 | 0.674 | 0.671 |
| RNN | | | | | | 0.766 | 0.842 | 0.764 | 0.654 | 0.73 |
| **Proposed approach** | 0.788 | **0.868** | **0.924** | 0.938 | **0.837** | **0.85** | **0.974** | **0.838** | **0.948** | **0.891** |

Bold values represent the highest values among all the techniques for each dataset

functions are clearly specified to make it easy for researchers to replicate the results [85, 86].

Conclusion validity threat ensures that the drawn conclusions are statistically valid. To address this concern, we use a non-parametric Wilcoxon test to ensure we obtain statistically valid results. Using stable performance measures such as F-measure, AUC, MCC, and G-mean to evaluate results avoids the threat of conclusion validity [87].

Construct validity threats involve the appropriate selection of evaluation metrics. AUC, F-measure, MCC, and G-mean are well-established metrics used in the previous SDP studies. Therefore, this threat should also be acceptable [88, 89].

**Table 19** Average performance values of the proposed approach against other models

| Average | Proposed approach | | | | CNN-WSHCKE | | | | BERT | | | | CodeBERT | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | F-measure | AUC | G-mean | MCC | F-measure | AUC | G-mean | MCC | F-measure | AUC | G-mean | MCC | F-measure | AUC | G-mean | MCC |
| AEEEM | **0.855** | **0.919** | **0.895** | **0.746** | 0.691 | 0.792 | 0.799 | 0.484 | 0.540 | 0.730 | 0.652 | 0.300 | 0.32 | 0.497 | 0.425 | 0 |
| PROMISE | **0.768** | **0.89** | **0.876** | **0.589** | 0.485 | 0.744 | 0.646 | 0.353 | 0.652 | 0.883 | 0.685 | 0.407 | 0.306 | 0.5 | 0.5 | 0 |
| NASA | **0.766** | **0.881** | **0.856** | **0.577** | – | – | – | – | 0.349 | 0.548 | 0.448 | 0.002 | 0.355 | 0.503 | 0.396 | 0 |

Bold values represent the highest values among all the techniques for each dataset

## 8 Conclusion and future work

In conclusion, this study offers a novel hybrid model for SDP utilizing a multiple filter wrapper feature selection (MFWFS) technique which combines the features obtained from three filter methods—Information Gain, Chi-square, and Relief-F and one swarm-based oppositional whale optimization algorithm. It effectively selects fewer optimal features by taking advantage of both the filter and wrapper methods. The features obtained from the MFWFS are integrated with the automatic features generated from the output of attention-based 1D-CNN to construct a powerful unified defect predictor. The hybrid model is evaluated on 17 different software defect datasets using four performance metrics—F-measure, MCC, AUC, and G-mean to validate its effectiveness in SDP. Wilcoxon test is used for statistical analysis to validate our findings.

The results emphasize the significance of feature selection (MFWFS) in SDP so that developers can mainly focus on those software modules that are defect-prone, which helps minimize the software maintenance cost. The attention layer is also able to capture more crucial features that ultimately enhance the classification performance of 1D-CNN. The MFWFS-1D-CNN-Attention model yielded a better SDP model than individual filter and wrapper-based models as well as other state-of-the-art methods from the previous literature studies across all the datasets used in the study.

In the future studies, it would be valuable to evaluate the effectiveness of our SDP models on a wider range of open-source and commercial projects utilizing various programming languages like Python. We can explore automated parameter tuning techniques to optimize the parameter configuration of our models. Additionally, we can extend our implementation to address cross-version SDP problems.

**Author contributions** All authors jointly contributed to the conceptualization and design of the study. Sonali Chawla undertook the data collection and analysis process. Sonali Chawla drafted the initial manuscript, which was subsequently reviewed and commented on by Anjali Sharma and Ruchika Malhotra. All authors participated in the review process and provided final approval for the manuscript.

**Funding** The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

**Data availability** The datasets that support this study are publicly available. They are taken from http://promise.site.uottawa.ca/SERepository/, http://purl.org/MarianJureczko/MetricsRepo, and http://bug.inf.usi.ch/.

## Declarations

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

**Ethical approval** This study does not contain any studies with human or animal subjects performed by any of the authors.

## References


1. Ma Y, Luo G, Zeng X, Chen A (2012) Transfer learning for cross-company software defect prediction. Inf Softw Technol 54(3):248–256. https://doi.org/10.1016/J.INFSOF.2011.09.007
2. Tumar I, Hassouneh Y, Turabieh H, Thaher T (2020) Enhanced binary moth flame optimization as a feature selection algorithm to predict software fault prediction. IEEE Access 8:8041–8055. https://doi.org/10.1109/ACCESS.2020.2964321
3. Zhu K, Ying S, Zhang N, Zhu D (2021) Software defect prediction based on enhanced metaheuristic feature selection optimization and a hybrid deep neural network. J Syst Softw 180:111026. https://doi.org/10.1016/J.JSS.2021.111026
4. Wahono RS, Suryana N, Ahmad S (2014) Metaheuristic optimization based feature selection for software defect prediction. J Softw 9:5. https://doi.org/10.4304/jsw.9.5.1324-1333
5. Liu S, Chen X, Liu W, Chen J, Gu Q, Chen D (2014) "FECAR: A feature selection framework for software defect prediction. In: Proceedings - international computer software and applications conference. IEEE Computer Society, pp 426–435. https://doi.org/10.1109/COMPSAC.2014.66.
6. Abe N, Kudo M (2006) Non-parametric classifier-independent feature selection. Pattern Recognit 39(5):737–746. https://doi.org/10.1016/J.PATCOG.2005.11.007
7. Laura Emmanuella LEA, De Paula Canuto AM (2014) Filter-based optimization techniques for selection of feature subsets in ensemble systems. Expert Syst Appl 41(4):1622–1631. https://doi.org/10.1016/J.ESWA.2013.08.059
8. Llobet E et al (2007) Efficient feature selection for mass spectrometry based electronic nose applications. Chemom Intell Lab Syst 85(2):253–261. https://doi.org/10.1016/J.CHEMOLAB.2006.07.002
9. Dash M, Liu H (1997) Feature selection for classification. Intell Data Anal 1(1–4):131–156. https://doi.org/10.1016/S1088-467X(97)00008-5
10. Hsu HH, Hsieh CW, Da Lu M (2011) Hybrid feature selection by combining filters and wrappers. Expert Syst Appl 38(7):8144–8150. https://doi.org/10.1016/J.ESWA.2010.12.156
11. Sarro F, Di Martino S, Ferrucci F, Gravino C (2012) A further analysis on the use of genetic algorithm to configure support vector machines for inter-release fault prediction. In: Proceedings of the ACM symposium on applied computing, pp 1215–1220. https://doi.org/10.1145/2245276.2231967.
12. Moussa R, Azar D (2017) A PSO-GA approach targeting fault-prone software modules. J Syst Softw 132:41–49. https://doi.org/10.1016/j.jss.2017.06.059
13. Kumar PR, Saradhi Varma GP (2018) A novel probabilistic-ABC based boosting model for software defect detection. In: Proceedings of 2017 international conference on innovations in information, embedded and communication systems, ICIIECS 2017, Institute of Electrical and Electronics Engineers Inc., pp 1–6. https://doi.org/10.1109/ICIIECS.2017.8276059
14. Arar ÖF, Ayan K (2015) Software defect prediction using cost-sensitive neural network. Appl Soft Comput J 33:263–277. https://doi.org/10.1016/j.asoc.2015.04.045
15. Manivasagam G, Gunasundari R (2018) "An optimized feature selection using fuzzy mutual information based ant colony optimization for software defect prediction. Int J Eng Technol 7:456–460
16. Kassaymeh S, Abdullah S, Al-Betar MA, Alweshah M (2022) Salp swarm optimizer for modeling the software fault prediction problem. J King Saud Univ - Comput Inf Sci 34(6):3365–3378. https://doi.org/10.1016/J.JKSUCI.2021.01.015
17. Mirjalili S, Lewis A (2016) The whale optimization algorithm. Adv Eng Softw 95:51–67. https://doi.org/10.1016/J.ADVENGSOFT.2016.01.008
18. Alamri HS, Alsariera YA, Zamli KZ (2018) Opposition-based whale optimization algorithm. J Comput Theor Nanosci 24(10):7461–7464. https://doi.org/10.1166/ASL.2018.12959
19. Balogun AO et al (2020) Impact of feature selection methods on the predictive performance of software defect prediction models: an extensive empirical study. Symmetry. https://doi.org/10.3390/SYM12071147
20. Balogun AO, Basri S, Abdulkadir SJ, Hashim AS (2019) Performance analysis of feature selection methods in software defect prediction: a search method approach. Appl Sci. https://doi.org/10.3390/APP9132764
21. Li J, He P, Zhu J, Lyu MR (2017) Software defect prediction via convolutional neural network. In: Proceedings - 2017 IEEE international conference on software quality, reliability and security, QRS 2017, Institute of Electrical and Electronics Engineers Inc., pp 318–328. https://doi.org/10.1109/QRS.2017.42.
22. Singh Y, Kaur A, Malhotra R (2009) Software fault proneness prediction using support vector machines. Eng Comput Sci 2176(1):240–245
23. Huang CL, Wang CJ (2006) A GA-based feature selection and parameters optimizationfor support vector machines. Expert Syst Appl 31(2):231–240. https://doi.org/10.1016/j.eswa.2005.09.024
24. Ceylan E, Kutlubay FO, Bener AB, (2006) Software defect identification using machine learning techniques
25. Ceylan E, Kutlubay FO, Bener AB, (2006) Software Defect Identification Using Machine Learning Techniques. In: 32nd EUROMICRO conference on software engineering and advanced applications (EUROMICRO06). https://doi.org/10.1109/EUROMICRO.2006.56
26. Manjula C, Florence L (2019) Deep neural network based hybrid approach for software defect prediction using software metrics. Cluster Comput 22(4):9847–9863. https://doi.org/10.1007/s10586-018-1696-z
27. Li Z, Li T, Wu Y, Yang L, Miao H, Wang D (2021) Software defect prediction based on hybrid swarm intelligence and deep learning. Comput Intell Neurosci. https://doi.org/10.1155/2021/4997459
28. Lecun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444. https://doi.org/10.1038/nature14539
29. Anvarjon T, Mustaqeem, Kwon S (2020) Deep-Net: A light-weight CNN-based speech emotion recognition system using deep frequency features. Sensors 20(18):5212. https://doi.org/10.3390/S20185212
30. Shen Y, He X, Gao J, Deng L, Mesnil G (2014) Learning semantic representations using convolutional neural networks for web search. In: Proceedings of the 23rd International conference on World Wide Web, association for computing machinery. Inc, pp 373–374. https://doi.org/10.1145/2567948.2577348.
31. Galvez RL, Bandala AA, Dadios EP, Vicerra RRP, Maningo JMZ (2019) Object detection using convolutional neural networks. In: IEEE Region 10 Annual International Conference, Proceedings/TENCON, Institute of Electrical and Electronics Engineers Inc., pp 2023–2027. https://doi.org/10.1109/TENCON.2018.8650517

32. Kiranyaz S, Avci O, Abdeljaber O, Ince T, Gabbouj M, Inman DJ (2021) 1D convolutional neural networks and applications: a survey. Mech Syst Signal Process. https://doi.org/10.1016/J.YMSSP.2020.107398
33. Avci O, Abdeljaber O, Kiranyaz S, Boashash B, Sodano H, Inman D (2018) Efficiency validation of one dimensional convolutional neural networks for structural damage detection using A SHM Benchmark Data. In: 25th International congress on sound and vibration, Hiroshima, pp 4600–4607
34. Kiranyaz S, Ince T, Hamila R, Gabbouj M (2015)"Convolutional Neural Networks for patient-specific ECG classification. In: Proceedings of the annual international conference of the IEEE Engineering in Medicine and Biology Society, EMBS, Institute of Electrical and Electronics Engineers Inc., pp 2608–2611. https://doi.org/10.1109/EMBC.2015.7318926
35. Avci O, Abdeljaber O, Kiranyaz S, Hussein M, Gabbouj M, Inman DJ (2021) A review of vibration-based damage detection in civil structures: From traditional methods to Machine Learning and Deep Learning applications. Mech Syst Signal Process 147:107077. https://doi.org/10.1016/J.YMSSP.2020.107077
36. Ince T, Kiranyaz S, Eren L, Askar M, Gabbouj M (2016) Real-time motor fault detection by 1-D convolutional neural networks. IEEE Trans Industr Electron 63(11):7067–7075. https://doi.org/10.1109/TIE.2016.2582729
37. Jiang H (2010) Discriminative training of HMMs for automatic speech recognition: a survey. Comput Speech Lang 24(4):589–608. https://doi.org/10.1016/J.CSL.2009.08.002
38. Jin C, Jin SW (2015) Prediction approach of software fault-proneness based on hybrid artificial neural network and quantum particle swarm optimization. Appl Soft Comput J 35:717–725. https://doi.org/10.1016/j.asoc.2015.07.006
39. Turabieh H, Mafarja M, Li X (2019) Iterated feature selection algorithms with layered recurrent neural network for software fault prediction. Expert Syst Appl 122:27–42. https://doi.org/10.1016/J.ESWA.2018.12.033
40. Hassouneh Y, Turabieh H, Thaher T, Tumar I, Chantar H, Too J (2021) Boosted whale optimization algorithm with natural selection operators for software fault prediction. IEEE Access 9:14239–14258. https://doi.org/10.1109/ACCESS.2021.3052149
41. Balogun AO et al (2021) A novel rank aggregation-based hybrid multifilter wrapper feature selection method in software defect prediction. Comput Intell Neurosci. https://doi.org/10.1155/2021/5069016
42. Zain ZM, Sakri S, Ismail NHA, Parizi RM (2022) Software defect prediction harnessing on multi 1-dimensional convolutional neural network structure. Comput, Mater Continua 71(1):1521–1546. https://doi.org/10.32604/cmc.2022.022085
43. Fan G, Diao X, Yu H, Yang K, Chen L (2019) Software defect prediction via attention-based recurrent neural network. Sci Program. https://doi.org/10.1155/2019/6230953
44. Akay R, Akay B (2020) Artificial bee colony algorithm and an application to software defect prediction. In: Nature-inspired methods for metaheuristics optimization, vol. 16, Springer, pp 73–92. https://doi.org/10.1007/978-3-030-26458-1_5.
45. Xu Z, Liu J, Yang Z, An G, Jia X (2016) The impact of feature selection on defect prediction performance: an empirical comparison. In: Proceedings - International symposium on software reliability engineering, ISSRE. IEEE Computer Society, pp 309–320. https://doi.org/10.1109/ISSRE.2016.13
46. Liu H, Setiono R (1955) "Chi2: feature selection and discretization of numeric attributes. In: Proceedings of 7th IEEE International conference on tools with artificial intelligence. IEEE, Singapore, pp 388–391. https://doi.org/10.1109/TAI.1995.479783
47. Kononenko I(1994) Estimating attributes: analysis and extensions of RELIEF. In: Estimating attributes: analysis and extensions of RELIEF. Springer Verlag, Berlin, pp 171–182. https://doi.org/10.1007/3-540-57868-4_57/COVER
48. Ahmed S, Zhang M, Peng L (2014) Improving feature ranking for biomarker discovery in proteomics mass spectrometry data using genetic programming. Connect Sci 26(3):215–243. https://doi.org/10.1080/09540091.2014.906388
49. Chandrashekar G, Sahin F (2014) A survey on feature selection methods. Comput Electr Eng 40(1):16–28. https://doi.org/10.1016/j.compeleceng.2013.11.024
50. Mafarja M, Mirjalili S (2018) Whale optimization approaches for wrapper feature selection. Appl Soft Comput 62:441–453. https://doi.org/10.1016/J.ASOC.2017.11.006
51. Rahnamayan S, Tizhoosh HR, Salama MMA (2008) Opposition versus randomness in soft computing techniques. Appl Soft Comput 8(2):906–918. https://doi.org/10.1016/J.ASOC.2007.07.010
52. Tizhoosh HR (2005) Opposition-based learning: a new scheme for machine intelligence. In: International conference on computational intelligence for modelling, control and automation and international conference on intelligent agents, web technologies and internet commerce (CIMCA-IAWTIC'06), pp 695–701. https://doi.org/10.1109/CIMCA.2005.1631345
53. Wang H, Wu Z, Rahnamayan S, Liu Y, Ventresca M (2011) Enhancing particle swarm optimization using generalized opposition-based learning. Inf Sci (N Y) 181(20):4699–4714. https://doi.org/10.1016/j.ins.2011.03.016
54. Rahnamayan RS, Tizhoosh HR, Salama MMA (2008) Opposition-based differential evolution. IEEE Trans Evol Comput 12(1):64–79. https://doi.org/10.1109/TEVC.2007.894200
55. Ibrahim RA, Elaziz MA, Lu S (2018) Chaotic opposition-based grey-wolf optimization algorithm based on differential evolution and disruption operator for global optimization. Expert Syst Appl 108:1–27. https://doi.org/10.1016/J.ESWA.2018.04.028
56. Malisia AR, Tizhoosh HR (2007) Applying opposition-based ideas to the ant colony system. In: Proceedings of the 2007 IEEE swarm intelligence symposium, SIS 2007, Honolulu, pp 182–189. https://doi.org/10.1109/SIS.2007.368044
57. Wang WL, Li WK, Wang Z, Li L (2019) Opposition-based multi-objective whale optimization algorithm with global grid ranking. Neurocomputing 341:41–59. https://doi.org/10.1016/J.NEUCOM.2019.02.054
58. Ghotra B, McIntosh S, Hassan AE (2017) A large-scale study of the impact of feature selection techniques on defect classification models. In: IEEE International working conference on mining software repositories. IEEE Computer Society, pp 146–157. https://doi.org/10.1109/MSR.2017.18
59. Emary E, Zawbaa HM, Hassanien AE (2016) Binary grey wolf optimization approaches for feature selection. Neurocomputing 172:371–381. https://doi.org/10.1016/j.neucom.2015.06.083
60. Mafarja M, Jarrar R, Ahmad S, Abusnaina AA (2018) Feature selection using Binary Particle Swarm optimization with time varying inertia weight strategies. In: ACM international conference proceeding series, association for computing machinery. https://doi.org/10.1145/3231053.3231071
61. Sumbul G, Cinbis RG, Aksoy S (2019) Multisource region attention network for fine-grained object recognition in remote sensing imagery. IEEE Trans Geosci Remote Sens 57(7):4929–4937. https://doi.org/10.1109/TGRS.2019.2894425
62. Qi X, Li K, Liu P, Zhou X, Sun M (2020) Deep attention and multi-scale networks for accurate remote sensing image segmentation. IEEE Access 8:146627–146639. https://doi.org/10.1109/ACCESS.2020.3015587
63. Li W, Liu K, Zhang L, Cheng F (2020) Object detection based on an adaptive attention mechanism. Sci Rep 10(1):1–13. https://doi.org/10.1038/s41598-020-67529-x

64. Hou Q, Zhou D, Feng J (2021) "Coordinate attention for efficient mobile network design. In: Proceedings of the IEEE computer society conference on computer vision and pattern recognition, IEEE Computer Society. pp 13708–13717. https://doi.org/10.1109/CVPR46437.2021.01350
65. Shepperd M, Song Q, Sun Z, Mair C (2013) Data quality: some comments on the NASA software defect datasets. IEEE Trans Software Eng 39(9):1208–1215. https://doi.org/10.1109/TSE.2013.11
66. Jureczko M, Madeyski L (2010) Towards identifying software project clusters with regard to defect prediction. In: ACM international conference proceeding series, pp 1–10. https://doi.org/10.1145/1868328.1868342
67. Ambros MD', Lanza M, Robbes R (2010) An extensive comparison of bug prediction approaches. In: Proceedings - international conference on software engineering, pp 31–41. https://doi.org/10.1109/MSR.2010.5463279
68. Lessmann S, Baesens B, Mues C, Pietsch S (2008) Benchmarking classification models for software defect prediction: a proposed framework and novel findings. In: IEEE transactions on software engineering, pp 485–496. https://doi.org/10.1109/TSE.2008.35
69. Jayanthi R, Florence L (2019) Software defect prediction techniques using metrics based on neural network classifier. Cluster Comput 22(1):77–88. https://doi.org/10.1007/S10586-018-1730-1/METRICS
70. Li J et al. (2017) Rare event prediction using similarity majority under-sampling technique. In: Soft computing in data science. SCDS 2017 communications in computer and information science. Springer, Singapore, pp 23–39. https://doi.org/10.1007/978-981-10-7242-0_3/COVER
71. Tan M, Tan L, Dara S, Mayeux C (2015) Online defect prediction for imbalanced data. In: 2015 IEEE/ACM 37th IEEE International conference on software engineering. IEEE Computer Society, pp 99–108. https://doi.org/10.1109/ICSE.2015.139
72. Farid AB, Fathy EM, Eldin AS, Abd-Elmegid LA (2021) Software defect prediction using hybrid model (CBIL) of convolutional neural network (CNN) and bidirectional long short-term memory (Bi-LSTM). PeerJ Comput Sci 7:1–22. https://doi.org/10.7717/PEERJ-CS.739
73. Srinivasan K, Fisher D (1995) Machine learning approaches to estimating software development effort. IEEE Trans Software Eng 21(2):126–137. https://doi.org/10.1109/32.345828
74. Wei H, Hu C, Chen S, Xue Y, Zhang Q (2019) Establishing a software defect prediction model via effective dimension reduction. Inf Sci (N Y) 477:399–409. https://doi.org/10.1016/J.INS.2018.10.056
75. Tantithamthavorn C, McIntosh S, Hassan AE, Matsumoto K (2017) An empirical comparison of model validation techniques for defect prediction models. IEEE Trans Software Eng 43(1):1–18. https://doi.org/10.1109/TSE.2016.2584050
76. Zhu K, Zhang N, Ying S, Wang X (2020) Within-project and cross-project software defect prediction based on improved transfer naive bayes algorithm. Comput, Mater Continua 63(2):891–910. https://doi.org/10.32604/CMC.2020.08096
77. Song Q, Guo Y, Shepperd M (2019) A comprehensive investigation of the role of imbalanced learning for software defect prediction. IEEE Trans Software Eng 45(12):1253–1269. https://doi.org/10.1109/TSE.2018.2836442
78. Ghotra B, McIntosh S, Hassan AE (2015) Revisiting the impact of classification techniques on the performance of defect prediction models. In: 2015 IEEE/ACM 37th IEEE International conference on software engineering. IEEE, Florence, pp 789–800. https://doi.org/10.1109/ICSE.2015.91
79. De Carvalho AB, Pozo A, Vergilio S, Lenz A (2008) Predicting fault proneness of classes trough a multiobjective particle swarm optimization algorithm. In: Proceedings - international conference on tools with artificial intelligence. ICTAI, pp 387–394. https://doi.org/10.1109/ICTAI.2008.76
80. Demšar J (2006) Statistical comparisons of classifiers over multiple data sets
81. Wilcoxon F (1945) Individual comparisons by ranking methods. Biometrics Bullet 1(6):80–83
82. Uddin MN, Li B, Ali Z, Kefalas P, Khan I, Zada I (2022) Software defect prediction employing BiLSTM and BERT-based semantic feature. Soft comput 26(16):7877–7891. https://doi.org/10.1007/s00500-022-06830-5
83. Pan C, Lu M, Xu B (2021) An empirical study on software defect prediction using codebert model. Appl Sci (Switzerland) 11:11. https://doi.org/10.3390/app11114793
84. Malhotra R, Khanna M (2018) Threats to validity in search-based predictive modelling for software engineering. IET Software 12(4):293–305. https://doi.org/10.1049/IET-SEN.2018.5143
85. Kondo M, Bezemer CP, Kamei Y, Hassan AE, Mizuno O (2019) The impact of feature reduction techniques on defect prediction models. Empir Softw Eng 24(4):1925–1963. https://doi.org/10.1007/S10664-018-9679-5/METRICS
86. Lu H, Kocaguneli E, Cukic B (2014) Defect prediction between software versions with active learning and dimensionality reduction. In: 2014 IEEE 25th International symposium on software reliability engineering. IEEE Computer Society, pp 312–322. https://doi.org/10.1109/ISSRE.2014.35
87. Malhotra R, Lata K (2021) An empirical study to investigate the impact of data resampling techniques on the performance of class maintainability prediction models. Neurocomputing 459:432–453. https://doi.org/10.1016/J.NEUCOM.2020.01.120
88. Li M, Zhang H, Wu R, Zhou ZH (2012) Sample-based software defect prediction with active and semi-supervised learning. Autom Softw Eng 19(2):201–230. https://doi.org/10.1007/S10515-011-0092-1/METRICS
89. Zhou T, Sun X, Xia X, Li B, Chen X (2019) Improving defect prediction with deep forest. Inf Softw Technol 114:204–216. https://doi.org/10.1016/J.INFSOF.2019.07.003


**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.

*MĀPAN-Journal of Metrology Society of India*
https://doi.org/10.1007/s12647-024-00798-3

ORIGINAL PAPER

# Stimulus Towards ESG in India and GHG Protocol with Measurable Metrological Tools: Progresses and Challenges

**S. Chourasia[1]* and S. M. Pandey[2]**

[1]Department of Mechanical Engineering, Delhi Technological University, Delhi, India

[2]Department of Mechanical Engineering, National Institute of Technology, Patna, India

Received: 16 August 2024 / Accepted: 03 December 2024

© Metrology Society of India 2025

**Abstract:** The up-surging of environmental, economic, and social-related matters has changed the view of corporate stakeholders' responsiveness towards profit. They have started thinking about the planet, people, and their wealth besides profit. Consequently, high hope is also focused on Environmental, Social And Governance (ESG) factors. The commitment towards the environment and social responsibility of the organization imitates the ESG performance of enterprises that are corporating ESG activities in their working style with a vigorous governance structure; furthermore, the Covid pandemic has exaggerated the importance of ESG parameters. Corporate stakeholders are fervent in knowing how ESG practices can be embraced into business policy, bringing more clarity and transparency to ESG reporting disclosures. The present study offers insight into the ESG concerns unfolding in terms of their prominence and the existing situation in India. In this study, we briefly discuss the need for ESG in Indian organizations and its transformation landscape for the future. Then, we discuss ESG adoption preparedness and its challenges in India. Then, we discussed the steps to be taken to achieve the aim of the sustainability strategy and its implementation, including the ESG Strategy, its linkage of ESG with GHG protocol with measurable metrology tools & the role of metrology in measuring Scope-1, Scope-2, and Scope-3 its emissions measurement with metrology measuring instruments. In the next section, we discuss the need for sustainable ESG reporting, its associated trends and challenges, and India's current status in ESG funds, ESG investing, and its current challenges.

**Keywords:** ESG; ESG reporting; BRSR; ESG rating; ESG investment; GHG protocol; Metrology

## 1. Introduction

For the last many years, industries have been focusing on production, gratifications of customer demand, services, and profit earned; only the business is called to be successful if it can generate profit from the refined products society needs at minimum cost. The benchmark for calculating the business's success was profit-based, i.e., how much profit the business and its owners were entirely focused on earning rather than their business impact on the environment and society. But nowadays, the customers and government operculum have changed towards the business process. Now, businesses are expansively responsible for environmental and social impacts, not merely for old economic and financial impacts. Currently, triple bottom line and stakeholder capital are used to take the responsibility of present-day business owners and managers toward multiple stakeholders, not just as investors [1]. This triple line concept is the business concept where businesses measure their environmental and social impact with financial performance rather than concentrating on only producing profit[2]. The three pillars of ESG are shown in Fig. 1. The world is getting prospects and many challenges in adopting ESG. During the COVID-19 pandemic, the Indian stock market fell in a fortnight, and environmental, social, and governance issues increased globally [3]. As a result, the discussion on sustainability and its inclusive development has started and become a burning topic for businesses, governments, and customers. In answerable to enhancing the exhaustive progress in problems related to the environment, society, and economy, the international government, organizations, and the entire world have come forward to develop the sustainable action plan for the

*Corresponding author, E-mail: shubhangichourasia7@gmail.com

Published online: 08 January 2025

welfare of countries like ESG to develop a sustainable and complete framework for human society[4].

For many years, the formation of strategic plans for businesses and governments became a top priority for the global community after the launch of UN SDGs (Sustainable goals) in 2015. Sustainable development can be defined as fulfilling today's requirements by protecting the earth and its life-supportive and balancing system[5]. Several authors gave different definitions of sustainability that could be putatively accepted across several disciplines.[6]. These principles are illustrated as follows based on different subsections:

(i). Dimensions of sustainability and their balance: It is a multidimensional concept that works on three dimensions: Environmental, Social, and Economic. Refrain from undermining each dimension to balance all three dimensions properly [7].

(ii). As per the view of stakeholders, This principle is defined by identifying, evaluating, and managing the need for future and present use by seeing today's decisions' impacts on the future by maintaining a proper balance between short and long-term goals[8].

(iii). Stakeholder Approach: In the stakeholder approach, sustainability implicates the needs of stakeholders and their outlooks on future needs.

(iv). Life cycle thinking approach: Sustainability involves managing the bearings of the organization's upstream and downstream doings by embracing the life cycle approach by considering environmental, economic, and social responsibilities.

The concept of ESG was initially proposed in 2004; the concept was enthused by developed countries such as America, Europe, and many more. Before 2004, environments such as (energy, water use, climate change, and carbon emissions), society (Human rights, product safety, health, security, and fair trade rules), Corporate governance(Bribery policy, reporting, disclosure, shareholder protection, board independence) was not part of ESG[9]. The first introduction came to light in the UN report on the global compact corporation with the World Bank heading "Who Cares Wins: connecting Financial Markets to a Changing World," for promoting and improving the condition of the environment, society, and governance by establishing an ESG evaluation and validation system, ESG reporting, ESG disclosure standards, and ESG indexing are promoted[10]. Table 1 shows the framework of ESG. The measurement of ESG has no perfect agreement on the fixed list of concerns and their materiality, but the concern is that it can hamper the value of company creation. These concerns are raised due to surging company values with its tangible assets[11] because of no formal definitions. Although ESG is not an immobile concept, apart from the three terms (Environment, Social, Governance), most ESG metrics have distinguished applications employed to develop corporate reports, different rankings, sustainable rankings, several group rankings, and many more[12].

In other words, ESG is defined by the different factors employed to quantify the non-financial impacts of individual investments and businesses/companies. At oncomes, ESG also allows businesses to invest in it [13].ESG contains three components: Environment, Social, and Governance. Meanwhile, ESG factors are co-related because

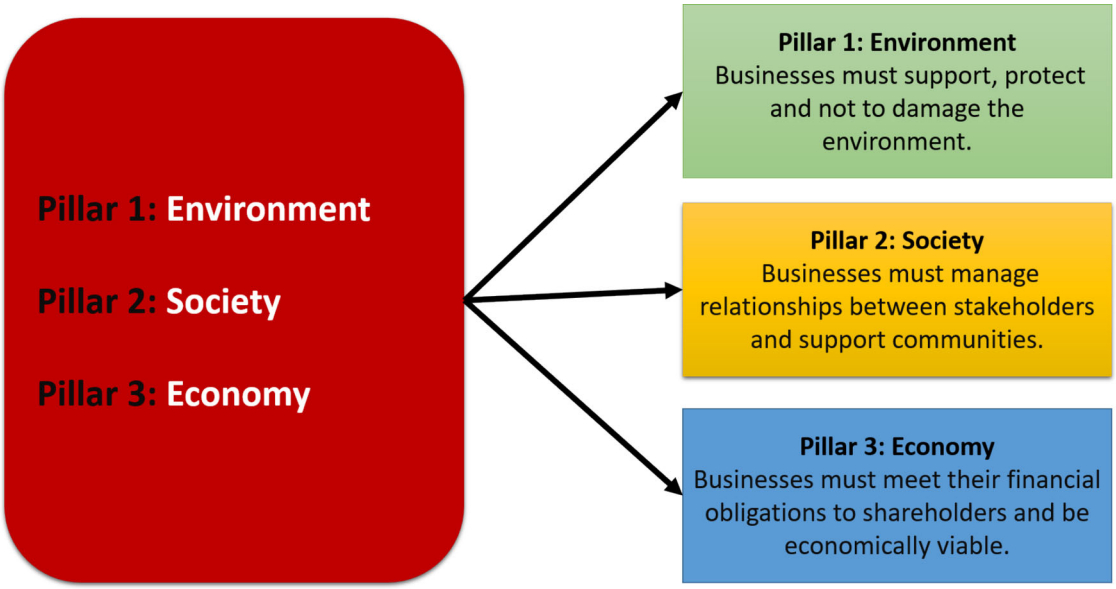



**Fig. 1** Shows Three Pillars of ESG

**Table 1** Shows the framework of ESG [11]

| ESG pillars | Factors |
|---|---|
| Environment (E) | GHG Emission |
| | Innovation in eco-friendly products |
| | Water usage and recycling |
| | Impact and dependency on biodiversity |
| | Air Pollution |
| | Waste Production and Management |
| | Impact on the Ecosystem |
| Social (S) | Workforce Freedom of Associations |
| | Community Impacts |
| | Supply Chain Management |
| | Child Labour |
| | Customer Health and Safety |
| | Training and Education |
| | Discrimination |
| | Work Place Health And Safety |
| Governance(G) | Code of conduct and Business Principles |
| | Shareholder Rights |
| | Board Diversity and Structure |
| | Executive Pay |
| | Transparency and Disclosure |
| | Accountability |
| | Bribery and Corruption |

sometimes, measuring one component of ESG is not feasible for companies.

1. Environment: The first component of ESG focuses on the conservation and preservation of renewable sources of the world; ESG helps the company identify whether their manufacturing comes under the scale of adoption for carbon footprint and eco-friendly methods. ESG components typically include air and water pollution, water stress, waste management, carbon emission, carbon footprint, climate change, and natural resource conservation[14].
2. Social: This aspect of ESG focuses on social features surrounding the people of the business/company and the relation between them. The life of any business is the people or employees who work in the organization; socially developed organization greetings to human development and growth. This component shows the company's relationship with its employees, customers, community, and business associates. This ESG component comprises customer satisfaction, labor laws, employee engagement, human rights diversity and equality, data protection, and gender[15].
3. Governance: This ESG component clarifies the compulsory standards to run an organization. The primary function of this component is to decide the board and management's drives to bring positive change in the organization with more transparency, ethically with governance consistently. The management caters to this component's stakeholders, shareholders, customers, and employees' needs. Apart from this, this component includes the development of the board of directors, whistle-blowing policy, executive compensation, bribery and corruption, etc.[16].Thus, the following question is asked in the present research work.

**Query 1:** To study the need for ESG in Indian organizations and its transformation landscape for the future.

**Query 2**: To study ESG adoption preparedness and its challenges in India.

**Query 3:** To study the ESG Strategy, the role of metrology in the implementation of ESG, its linkage of ESG with GHG protocol with measurable metrology tools & role of metrology in measuring Scope-1, Scope-2, and Scope-3.

**Query 4:** To study the present status of ESG reporting in India with different ESG disclosure practices.

**Query 5**: To study the need for sustainable ESG reporting and its associated trends and challenges.

**Query 6**: To study the need for ESG reporting, the current state of ESG reporting in India, trends, and challenges.

**Query 7:** To study India's current status in ESG funds, ESG investing, and its current challenges.

## 2. Need of ESG in Indian Organisation

In the present scenario, the ESG regulations are gradually explaining the path where businesses/companies can espouse their operations [6] with no economic risks, more health safety, and a healthy environment [17]. Now, the journey of ESG has moved from an unclear concept to a business and investing priority. ESG helps companies/businesses to increase the quality, reliability, and operations of business[18]. Starting companies may have to abide by initial ESG investing costs, but in the long run, the company and investor will receive a higher economic return. [19]. Companies should give proper attention to ESG; they should not neglect it as it can lead to enormous financial repercussions, besides the breach of public trust and ESG risks in the ESG process. These risks can be physical risks like ruining company infrastructure and properties that could suffer from natural disasters, multiplying the financial risks in terms of losses and liabilities.[17]. Due to this negligence, they may be fined because of non-obeying environmental and social laws [20]. On the ending note, the entire reputation of a company can be laid

at risk in any situation discussed above [21]. Therefore, uniting ESG into the business/company corporate structure is most important.

Companies' topmost responsibility is following ESG regulations to minimize environmental, social, and economic risks. Among all the risks, the GCC (Global Climate Change) is the utmost risk for many businesses and companies [22]. The thrilling weather conditions lead to natural disasters. As per the Brown to Green Report published in 2019, India is among the five G20 nations that experience high economic losses due to awful weather conditions [23]. According to one report, Kerala has suffered from adverse flooding in its long history and has reported approximately a loss of USD 4.25 billion. The second most affected state is Bihar, where about 7.2 million people are severely affected by floods, collapsed infrastructure, and worse agricultural activity[13]. The wickedest impact of these disasters increases with time because of companies' carelessness towards the government regulations implemented by authorities [19]. The increasing level of sea results in floods, and because of this, India has to face many natural disasters because of climate change, which is one of the buckets of ESG. For instance, In 2014, the north India-situated bottling plant of Coca–Cola was bidden to close by the state government of Uttar Pradesh after receiving complaints from farmers. According to a farmer who complained, the Coca-Cola bottling plant was using too much water. This was the 2nd time the Coca-Cola company bottling plant got the order to close in India; earlier, it was also closed in the South Indian state of Kerala. In 2020, because of full rain and flood in Assam and WB (West Bengal), a 10% decrement in tea production was noted due to climate change[24]. That's why there is a need to consider the fast track to save the entire ecosystem, and industries, Government, academia, and individuals must come together and consciously try to mitigate climate risks. It is better to get things done late than never; luckily, the Government has now started paying attention across the globe to handling the issue concerning Global Warming [25].

### 2.1. ESG Transformation Landscape to Be Future-Ready

The world felt the first sustainability crisis post-pandemic, which has shaken the globe and resulted in high investment in ESG. According to the Ernst & Young report, approximately 86% of investors are investing in companies focusing on decarbonization, and 90% of the investors see the ESG performance of the company for their investing purpose. In India, the condition is in the starting phase where only 25 out of 5180 investors are becoming signatories in UNPRI (United Principles of Responsible Investing). During the Paris visit of Honourable Prime Minister Narendra Modi, he committed to attaining zero-emission goals by 2070[26].To fulfill the "0" emission goal, the corporate group of India must assimilate ESG principles into their business to protect the environment and raise the interest of several stakeholders. In India, GOI has maintained a sharp vigilance with companies emitting high carbon. To keep this in view, SEBI (Securities and Exchange Board of India) also developed ESG disclosures that are compulsory for all companies or businesses under 1000 listed companies under its BRSR reporting start-up. In India, it is compulsory for CSR (Corporate Social Responsibility) for businesses/companies whose net profit is Rs.50 million, Rs.10 million turnovers, and 5 billion net worth to file the BRSR reporting to show their status towards sustainability. These companies strictly spend at least 2% of their profit on CSR activities for the welfare of society(social) and also file ESG reporting to facilitate capital from ESG investors globally [15].

In the current state, the Indian organizations have started to file several reports with ESG performance conforming with globally ascribed frameworks such as GRI(Global Reporting Initiatives), TCFD(Task Climate Force Disclosure), and IR(Investment Responsibility), even here the non-listed companies are also voluntarily disclosing their ESG reporting in BRSR-LITE format. [27]. Some global investors have already gone one step ahead and adopted clear ESG policies in their hard work and investment, observing the process. They capitalize on prospects to endorse environmentally impactful investing and environmental sustainability. On other prospects, investors are investigating companies that are good in ESG and poor in ESG and avoid those entities that are poor in ESG parameters. India's corporate system is still in the budding stage of optimizing its changing strategy, ESG profiles, and financing requirements[28].

## 3. ESG Adoption Preparedness in India

The big institutional investors (an organization that invests money on behalf of other people) identify the influence of ESG on the long-term feasibility of businesses, enhancing the integration of ESG areas into their investment system and pushing businesses to disclose their ESG-related risk. Though many advantages of a robust corporate ESG program are undeniable, Indian companies face numerous challenges in building one. There are some critical questions in front of India for adopting the ESG where SEBI (Securities and Exchange Board of India) recently published a conference paper on ESG (Environment, Social, and Governance) and EGS ratings to understand securities markets. SEBI clears in this report that companies are

passing through high pressure to focus on integrating ESG in their business practices and growing many expectations with companies transitioning towards a sustainable environment for social and economic business activities. This report is aligned with observations developed by the WEF (World Economic Forum), where the list of the world's biggest threats was mentioned in the list of 2021 i.e. climate risks such as bad weather, drought, and human damage become the biggest fear in front of the world by chance and impact. Despite increasing awareness of ESG, some obstructions exist to its adoption in India. Before implementing ESG in India, companies have to clear their mind with the following questions:

**Q.1 Are Indian businesses considered suitable material issues relevant to their stakeholders?**

It has been seen that ESG issues keep changing from company to company. The material issues should be a company focus because they directly impact their stakeholders, influence society, and subsidize the bottom line. The first step of ESG adoption is the materiality assessment. This assessment is conducted by assessing the needs of customers and stakeholders and providing a complete blueprint of the organization's ESG strategy for the future. When these company-specific materials are recognized through supplier meetings, stakeholder engagement, and many more, management must set goals, develop a unified ESG strategy, and start working to achieve the goals. In India, it is crucial to make these strategies through leadership and management by inspiring them to include ESG strategies in daily manufacturing operations. When the ESG strategies and their road map are complete, it is vital to increase communication with stakeholders to provide them with proper and authenticated information with updates [29].

**Q2. Are Indian companies capable of gathering data that is compulsory to report on key material topics?**

At present, there is a lack of data on Indian industries. Data is vital for ESG adoption for the organization because data is the only information through which the government of India keeps the organization track and report on ESG issues. Due to the need to fix standard rules for ESG, reporting usually needs to be revised to collect and recognize data that can form dynamic disclosure. Many companies, not only in India but also outside India, find it very difficult to gather and update ESG details in their domain operations, mainly when companies are collecting data on scope three emissions, which is entirely dependent on a collection of information from upstream to downstream cradles which are outside the control of the company. This fact was again emphasized interiorly by research conducted by SG analytics, where they found that only 20% to 30% of ESG data for ESG KPI was published online[29].

**Q3. Can India's small and medium-sized businesses adopt ESG carrying costs after the pandemic?**

It is a conception that only mega companies can adopt ESG reforms due to their heavy pocket because ESG demands higher investment in assimilating and reporting on ESG topics. After the COVID-19 pandemic, businesses in India are suffering and struggling to gratify the ESG wants of their stakeholders in the scarcity of well-trained staff and suitably trained staff for ESG departments. Small-scale and mid-scale industries are past a cash shortage, and these companies need to be more mature and require money to develop their business. That's why the BRSR reporting is mandatory for 1000 listed companies whose annual turnover is more from financial year 22–23 and not compulsorily(volunteering) to all[29].

**Q4. Do companies in India have a suitable set of skilled ESG staff and professionals?**

There is no doubt that in India, ESG is growing with good speed and searching for professionals and qualified people in the ESG field. Earlier, the only need of the ESG staff was data collection. Still, it has become complex to deal with several stakeholders, keeping them current concerning ESG with continuously changing regulations, assigning money, and many more of the diverse functions of ESG professionals[29].

In ESG, training on ESG and ESG education for staff and management needs to be improved. To speed up the ESG program, companies must require highly skilled ESG professionals capable of working from an operational business perspective. The readiness of skilled professionals with their teams is still a big concern in India. In simple words, in India, there is a need for more experienced ESG professionals who can improve the ESG performance of a company. The ESG preparation will continue to be a noteworthy teamster of stakeholder satisfaction, whether it be investors, employees, regulators, or customers. The ESG in Indian organizations can not be ignored in long-term value, and it is supreme that they have to overcome these hindrances to stay in step with the constantly changing ESG landscape [30].

### 3.1. ESG Adoption Preparedness in the European Union(EU)

It has been observed that India has no binding laws or rules for the companies in ESG, except the top 1000 listed companies by market capitalization [31]. The European Union's parliament has made strict rules for ESG implementation, and action has been taken to make it part of the EU economy. In this section, we will analyze India's current situation in ESG and the European Union and how it can be implemented in India. The European Union Commission has developed an action plan for sustainable

financing [32]. According to the EU Commission, sustainable financing is the process of considering environmental and social areas while making the investment decision, which helps enhance the investment in the long term and sustainable actions that incorporate ESG investing[33]. The sole target of the EU action plan is to develop a working framework of ESG investment to enhance sustainable investing in the EU and also create a sustainable EU framework that imparts environmental, social, and factors elements into the European economy, which should have the capability to give Europe greener and efficient economy system[19].

The European Commission appointed HLEG (High-Level Expert Group) on sustainable financing. Their last report mentioned the need to increase finance-sustainable progress that can be attained by the long-term needs of society and by reinforcing financial stability by integrating ESG factors into investment-making decisions [34]. Europe has also projected to create a combined EU classification system (EU Taxonomy) and EU ECO labels and establish the green bond [34]. In this taxonomy, the EU target is to elucidate which type of activities will be considered sustainable[35]. This new EU taxonomy would be employed as the foundation of other additional features of the EU action plan, such as ECO labels 79. This classification would aid investors and stakeholders in measuring the earnings of companies created using sustainable activities. Initially, EU Taxonomy would be voluntary, but even though it would be voluntary, expectations of higher participation of companies would help the financial instruments of companies with sustainable credentials that help the investors and stakeholders. These sustainable taxonomy classifications are more challenging and will take time because the EU faces many difficulties in developing a fully functional EU sustainability taxonomy that covers environmental, social, and climate aspects[35]. That's why the EU Commission has taken a step-by-step approach to starting the EU sustainable taxonomy on climate change mitigation and other environmental-based activities. The remaining part of the taxonomy would be covered after the starting steps were completed. This EU taxonomy started on 2nd July 2020 in Europe. In Europe, companies must submit two extra reports, such as Non-Financial Directives and Sustainable Financial disclosure regulations[36]. The aim of reporting in the EU is to assist and encourage the sustainable investment of companies that confirms the investors for company transparent activities that help them make their investing decisions.EU taxonomy helps investors understand what proportion of company revenue comes from sustainable activities, how many green activities, and how ECO labels are provided to them[37]. The European Commission Taxonomy, with ECO labels, fixed the verge for a company where it helps the company pursue ESG practices[35]. Companies would be considered "GREEN" or "SUSTAINABLE" when the company's activities cover the above-said statements. The sustainable financial market would be protected by developing EU standards & labels and providing easy access to investors interested in these products.

The next step to be introduced by the EU in their action plan is the EU Green Bond Standard (GBs) [32]. Green bonds are also called climate bonds. The fixed-income financial bond is used in green bonds to fund projects with positive environmental and climate benefits[38]. When an issuer floats a green bond, no regulations require the issuer to specify the "green" purpose of floating the bond, and there is no periodic reporting requirement. The EU GBS(Green Bond Standard) is a regulation regarding the prospectus that will include assistance on a green bond standard 96, and EU Taxonomy will also be linked to the EU GBS97.

### 3.2. Comparative Study of ESG Adoption in India and ESG Adoption in EU

When we compare the ESG in India and ESG in the EU, the EU has announced mandatory rules and regulations for ESG implementation and reporting requirements to fight against climate change and for bringing significant sustainable financing. In India, it has been seen that in 2017, the SEBI circulated a disclosure for listing Green Debt Securities. After that, in 2020, SEBI released the BRSR (Business Responsibility & Sustainability Reporting) reporting, a form of ESG reporting that only has to be filled by the top 1000 companies by market capitalization to disclose sustainable factors, but it was made mandatory only for showing their company transparency to the investors. This reporting is still nebulous because it needs to fix the standard through which the green and sustainable activities of the company can be identified. The current reporting only targeted the disclosures of companies' sustainable activities but did not encourage the company to incorporate the ESG criteria into the companies' practices. But now, due to minimal regulations in ESG investing, fewer regulations in ESG governance would discourage investors from arriving in the Indian markets. That is why Indian Legislation needs to permit mandatory laws on ESG so that a high source of finance can flow into the country to help promote sustainability among companies.

## 4. Present Status of ESG & BRSR Reporting in India

With the increased drive for sustainability worldwide, the corporate plan has been detected in India over the years. Yet, the face of India is changing and tends to become

highly comprehensive towards the social needs and their expectations, along with the environmental point of view. Many years back, SEBI (Securities and Exchange Board of India) made the integrated reporting framework mandatory for businesses; 10 top companies in India filled out this report to show their business ESG performance. However, many Indian companies have been publishing ESG reporting at a mandatory and voluntary level, following the standard SEBI framework from the 2000s, and can be steered rigorously in the future. The MCA (Ministry of Corporate Affairs) issued the NVGs (National Voluntary Guidelines) business responsibilities towards social, economic, and environmental in 2011. This reporting helped the business deliver busy business-related formation in a simple and all-inclusive way by highlighting that In ESG disclosure, "E" emphasizes the environment-related issues of the business non-financial discoursers. Presently, the most significant challenges faced by the world are global warming, waste, pollution, and the diminution of non-renewable resources. To see these challenges, most countries of the world have dedicated themselves to the reduction of their GHG (Green House Gases), which were discussed in the Paris Agreement and recently discussed in the UNCCC (United Nations Climate Change Conference (COP 26) in 2021. Grounded on these guidelines, ISMR (Indian Stock Markets Regulator), i.e., SEBI, mandated the BRR (Business Responsibly Report) to the top 100 Indian companies in 2012, which stretched to 1000 companies in 2019. Seeing upraised worries about the quality and reliability of received information through BRR, SEBI circulated the circular on 10 May 2021 that made ESG reporting mandatory, called BRSR(Business Responsibility and Sustainability Reporting). The Government also accomplished plugging the gaps in the accuracy and depth of reporting sustainability. The critical differences between BRR(Business Responsibility Reporting) and BRSR are discussed in Fig. 2 [31].

BRSR reporting was voluntary for the financial year 2021–22, but in FY-2022–2023 [39], SEBI introduced the sustainability parameter in BRR(Business Review Report), and it became BRSR(Business Responsibility and Sustainability Report), making it mandate by reporting and its disclosure for 1000 listed companies on an annual basis. The BRSR reporting works on the guidelines of NGRBC (National Guidelines for Responsible Business Conduct) and is also proposed to be prolonged to all businesses (companies) by including LLPs (Limited Liability Partnerships) in the coming five years. Limited Liability Partnerships work on flexible legal and tax entities, which allow partners to profit from economies of scale by working together to reduce their liability toward other partners' actions. Table 2 shows the nine principles of BRSR.

The BRSR is wider than the BRR. The BRR is based on ESG discloser requirements, which almost works on nine sustainability principles issued by MCA ESG Guidelines and needs more meaningful ESG data. Brief details of the ESG metrics and BRSR critical discourses are shown in Table 3 [41].

This reporting focuses on SDGs (Sustainable Development Goals), which the United Nations Sustainable Development drives in the Paris Agreement and ILO (International Labour Organisation) pact. However, in India, giant businesses such as Reliance Industries Limited, TATA and its associated forty groups companies, Wipro and Mahindra & Mahindra, and many more have been publishing their sustainability reports since the 2000s[16]. The ESG disclosures can support and help the corporates to excellently retort to stakeholder needs and increase their effectuation across countless ESG ratings. However, 1000 listed companies in India are presently reporting on the non-financial traits of their business in the form of sustainability reports or integrated reports. The journey of ESG reporting is shown in Table 4.

At present, the ESG sustainability framework follows the GRI (Global Reporting Initiative), SDGs (Sustainable Development Goals), TCFD (Task Force on Climate-Related Financial Discourses), SABA (Sustainability Accounting Standards Board), etc. These factors help develop a new outline of sustainable development. That's why the concept of ESG is swiftly increasing and becoming mainstream, practiced, and famous in the working field. Also, scholars and investors are moving to concentrate on non-financial factors such as social and environmental to identify the investment with progress prospects in their business and reduce the risk accompanying decisions.

## 5. Need for Sustainable Reporting

2015, the Paris Agreement was held, which was legally requisite by the international climate change convention. One hundred ninety-six parties accepted this at COP-21, held on December 12, 2015, and came into existence on November 4, 2016. The Paris Agreement aims to keep the environment safe from high temperatures due to global warming and keep the GW below 2C. Due to global warming, the earth's temperature has increased by 1.5 °C since 1980, but the maximum warming occurred since 1975, at a speed of approximately 0.15 °C to 0.20 °C decade[42]. This aim can be achieved by setting a long-term goal to reduce greenhouse gas emissions as rapidly as possible. It defines a multidimensional climate change process that binds all nations together for the first time in protecting the environment from global warming. To adopt the Paris Agreement, businesses must focus on novel

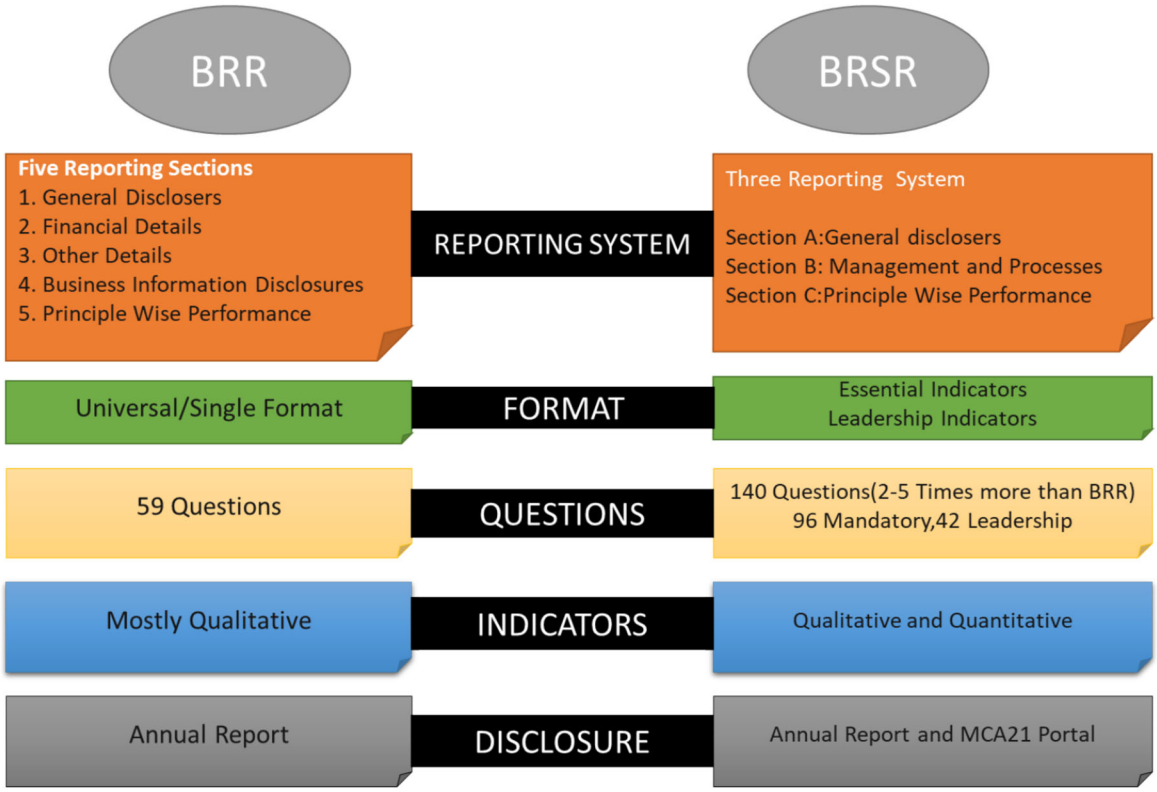



**Fig. 2** Difference between BRR and BRSR Reporting

**Table 2** The nine principles of BRSR[40]

| Principles | Explanation |
|---|---|
| Principle-1 | Businesses should conduct and govern themselves with integrity in a manner that is ethical, transparent, and accountable |
| Principle-2 | Businesses should provide goods and services in a manner that is sustainable & safe |
| Principle-3 | Businesses should respect and promote the well-being of all employees including those in their value chains |
| Principle-4 | Businesses should respect the interests of and be responsive to all their stakeholders |
| Principle-5 | Businesses should respect and promote human rights |
| Principle-6 | Businesses should respect and make efforts to protect and restore the environment |
| Principle-7 | Businesses when engaging in influencing public and regulatory policy, should do so in a manner that is responsible and transparent |
| Principle-8 | Businesses should promote inclusive growth and equitable development |
| Principle-9 | Businesses should engage with and provide value to their consumers in a responsible manner |

available science to bring social and economic transformation. This agreement is based on a 5-year timeline (2020) to achieve determined climate action by governments. This is why it has become necessary for every organization to perform sustainable reporting and climate change reporting to achieve the target of the Paris Agreement; through said reporting, organizations and governments both came to know if their business is attaining the Paris goal. Though momentous increases in climate change policy are required to convene the Paris Agreement goals,

**Table 3** A brief details of the ESG metrics and BRSR key discoursers [41]

| Matter of subjects | Key disclosures | Matter of subjects | Key disclosures |
|---|---|---|---|
| | Key KPI's-Quantitative & Qualitative | | Key KPI's-Quantitative & Qualitative |
| General | Principal Business Activities & Product Services | Social | Median Remuneration & Minimum Wages |
| | Trade & Industry, Plants, Offices Types of customers | | Return to Work & Retention Rates |
| | R & D & Capex investment in technologies to improve ESG impacts | | Union /Associations Memberships |
| | Overview of Materials Responsible Business Conduct Issues | | Occupational Safety & Health Matters, Including Risk Assessments & Reporting |
| | Specific ESG Commitments Performance | | Human Rights |
| Environment | Sustainable Input Sourcing | | Training on ESG, Human Rights & Skill Upgradation |
| | Environment Impact Assessment | | Safety Linked Product Results |
| | Extended Producer Responsibility | | Rehabilitations |
| | Life Cycle Assessments | | Product Labelling |
| | Operations in Ecological Sensitive Areas & Biodiversity Impacts | | Sourcing from Local/Small producers & Marginalised /Vulnerable Groups |
| | | | Beneficiaries of CSR Projects |
| | | | Data Privacy Issues |
| | | Governance | Anti Bribery/corruption disciplinary Sanctions involving directions & employees & related corrective actions |
| | | | Conflicts of Interest involving directors or KMPs,& related corrective actions |

in some years, new alternatives have come on the market to reducing carbon emissions. The 25% emission results from carbon emission, so carbon neutrality goals are critical and practical in markets. By 2030, zero-carbon strategies will be sustainable in markets that account for more than 70% of global emissions[43]. As a result, the World Organization must report sustainability and ESG reports to attain the above criteria.

## 6. ESG Estrategia

Before 18 years ago, the ESG (Environment, Social & Governance) was a Nobel idea and was solely an initiative published in a UN report. Now, the worldwide industry focuses on people's development and a greener environment rather than only getting good returns, revenue, and CAGR. That's why, in the present time, it is a must for every company or business to know the basics of ESG, understand what it means, and also realize their responsibilities towards the environment and society. ESG is hardly about profit-making but is concerned about almost everything else: how the company works, interacts with the market, treats its employees and workers, and impacts society. At this time, an effective ESG strategy provides excellent guidance to industries or companies. ESG strategy lies around the three principles of ESG, i.e., Environmental protection, acceptance of social issues, and ensuring the industry/company is running on the right path without affecting the environment and society [44]. These environmental factors include energy consumption, water use, carbon footprint, GHG (Green House Gases) emissions, biodiversity, and climate change. The social factors comprise how any organization treats its employees, customers, groups of people, and society. The governance part of ESG involves all the management practices and the formation and implementation of policies. Companies must look at several ESG trends, new ideas, and practices before making an ESG strategy. For instance, companies must focus more on reusing and recycling the product, its components, and materials rather than throwing them out while creating a responsible supply chain, reducing greenhouse gases, and adopting and implementing climate measures. Figure 3 shows the assumed ESG strategy, implementation, and reporting & communication flow. The following points should be considered while implementing the ESG strategy. The first important thing in making a successful ESG strategy is the integration of inputs from internal as well as external stakeholders. The future-formed ESG strategy aligns with the organization's primary values and operational performance and confirms the community and market requirements. In this step, access to top

**Table 4** The journey of BRSR reporting [39]

| |
|---|
| 2000 |
| Large companies (TATA group companies, reliance in india has published a sustainability report) |
| 2007 |
| RBI issued an advisory to commercial banks on their role in sustainability development, CSR & non-finacial reporting |
| 2008 |
| KLD Research & analytics, S & P launched ESG India index |
| 2009 |
| CSR Voluntary guidelines published by the ministry of corporate affairs |
| 2010 |
| KLD Research & analytics, S & P launched ESG India Index |
| 2011 |
| MCA published NVGs on environment, social, economic responsible of business companies bill 2021-CSR report to be included stating CSR policy |
| 2012 |
| Publishing of BRR made mandatory by SEBI for top 100 listed companies |
| 2013 |
| MSCI India ESG leaders index was launched |
| 2014 |
| Companies Act 2013 mandated companies of particular turnover and profit to spend 2% of the average net profit for the past 3 years on CSR |
| 2016 |
| Publishing of BRR Extended to Top 500 Companies SEBI is used in Green Bond Guidelines |
| 2018 |
| Insurance of Guidelines for ESG Reporting Voluntary Basis By Bombay Stock Exchange in 2018 |
| 2019 |
| BRR extended to Top 1000 Companies NVGs were Revised & Released as National Guidelines For Responsible Business Conduct |
| 2020 |
| MCA proposes a new format for BRSR |
| 2021 |
| BRSR made applicable to the top 100 listed companies by market capitalization voluntarily & mandatory from FY 22–23 |

management members, such as the board of directors, business heads, and business executives, is crucial because they all have deep knowledge of business and operational goals and their challenges. Their insights can help prioritize ESG issues directly impacting business performance and sustainability. With their fiduciary responsibility, board members can provide governance perspectives that ensure the ESG strategy aligns with long-term shareholder value and risk management. Business executives can offer operational insights, identifying feasible and impactful ESG initiatives that can be integrated into the company's operations.

In ESG issues, the materiality assessment is the strong pillar of ESG, and the integration of this can help identify and prioritize all the ESG issues that are most important to the business, its customers, and stakeholders. Some topics are less important to businesses or stakeholders during the materiality assessment, so these topics are separate from the overall ESG strategy. However, in the future, they will be considered in the starting phases.ESG performance selection is based on the baseline. The ESG baseline is the summary shot of the organization's performance across several ESG factors. It contains information on present policies, performance positions, practices, and essential statistics. This baseline helps to set a benchmark for future comparisons, which helps the organization assess the progress made over time and helps to make data-driven decisions. In the next step, the organization defines its measurable goals for ESG initiatives, develops an ESG roadmap, and develops a reporting framework for ESG reporting [45].

### 6.1. Role of Metrology in the Implementation of ESG & Linkage of ESG with GHG Protocol with Measurable Metrology Tools

Metrology is the science where measurements play an important role during the implementation of Environmental, Social & Governance practices through using metrological tools and a unique framework necessary for

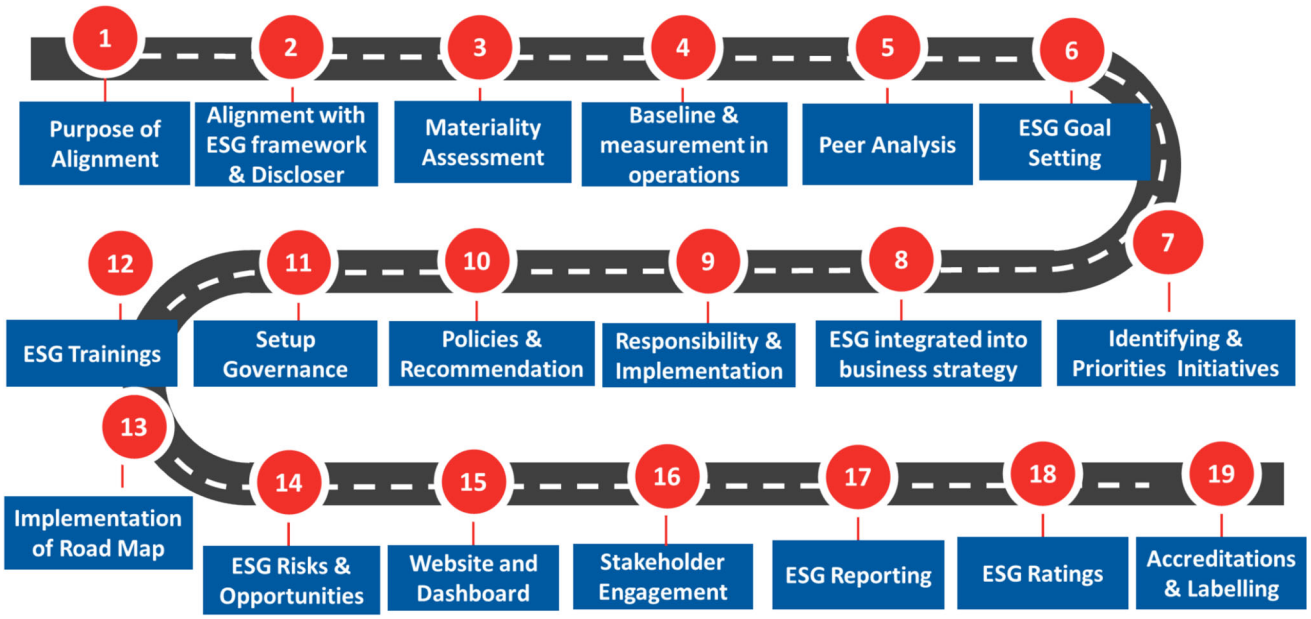



**Fig. 3** Assumed ESG strategy, Implementation, and Reporting & Communication Flow

accurately measuring, monitoring verifying, and validating ESG-based data. Thus businesses & industries are rigorously striving to attain ESG goals with more precise more reliable data capturing where the data progress, its transparency, and accountability can be measured and showcased during the reporting. Metrology helps in measuring the environmental impact such as carbon emission monitoring, resource efficiency, and air and water quality. For instance, in the carbon emission case, metrology instruments are used for measuring the carbon emission and can track organization GHG emissions by using several metrology-based tools per the reduction target. In the case of resource efficiency, the accurate measurement of water, energy, and raw material could support the organization in tracking their water, energy, and raw material usage to enhance their operational function to reduce their waste and improve their operational efficiency which could help the organization to attain their environment based sustainable goals. Another aspect of ESG is Social, where metrology helps in measuring the social metrics inside the social dimension related to social impact, labor condition, workplace health, and safety measurements such as noise measurement, air quality assessment, and ergonomics. In industries the working environment is very important if there is a lot of noise in the workstation, it adversely affects the employee well being and deteriorates their health. In this case, metrological instruments such as sound level meter, noise dosimeter, octave band analyzer, and many more could be employed to measure the level of noise on the ground floor to fulfill occupational safety standards, it certifies the safety of the environment, reduce risks of hearing and recovers the worker health. These metrological tools help measure whether the work environment meets health and safety standards, preventing respiratory issues or other long-term health problems[46].In the current scenario, governments & industries are highly emphatic on corporate sustainability, which has increased the adoption of the ESG framework & under these frameworks, the exact measurement of greenhouse gas emissions is also a must. The GHG protocol was initiated by WRI (World Resource Institute) and WBCSD (World Business Council for Sustainable Development) which provides the standardized guidelines for GHG accounting and its reporting after implementing the ESG strategy. Organizations can consider the GHG emission data in their ESG reporting, which would provide transparent information about their environmental impact. The GHG protocol and ESG framework are interrelated; they provide essential tools, instructions & guidelines for gauging and handling GHG emissions. Organizations have good options to integrate the GHG protocol into ESG strategies; by doing so, they can increase their E (Environment) performance, engage stakeholders, expand government practices & ensure regulatory compliance. This inclusive approach helps achieve sustainability goals and long-term accomplishment of business and resilience. After formulating the ESG strategy, ESG reporting comes into the picture, where GHG protocol is the most essential tool for measuring and managing GHG emissions. It is divided into Scope-1 (Direct Emission), Scope-2 (Indirect Emission), and Scope-3(Indirect Emissions throughout the value chain).Scope 1 is the direct emission from industry and manufacturing

facilities owned or controlled sources. Direct emissions occur from the direct sources owned by the company on their premises. At the same time, Scope 2 emissions are the indirect emissions from generating purchased electricity, heating, steam, and cooling. Scope 3 includes all the things that need to be covered in the Scope 1 & Scope 2 boundary for instance, supply chain, suppliers' relationship, stakeholders, and other industry-related activities. Figure 4 shows the GHG interiorization across the dairy firm.

### 6.1.1. *Role of Metrology in Measuring Scope-1, Scope-2, and Scope-3*

Metrology is the science where measurement is essential in quantifying the GHG emissions across Scope-1, Scope-2, and Scope-3 classes as distinct by GHG protocol.Scope-1 emissions come from owned sources or sources the organization controls and depend highly on metrology for accurate measurement. Continuous emission monitoring systems(CEMS), NDIR(Nondispersive infrared) & FTIR(-Fourier transform infrared spectroscopy) are examples of some techniques that require rigorous calibration and maintenance to get accurate results. The certified calibrated gages concerning established methods are vital to maintaining & ensuring the measurements' rigorous consistency and reliability. Metrology comprises both experimental & theoretical parts of measurement and finding the serval stages of the uncertainty of these features. It plays a vital role in quantifying Scope 1 emissions. Scope 1 is the direct emission from industry and manufacturing facilities owned or controlled sources. The following are the essential metrology tools and technologies used to measure GHG emissions accurately.

*6.1.1.1. Nondispersive Infrared Gas Sensing* IR analyzers are speedily employed to detect and measure several gases present in the environment and make them measurable.IR can give precise and real-time data on monitoring GHG gases, essential components of the ESG framework. This technique is based on the optical sensing principle, where the infrared radiations are intermingled with the target analyte and absorbed during processes. The absorption parameters are not the same for all the gases. It is different for every gas; hence, due to the different adsorption characteristics, every gas molecule is fingerprinted and differently identified. A simple nondispersive infrared gas sensing system comprises a detector, optical filters, IR emitters, and circuit elements for conducting signals[46]. Figure 5 is a schematic diagram of nondispersive infrared sensing. There are some basic steps for calibrating the nondispersive infrared gas sensing. Before calibration of nondispersive infrared gas sensing, the first basic step is to prepare for calibration. In this step, calibration gases are chosen with the known quantity of concentrations or regularly traceable to the fixed standard of national metrology institutes. For proper calibration of the instrument, the working environment should be under controlled pressure, temperature, and humidity because if

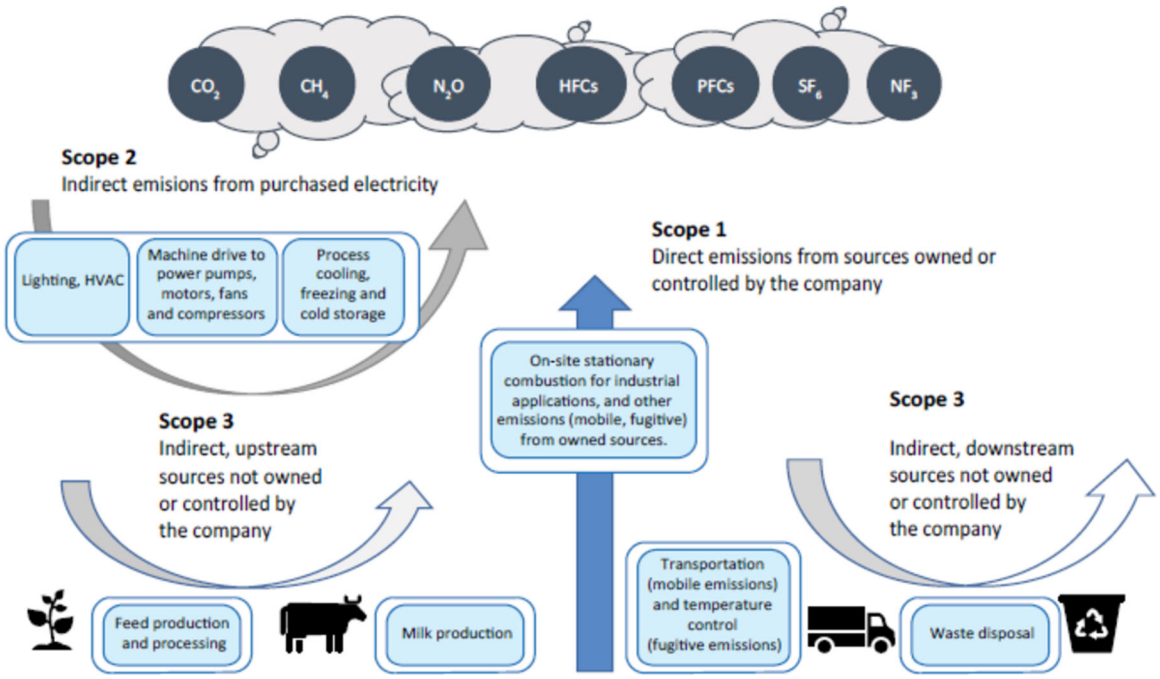



**Fig. 4** GHG Interiorization across Dairy Farms [47]

these parameters are not controlled, they can influence the performance of sensors. In the second step, baseline adjustment needs to be fixed where the sensors are exposed to the gas-free environment of either nitrogen or filtered air to fix the zero calibration or zero point. This step is very important for any sensor calibration from their baseline. The next calibration step is fixing the calibration span where the sensor is introduced in one or more than one target gas concentration. As the span calibration is fixed, obtained records responses are compared with the predicted response based on the known gas concentration, and the curve/scaling factor is adjusted to match the expected values. If the sensor does not give a linear curve, then mathematical corrections are applied to linearize the output. After the sensor performance calibration, the validation & verification is the next step, where sensor responses are maintained by adjusting the environment condition using the calibration log records for future reference and traceability. After validation & verification, regular checkups or recalibrations of sensors are performed to note the changes in sensors' characteristics over time. These sensor calibrations are integrated with advanced software to process the data and ensure the calibration software is updated according to the new calibration updates [48].

*6.1.1.2. Gas Chromatography* Gas Chromatography is a systematic technique employed to identify volatile & semi-volatile compounds, their separations, and their quantifications.GC also has tremendous capability to measure greenhouse gas emissions and provides precise and correct data on GHG gases; these gases are carbon dioxide ($CO_2$), methane ($CH_4$), and nitrous oxide ($N_2O$).In gas chromatography, the sample is introduced into the sample, where it starts to vaporize & taken by an inter gas via a chromatographic column. The chromatographic column is filled with the non-movable phase, which helps separate compounds that work on the interactions with non-movable phases and their B.P.(Bolling Points). The separated compounds are detected by instruments such as an electron capture detector for $NO_2$ and FID (Flame Ionization Detector) when these compounds leave the column. Before performing the test in gas chromatography, the basic step is the calibration. GC requires calibration to get accurate results. The following steps are required to be followed before performing the tests. The first step is the preparation of standards. In this step in calibrating the GC is the selection of a sample, which means identifying an air sample that would have gone through the under-experiment process. As the sample is decided, the standards must be prepared using certified reference materials or by preparing their standards. However, the important thing is that these prepared standards should contain known concentrations of the sample of interest. As the reference material is selected, the next step is preparing a series of standard solutions with changeable concentrations, which must cover the varying range of samples. After that, the second step is the setting of the GC experiment where before experimenting, it must be required to ensure the GC is contamination-free by cleaning the injector column and detectors and checking the flow of gases by properly setting the column parameters such as temperature program, port temperature (Injection), and detector setting. The third step is a blank sample running, after ensuring the contamination-free chamber, initiate the experiment by injecting a selected solvent while finalizing the standards on the blank sample. This step helps to ensure that contamination is carried over in the experiment setup. Before conducting the experiments, the baseline checkup is also a must to ensure that no unwanted peaks are present. In the fourth step the inject calibration standards**.** In this step, the lowest concentration and record each sample retention time, peak area, and height. Following this, continue injecting the remaining standards from the lowest to the highest concentration, ensuring the GC returns to baseline between each injection and

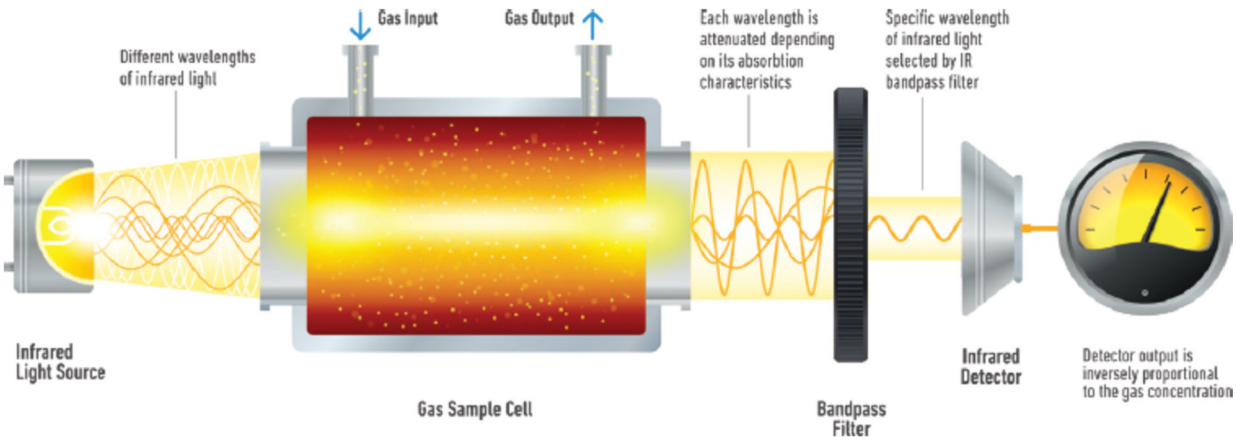



**Fig. 5** Schematic Diagram of Non-Dispersive Infrared Sensing [46]

maintaining consistency by using the same injection volume for each standard, which is crucial.

The fifth step is the generation of a calibration curve & data validation, where as the data is received through the experiment, the GC software starts to plot the graph of peak area vs the standard concentrations; this obtained graph is called the calibration graph. For the data validation, the obtained results are validated to check the consistency and accuracy of the calibration. The last step is the preparation of documents for calibrating results such as calibration curves, response factors, chromatograms, etc., which are saved for future reference at the last obtained results [49]. Figure 6 shows the gas chromatograph.

The signals generated by GC are in the form of peaks,& these peaks are different for different compounds. The GC is highly recommended for measuring GHG emissions because of its high sensitivity, high quantitative accuracy, and excellent selectivity, making it perfect for ambient air & more concentrated emissions. Future advancements in GC technology, including enhanced detectors, faster columns, and integration with real-time monitoring systems, promise to improve its effectiveness in GHG measurement and management [50].

The global economy is expanding; there is not only one challenge against climate change but also enormous challenges, which include GHG emissions, pollution, inflation hikes, supply chain disruptions, geopolitical conflicts, and many more. Lastly, all these challenges impact the country's economic growth and force governments and businesses to eliminate the risks and adapt to fluctuating circumstances. To solve all these issues, India has committed to net zero emissions by 2070, and to achieve the said target, India must invest a significant amount. In 2019–2020, green finance issued INR 309 thousand crores per year, whereas India requires only a quarter amount to become green. India is poised to emerge as a frontrunner in sustainable development with a firm dedication to minimizing carbon emissions and bolstering the use of non-fossil fuel alternatives in electricity production. This comes under Scope-2 emissions, which are concerned with indirect emissions such as the consumption of purchased electricity; here, metrology plays a vital role in accurately accounting for energy consumption. For measuring energy consumption, electricity meters, their calibration, and maintenance under the strict metrological standards give precise readings for calculating emissions. Furthermore, the standardized emission factors can be developed and checked using metrological practices and converting energy usage data into GHG emissions. The reliability and accuracy of the data are increased by solely verifying these factors, which enhances the credibility of Scope 2 emissions in ESG & BRSR reporting.

At the COP 26 summit, India presented five step plan to become greener called "Panchamrit."One of its plans is based on solar installation in every home and industry. The Indian Government issued INR 19500 crore to fix photovoltaic modules to push the country's shift towards renewable sources. The installation of solar panels reduces Scope 2 emissions due to the generation and purchase of energy. These solar panels are passed through many diagnosis systems, and many metrological instruments do this system check. These instruments are a thermal imager, digital multimeter, solar power meter, battery analyzer, and nanocontact infrared thermometer where the thermal imager detects the hotspot during electrical and breaker panels of PV modules. It is also employed to monitor leakage of cooling fluid levels, such as dielectric oils for transformers and the digital multimeter is the instrument used to detect the voltage switch. The next is the solar power meter or irradiance meter, which measures the overall energy efficiency of all solar systems. This solar power meter helps determine the amount of sunlight a particular area receives. It is used to find the optimum placement of solar panels. The battery analyzer is also a critical metrology instrument for checking the solar storage system batteries. The next is non-contact infrared thermometers, which measure accurate temperature over the surface of solar cells, panels, and areas that are also difficult to reach[51]. Figure 7 shows metrology instruments for measuring the solar panels.

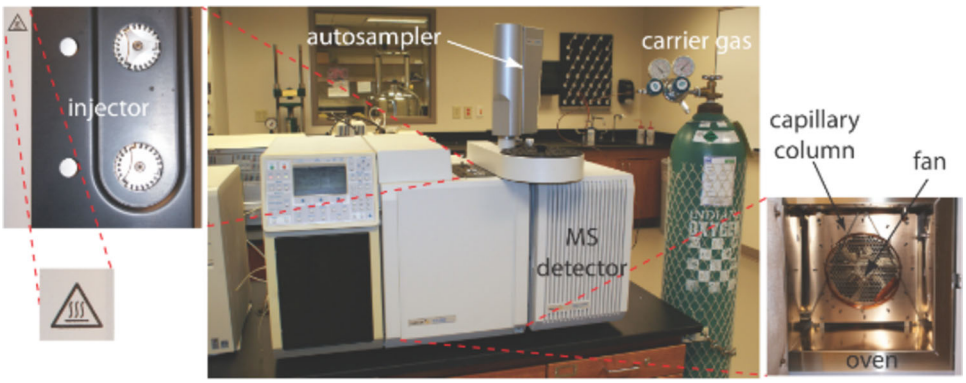



**Fig. 6** Image of Gas Chromatograph [50]




**Fig. 7** Metrology Instruments used for Diagnosis System **a** Thermal imager, **b** Digital multimeter, **c** Solar power meter, **d** Battery analyzer, **e** Nanocontact Infrared Thermometer

Whereas in Scope -3 emissions, all the indirect emissions in the value chain are part of it and possess lots of variability & complexity. The precise measurement of Scope 3 emissions depends on accurate data collection & estimated methodologies. In this case, metrology highly supports the development of tools and standards, which help collect all Scope -3 activity data and emissions from the different activities. Moreover, metrology also plays a vital role in standardizing the methods for conducting life cycle assessments to ensure the accuracy and quantification of emissions during the entire product life cycle. However, in the future, metrology can also include the technological advancements & development of international standards, the ability to increase the accuracy and affluence emission measurement. Digital technologies such as IoT(Internet of Things) & blockchain could deliver real-time monitoring & more transparent emissions reporting

### 6.2. ESG Reporting Disclosure Practice

In today's scenario, ESG become the most prominent topic among businesses, investors, and regulators. Its influence proliferates in business, asset managers, boards of directors of companies, deputation advisors, activists, and rating agencies. Small or large business organizations are affected by ESG issues. Businesses/Companies are continuously polishing their ESG credentials to reply to stakeholders and explain their ESG strategy, compensation, governance, and impact on society, the environment, profit, etc. These disclosures are also increasingly discussed in third-party analyses of ESG track records. The demand for ESG disclosures and standard frameworks in numbers is continuously increasing. For instance, according to the CDP (Carbon Disclosure Project) report published in 2020,515, investors whose assets value is $106 trillion and more than 147 significant purchasers whose assets value is $4 trillion have requested CDP to disclose their environmental data. Most ESG disclosures were filled voluntarily; now, the scene is being changed, and regulators and businesses are becoming proactive toward ESG disclosure. Third parties progressively deliberate these disclosures to analyze the ESG track records, etc. This section discusses the various ESG disclosures businesses undertake in several countries[13]. These are described as UN sustainable development goals, TFCD (Task Force on Climate-related Financial Disclosures), CDP (Carbon Disclosure Project), CDBS (Climate Disclosure Standards Board), GRI(Global Reporting Initiative), IIRC(International Integrated Reporting Council), SASB (Sustainability Accounting Standards Board. The various ESG disclosures are shown in Fig. 8.

### 6.3. UN Sustainable Development Goals

17 SDGs were adopted by the UN General Assembly in 2015 to complete the sustainable development agenda in 2030 to end poverty and hunger, save the environment, and ensure peace among the people. The 169 targets are linked with SDGs and 232 agreed indicators. To achieve the decided goals, many indicators will be added between 2020–2030. The SDGs are a good framework for developing ESG shaping, reporting, and ranking business strategy. Over time, SDGs would be amalgamated into many ESG assessment frameworks and form investment

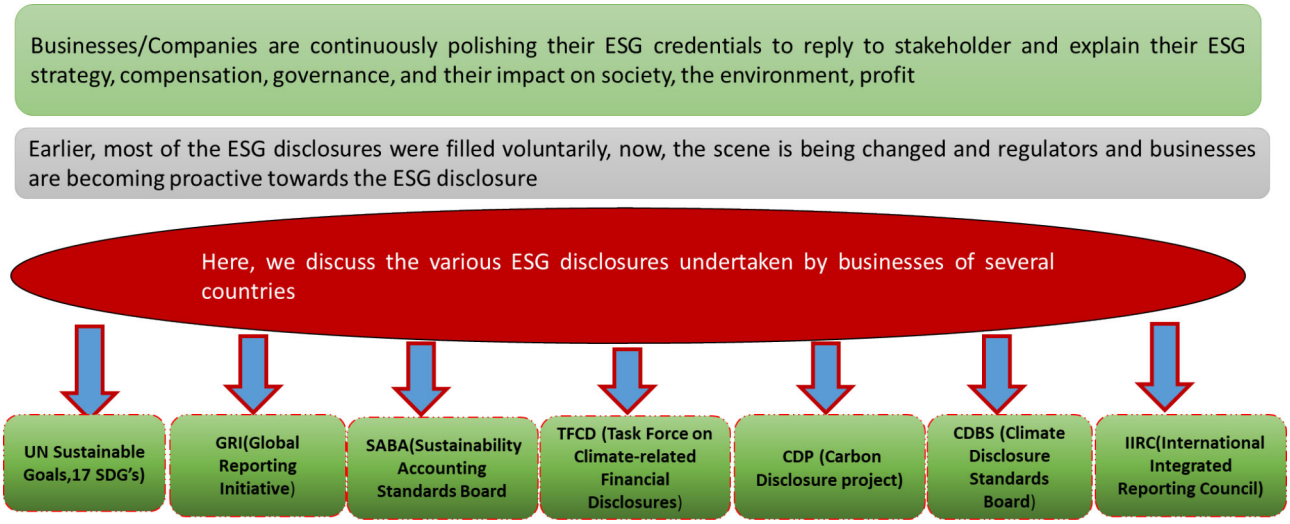



**Fig. 8** Various disclosures under ESG reporting

structures[52]. Nowadays, SDGs are expanding wings to figure out investment configurations. For instance, investors, customers, and issuers have started investing in SDGS bonds to increase the money for projects aligned with SDGs.In June 2020, in one sustainability program, the UN development deployed new principles(standards) for consultation, which targeted to guide them regarding the private equity fund managers and bond allocators for investment purposes in SDGs.

### 6.4. Task Force on Climate-Related Financial Disclosures(TFCD)

In 2015, at the G20 summit, the G20 Financial Stability Board (FSB) suggested cultivating the CRFR (climate-related financial risk) disclosure related to financial risk disclosures that investors, insurance companies, and businesses can use. G20 took this initiative to recognize the increasing authority of the financial system to ensure the entire financial system is strong enough to emergent climate-related risk. The concept of TCFD was launched at the end of the year to develop endorsements on CRFD (climate-related financial disclosures).In June 2017, the TCFD published its concluding endorsement[53]. According to FSB(Finacial Stability Board), climate risk-related financial information is not only a means for financial firms to manage their price correctly but can also be used as lending, insurance, or investment guaranteeing decisions based on their view of the transition scenario. About 2/3 of the FTSE(Finacial Time Stock Exchange) and 100 companies took TCFD referenced in their annual report which was published in 2019 has increased from 39% as compared to 2018[54]. The primary purpose of TCFD guidelines is to collect authenticated, consistent, helpful decision-making information that impacts material financial conditions, climate-related challenges, and opportunities, particularly those related to the global revolution of the low-carbon economy. They can be perceived in most financial filings for any company with public debt or equity in G20 jurisdictions.

### 6.5. Global Reporting Initiative (GRI)

GRI was established in 1997 with the association of UNEP, i.e., the United Nations Environment Programme. It is an international independent organization developed for standard reporting on global ESG impacts and developed with multi-stakeholder contributions. This organization was established to develop sustainability reporting standards to enhance the ingenuousness and accountability of finance, social responsibility, environment, and its outcome impacts on society; it also provides a robust sustainability reporting system to businesses that could be used globally. Since 2016, the GRI recommendations have also included references to the United Nations Sustainable Development Goals (SDGs)[4]. The GRI standards are divided into universal, sector, and topic-specific standards. The universal standards are classified as GRI 1: (Foundation), GRI:2(General Disclosures), and GRI 3: (Material Topics); GRI 3 gives a crystal clear view of the process of identifying several material topics starts by understanding the organization's context, identifying actual and potential impacts, assessing the significance of those impacts, prioritizing the most significant impacts for reporting and finally determining relevant material topics. Meanwhile, standards are classified as GRI 200 s, which cover economic topics; GRI 300 s, which cover environmental topics; and GRI 400 s, which cover social topics. In GRI

reporting, most organizations use all three sections but can choose the sector and topic standards that are most suitable for their organization. [55]. Figure 9 shows the outline of the GRI standard [56].

### 6.6. Sustainability Accounting Standards Board (SASB)

Climate change impacts the global economy's environment and social and governance issues. At present, investors have to understand the impacts of business on their business and its financial performance. But in reality, companies struggle to get standardized, akin to ESG reports showing their capital decisions. To solve this problem, sustainability disclosures such as SABA standards are coming to light to help businesses/companies identify, measure, manage, and report the subset of ESG-related topics of their business that would directly impact the long-term value creation of business. SABA is a non-profit organization founded in 2011 that tries to develop and maintain industry-particular standards for directing and guiding the disclosure related to financial materiality information of companies to investors and many other financial stakeholders[57]. It is an autonomous private sector standards-setting agency aiming to help companies worldwide identify, manage, and report on sustainability issues vital to investors, per SABA's view. The SABA is based on the grounded approach, which is industry-specific and financial materiality-specific, whose target is to provide a solid base for revealing (disclosing) sustainability information. The basic approach of SABA for standard setting is discussed below:

1. Global Applicability: The target of SABA standards is to provide authentic sustainability disclosure to the investor, which is reliable, pertinent, and comparable across global companies. The pertinent SABA metrics help investors invest in specific businesses globally and increase their global applicability.
2. Financial Materiality: This standard tries to recognize the ESG-related issues relevant to the financial performance of 77 industries. The selection of relevant sustainability information is very much required; neglecting or mispresenting it can change the entire risk profile of the business /company, influencing the complete capital allocation of the business.
3. Evidence-Based: In this phase, the SABA Board collects the financial information of a specific business from an external source to develop the financial impact of every sustainability-related issue that could affect the business. To ensure the significance of the ESG questions to the industry over time, it also studies the regulatory, environmental, and financial drivers for the given specific industry.

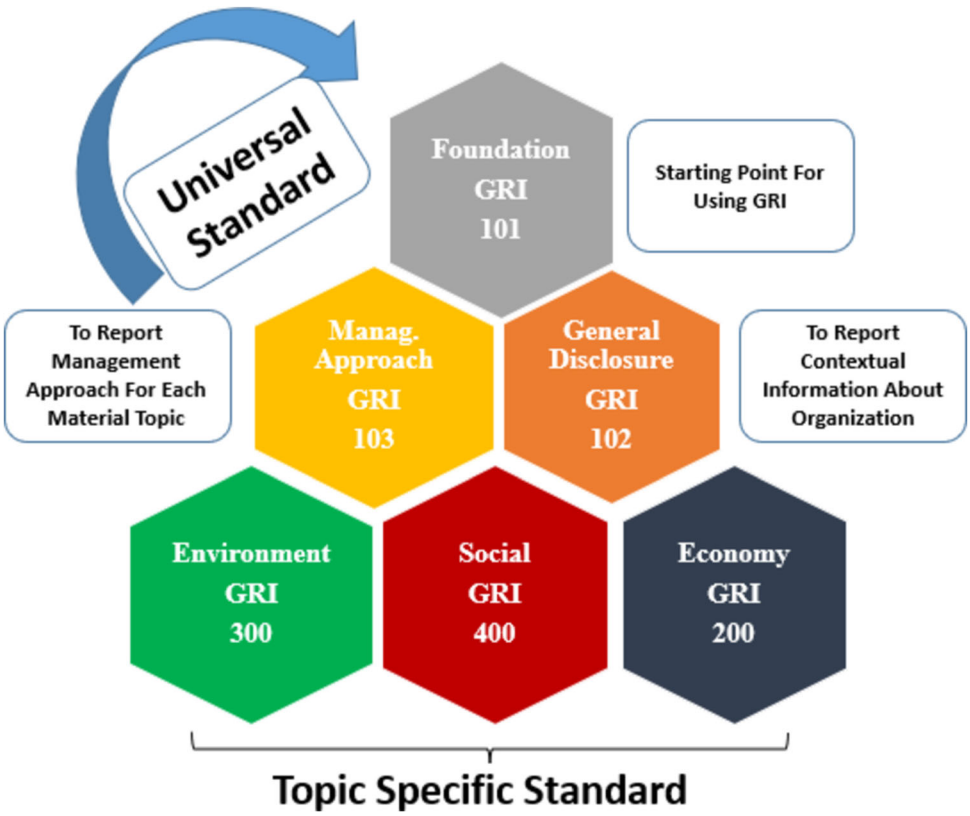



**Fig. 9** The outline of the GRI Standard

4. Industry Specificity: This standard believes in improving the specific industry sustainability disclosure, which concerns ESG issues.
5. Market Informed: The SABA standard boards ask for inputs from the appropriate companies, market participants, stakeholders, and other investors by keeping complete concern on sustainability issues that would be disclosed for the industry[57, 58].

## 7. Trends and Challenges of ESG Reporting

In recent years, there has been a change in the frequency and quality of ESG reporting by Indian corporates. Reporting started because of pressure created by investors, especially from foreign institutions, and it became a mandate over the last few years. Still, the number of challenges remains the same. The global standard of ESG reporting is growing at very speed; even the BRSR reporting framework takes some stimulus from Global Reporting Initiates (GRI) at a global level[59], but there is no clearness that Indian businesses/companies are conforming with GRI on an unvarying basis[60]. Secondly, the BRSR needs to improve its generality compared to international standards. BRSR disclosure has observed that disclosers only contain minimal, generic, and improper structures related to international working frameworks such as TCFD/EU[61]. More precisely, the BRSR framework looks like a generic and copy-and-paste arrangement of disclosures. It is a pretty over-distinct part of the area[61]. Thirdly, there needs to be interoperability regarding incongruent ESG reporting standards. As per Anderson's point of view[62], the concept of ESG is detrimental, more detrimental than the approach the ESG used, and its reporting is damaging. This is happening because of unhealthy thought-through metrics, damaging and no standard format and interoperability., The main root of this problem is the different methodologies, rules, regulations, and usage of standards, which creates confusion among the organization and develops many challenges in the ESG reporting concept. For example, the present world uses 600 distinguished ESG approaches, which causes a lack of harmonization. It isn't easy to get an exact figure of ESG reporting, which organization is doing what, and what will be their genuine ESG sources, ratings, performance, etc. The first goal of ESG reporting should be interoperability, and there should be a standard format for companies to collect ESG data and then report the data in a standard format that can be recognized as the ESG language used globally[63]. The fourth drawback in ESG reporting is the use of spreadsheets. As we know, the maximum business still depends on spreadsheets to chase the ESG data to develop the report and perform calculations. These spreadsheets pass from many data entry errors, break lines between spreadsheets, lack of data security, and failure to collaborate enormous amounts of data. Another problem with a spreadsheet is that one person uses it by applying formulas, functions, etc., that could be understood by that person only and perceptibly; it is not suitable for business continuity if the person (Excel Expert) decides to retire or leave from his job[64]. The fifth challenge in ESG reporting is the selection of KPIs and ESG metrics. The selection of ESG metrics and KPIs is always challenging to fix. Because of many ESG reporting frames such as GRI, SASB, etc. [65]. It needs to be clearer among the companies when the proper framework and KPIs for their entity are selected because the individual framework has its perspective on collecting data[64]. The next challenge in ESG reporting is the engagement of frontline workers for data collection. The effective ESG reporting system works on top-down reporting, involving all employees at every level of the organization daily. That's why many companies find it challenging to get the right person for data collection at every level of the organization.

## 8. Role of ESG Rating

ESG ratings are envisioned to give information to the market participants, and these market participants are analysts, corporate managers, and analysts regarding the relationship between corporations, investors, and non-investors interests of stakeholders. These markets examine the mass data and take out insight into the various elements such as environment, social and governance practice, performance, and risks. Investors 99% trust this information from market participants to create their investment decisions, whereas companies/businesses/corporations are using ESG ratings to increase third-party response on the developed quality of their sustainability initiatives. Lately, ESG rating provider companies have also been coming into inspection over trepidations of assessment reliability. These assessments are performed informally and on the concept of shareholder proposals related to investors' engagement with the organization on ESG matters. Many ESG rating provider companies boost input and engagement with their respective companies to enhance or sometimes correct the data. In the world, there are several ESG data provider organizations; some of the renowned ESG rating companies discussed in this paper are A) Bloomberg ESG Data Service Provider) Dow Jones Sustainability Index, C)MSCI ESG Research D)Sustainalytics Company ESG reports, E) Thomas Reuters ESG Research Data. This report will provide an overview and analysis of these providers. Bloomberg was established in 2009, and a new energy finance system was developed. This ESG data

service organization supplies information about renewable energy and the carbon market. Afterward, they launched the Bloomberg ESG data service center. This organization gathers ESG data for 10,000 published ESG discourses of listed companies. It scores the companies based on their published ESG disclosures after integrating the ESG data into Bloomberg Equities and Intelligence services. This organization has a mid-market cap coverage of 2$ billion to a large market cap of 10$ billion.It rates the companies out of the 100 rating scale. It also provides third-party assistance to companies through rating agencies such as Bloomberg, RobecoSam [66], Sustainalytics [67], ISS Quality Score [68], and CDP Climate Disclosure Sore[69]. Figure 10 shows the summary of ESG data and rating provider organizations. Bloomberg evaluates companies annually; it collects public ESG information from CSR (Corporate Social responsibility), sustainability reports, public reports, websites, annual reports, and direct collection through the company[70]. This data is checked and standardized by Bloomberg. Bloomberg considered 120 environmental, social, and governance pointers, including climate change, pollution, carbon emission, water disposal, etc. Bloomberg ESG rating will penalize companies for "missing data." Reputation-wise, in 2006, Bloomberg had over 12,200 ESG customers, providing ESG data to mainstream investors worldwide. Second is the Dow Jones Sustainability Index. It launched in 1991, and the Dow Jones Sustainability Index was one of the first global organizations to trail sustainability-focused public companies using RobecoSAMs ESG-based analysis. Dow Jones Sustainability and RobecoSAMs jointly worked to calculate and publish ESG indices for organizations. DJSI (Dow Jones Sustainability Index) has been divided into two parts: DJSI world, which tracks the world's leading companies, and DJSI country and DJSI Regions. DJSI has a market cap coverage of less than 2 billion dollars and a large cap of over 10 billion dollars. DJSI ranks the industry out of 100, contrary to other companies in their respective industry[71]. The DJSI World Index shows that 10 percent of the largest 2500 companies through 60 industries are in the S&P Global Board Market Index (BMI). The S&P Global Board Market Index (BMI) was float-attuned completely in 1989 and is merely a global index group that works on transparency with a modular structure. This more inclusive and highly transparent rating-centered index series works on a steady methodology across all the countries, including more than 14,000 stocks from 25 developed countries markets and 24 from developing countries markets[72]. The DJSI World index scores are developed by RobecoSAMs CSA (Corporate Sustainability Assessment), where 2500 publicly merchandised companies are requested to participate in the Corporate Sustainability Assessment. DJSI sent industry-specific questionnaires to participants that covered important environmental issues. Economic and social factors in questionnaires. These questionnaires contain 80 to 120 questions to judge their sustainability index. Basically.DJSI World considered 100 ESG issues. This questionnaire was sent to participants at the beginning of April, and their results were announced in September[71].The third is the ISS Quality Score, which gives ESG solutions to industries for developing and integrating responsible investment policies. It also provides several practices for their investment decisions for informing company activities and implements these developed policies by end-to-end voting[73].In 2015, the ISS quality index organization developed a strategic partnership with RepRisk and an advisory partnership with Ethix SRI, which helps the ISS explore and expand the ESG and SRI together (Socially Responsible Investing).The RepRisk is an ESG solution specializing in business conduct risk research and their quantitative answers (Environmental, Social, and Corporate Governance) provider company in Zurich and Switzerland[74].ISS also supports providing climate change data and analytics from its latest procurement of climate-neutral investments. It provides deep research on corporate governance in over 5600 public business companies globally. ISS quality index gives a rating from the 1st to 10th quintile; here, the 1st quintile means higher governance practices with low governance risks, whereas the 10th quintile shows higher governance risks.ISS examines more than 200 factors, which are subcategorized into four strong pillars; first is board structure, shareholder rights, risk oversight, audit, and remuneration, which depend on each region's governance standard, its ISS voting policies and its future impact on governance practices[73].The next rating organization is MSCI ESG research. It is the largest self-governing ESG rating provider. MSCI provides ESG ratings to over 6000 global companies and over 400,000 equity and constant income securities. The market cap of MSCI for the small-cap is less than 2 billion dollars, and for the large market cap is 10 billion dollars[75].MSCI works on AAA-CCC scale.The scale is again divided into sub-parts: CCC, B, BB, BBB, A, AA, and AAA. If a company lies on the scale of (CCC-B), it comes in "**Laggard Criteria**." This means the company is lagging in its high exposure and failing to manage the important risks of ESG. The second scale is subdivided into three points:- BB, BBB, and A. If a company lies on the scale of (CCC-B), it comes under the Average ranking. In this scale, the company has a mixed track record of handling the most vital ESG risks and prospects related to industry peers. The last scale is (AA-AAA).A company that attains this scale is a leading company that handles the most vital ESG risks and opportunities[76].However, in Methodology 37, MSCI considers ESG's main issues and divides them into three pillars(Environment, Social, and

Governance).The data is fetched from government databases, company disclosures, and academic and NGO databases, and the company's ranking is decided on a basis.

### 8.1. Challenges in ESG Ratings

Non-profit, non-government, and commercial organizations calculate the ESG rating to measure the performance, business models, corporate commitments, and structure alignment of sustainability goals of the organizations. The challenge is measuring the company's environmental, social, and governance impact, and reporting standards are still in the initial stages. The significant challenges in front of rating organizations are discussed below:

**Challenge 1: ESG Ratings are Costly**

The demand for sustainable investing and ESG rating is high among big and small asset owners greedily. According to the ERM research study,33 official investors have spent an average of 487,000 dollars annually on external ESG ratings and consultants. ESG rating consulting organizations generally charge high fees on sustainable investment product customers through high management fees, which can impact investors' returns rather than asset managers' BL (bottom lines). Though ESG ratings are supposed to be straightforward and constructive, excess data requirement in ESG ratings enhances the costs of raters. The companies must pay a handful of amounts to consultants or ESG raters because investors rely on ESG ratings.

**Challenge 2: ESG Rating Contradictions**

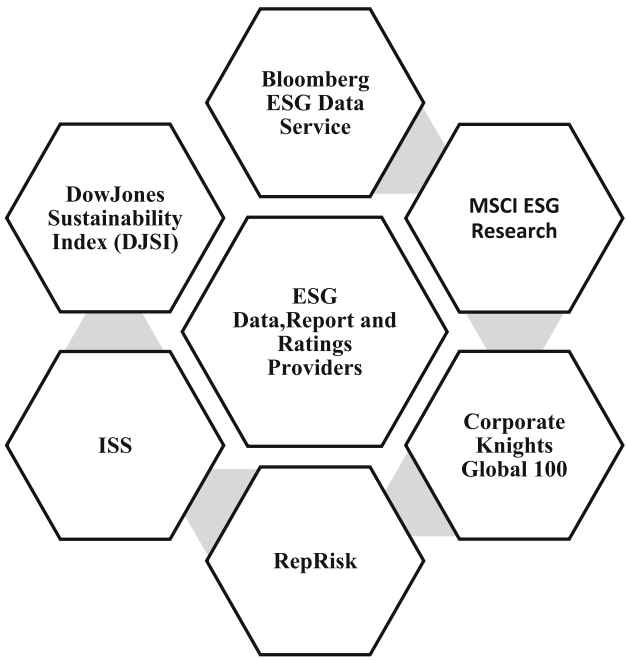



**Fig. 10** Summary of ESG data and rating providers' organization

The second challenge in variations in ESG rating is due to different methodology analyses of data done by ESG raters to give ESG ratings that usually are pretty secretive. There needs to be more motivation in the market for ESG rating firms (competing against each other to gain market share), so they attempt to develop more transposable results scores to make their position in the market. These ESG rating firms' ratings for the same firm are diverse. This variation in ESG rating needs to be clarified among investors to understand the sustainability and performance risks at the equity and portfolio levels. In one research, MIT researchers observed that only 61% of the average correlation in results was found in 6 big ESG rating agencies, whereas 91% similarity was seen in credit rating scores. According to one survey in 2022 of 1000 global investors, more consistency in ESG scoring is needed in implementing ESG investment strategies for investors[77].

**Challenge 3: ESG Ratings Are Not Evenly Distributed**

Investors and corporations need to be able to access the ESG ratings equally. The companies under the high market cap only list their ratings publically and are rated by raters. This irregular dispersal of rater's attention towards ESG ratings suffers the entire ESG ecosystem. Companies that just came into newly listed stocks for any area and market, in their case, institutional investors, make do with unclear coverage of ESG ratings. Due to the high cost of ESG consulting agencies, non-dependent investors and fewer asset managers have to work with limited admittance to ESG ratings. Private companies may not know about their employees' sustainability performance load, which is developing and must be improved [78].

**Challenge 4: ESG Ratings Are Not Predictive**

Many investors think that ESG ratings give accurate results to predict a company's ESG risk and opportunities, as credit ratings do for predicting credit risk. Maximum ESG raters have launched their rating software, products, and methodology for calculating risk ratings with the indirect or direct promises of predicting companies' profit possibility in the future when calculating ESG risk. Present-time sustainability expectations are increasing among companies and investors; companies can be limited to using old historical data for calculating ESG risks and opportunities. That's why the best ESG risk-based ratings can be achieved by using multiple new data from companies to construct robust predictions of future risk[49].

## 9. ESG Funds in India

ESG funds are mutual funds banks that invest in companies that strongly practice environmental, social, and governance. This funding supports business and companies'

practices or their work culture so that they can do better things for society[79].In the current scenario, India has ten ESG fund providers and individual funds with different characteristics.

In ESG funding, some funds are based on their capital. Some are passive funds, global stocks, and many more. These ESG funds in India are allocated to businesses as per their requirements. The ESG funds providers in India are as follows: SBI Magnum Equity ESG Fund, Aditya Birla Sun Life ESG Fund, Invesco India ESG Equity Funds, Kotak ESG Opportunities, ICICI Prudential ESG Funds, Mirae Asset ESG Sectors Leaders FoF, HSBC global Equity Climate Change Funds and Kotak ESG Opportunities[80]. Table 5 shows the ESG Funds in India comparison return(%) as of March 12, 2021 [81].

SBI Magnum Equity Fund was established in May 2018. It has also been repositioned as India's SBI Magnum Equity ESG Fund. This works in partnership with SBI and AMUNDI (France). AMUNDI is the world's prominent fund management fund organization. SBI Magnum Equity Fund has 41 equity holdings. Equity holding is the type of token share capital in a business or company where the shareholding allows the shareholder to vote, the right to profits distribution to shareholders, and to assets present for distribution on a winding up of that business or company. It holds 5 top companies: HDFC Bank, TCS (TATA Consultancy), Infosys, ICICI Banks, and Tata Motors.SBI Magnum Equity Fund invests at least 80% in stocks fitting ESG criteria and up to 20% in other equities and money markets [81]. According to Ruchit Mehta (SBI Magnum Equity Fund Manager), ESG investment faces challenges in India. The biggest challenge in ESG investing is the need for more data quality. Per his experience, the lack of data is not because companies have sustainability practices but because of the deficient (No Standardization) ESG disclosure[82]. Axis ESG Equity Fund was established in February 2020. It is a thematic ESG mutual funds scheme. Thematic mutual funds are those invested in socially responsible companies concerned with the environment, society, and governance. The sole aim of these funds is to attain sustainability growth. It has 51 equity holding and deals with five holding companies: Avenue Super Mart, Kotak Mahindra Bank, TCS, and HDFC Bank. Axis ESG Equity Fund invests 80% of the assortment in stocks with a high ESG score (High Rating) on ESG factors the allocations given to companies after a detailed review of the fundamental analysis of individual companies. According to Jinesh Gopan (Head of Equity), financial factors are also at the forefront of India. Companies capable of handling these non-financial factors through their corporate practices will be in a long-term race regarding investment prospects. They would ominously face less disruptive risks to their business model. Good ESG scores with good ESG practice also enhance their reputation in the market and lower risk probability[83]. Quantum India Equity ESG fund is also a thematic ESG mutual funds scheme. It was established in July 2019. It holds 45 equity and deals with five holding companies, including Avenue TCS, HDFC bank, Tata Motors, and Wipro. Quantum India, an Equity ESG fund, invests in companies after groundwork research companies. The Quantum India, Equity ESG funds, aim to develop a sustainable business based on E and G factors and make sense as an investment to develop prosperity over a long period. According to Chirag Mehta(Chief Investment Officer in Quantum India Equity), ESG phases help us know how the companies are ready to mitigate the various material risks and gain opportunities for capitalization [84].

### 9.1. ESG Investment

In the present scenario, ESG investments have quickly attracted customers and investors. ESG investments usually follow the triple bottom line approach, bringing financial returns together with social and environmental standards. The demand for ESG investments has increased due to raised environmental(climate) threats and increasing social and economic challenges. These factors pushed companies and businesses towards ESG investment. In addition, to attain these goals, businesses are working hard to map their goals with declarations.ESG investments are defined as "Socially responsible Investments," which means optimal ranking factors such as environmental, social, and economic factors and their outcomes. ESG investing focuses

**Table 5** ESG Funds in India comparison return(%) as of March 12, 2021 [81]

| Fund (Growth option) | 3 Months | 6 Months | 1 Year | 3 Year | 5 Year |
|---|---|---|---|---|---|
| SBI magnum equity (ESG) | 11.60 | 30.52 | 49.90 | 13.55 | 14.83 |
| Axis ESG equity | 8.53 | 30.02 | 48.04 | | |
| Quant ESG equity | 17.04 | | | | |
| Mirae asset ESG sec Ldrs FOF | 9.66 | | | | |
| Quantum india equity | 13.63 | 35.97 | 64.30 | | |
| ICICI Prudential ESG | 7.45 | | | | |

on sustainable investing, where the investment has to be made with the deliberation of its impact on the environment, humankind, and society [85]. One of the essential ingredients of sustainable investing is transparency in ESG reporting and its disclosure of information. This reporting has different legal faces in many countries; in some countries, it is voluntary, and in some, it is mandatory. This reporting and disclosure information is essential for customers and investors to comprehend the long-term business investment strategies of businesses, long-term business strategies, and their sustainability measures. This information is important to investors because it allows them to manage their investment risks, business strategy, and purpose. In fact, in the past, many ratings and scoring rating systems have indulged in offering companies a proper understanding of ESG performance in several rotations across the globe. This information has become essential for investors to manage investment risks and understand companies' strategies and purposes (PWC 2020). Many scoring and rating systems have evolved to offer an understanding of the ESG performance of companies in varying degrees and forms across countries[86].

### *9.1.1. ESG Investment: The Global Scenario*

Globally, ESG investments have gripped pace with speed among investors and consumers. The three main driving forces that give ESG investment rise are the Integration of ESG factors, which could positively affect investors towards risks and returns. The 2nd consideration is the transparency of money to customers in which companies invest. The 3rd consideration is the development and connection of the legal frame for ESG investment with legal necessities such as investments of many countries worldwide.One observation shows that about 24% of annual growth in ESG investment decision-making is seen worldwide. If we look geographically, the most significant contributors to ESG investment are the US and Europe[87], but in Asia, the acceleration of ESG investing is needed. Japan-like countries are at the top in Asian countries, and others are outspreading their ESG interments. As per the GSIA (Global Sustainable Investment Alliance) report, "DO GOOD" investment is proliferating, and investors are also considering it a profitable business opportunity. As per the information received from MSCI reporting, IT (Information Technology) is first in leading the ESG investment decision-making process, trailed by the healthcare system and financial sectors[75].The announcements of sustainable goals developments have also increased the drive of ESG investment in a further sustainable manner. According to one study, USD 7 trillion per year of investment is required to protect the climate as per the Paris Agreement. In the last four years, ESG investment has risen to 40.5 USD Trillion globally in 2020. Several new initiatives have been assigned and started in 2004 to enhance ESG investment in the US. The network for ESG investors has been established by UNRPIs (UN principle of responsible investment). The current information shows that the number of signatories has increased to 7000 across 135 countries [88], as shown in Fig. 11. The UN has also taken a horde of other initiatives, such as climate change, i.e., more than 100 + 167 leading corporate GHG emissions, to take pre-emptive action to attain net zero emission[87].

More than 500 investors with investment assets of USD 47 trillion aim to achieve net zero emission transitions under this scheme[89]. From above, it is easy to predict that the promotion of ESG investment is accelerating swiftly globally. The UN has taken good initiatives on ESG investment and declared SDGs, which imitate the requirements and primacies that the ESGs should underline.

### *9.1.2. ESG Investment: India Scenario*

The situation of ESG investment in India is in the budding stage of development. However, they have been embryonic very fast. Incorporating ESG components into the investment decision process is critical to reporting the continuing societal conundrums and growing environmental and climate threats[90]. The ESG elements are also synchronizing with 17 SDG goals. One study has shown that India entails a USD 0.6 trillion annual investment to meet the SDGs goal[91]. To attain growth in green energy and climate change investments, the country needs to focus on ESG investments. The country has made a global commitment to reducing carbon emissions by up to 30% by the end of 2050 and saving 40% of energy by using renewable sources

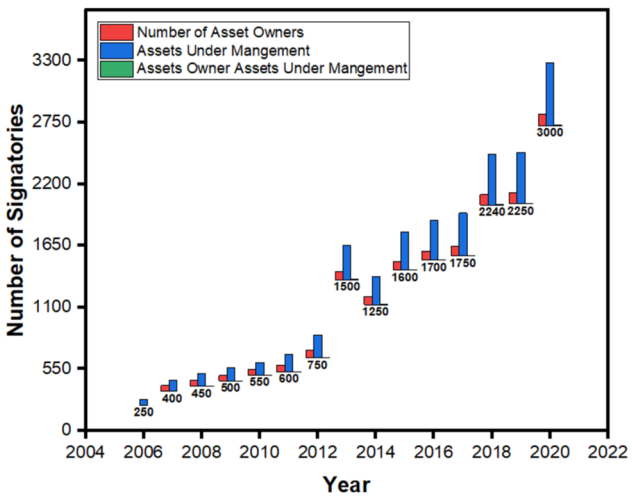



**Fig. 11** Information on the number of signatories and asset owners[87]

by 2030. This indicates that India focuses on the ESG investment process to get the results. Lately, GOI (Government of India) has understood the importance of ESG elements in business processes and has actively promoted and started to work on them with various improvements and regulations [90].The Indian corporate leaders have taken some noteworthy sustainability initiatives; one initiative is the participation of IC (Indian corporations) in the RE 100(Renewable Energy 100) global movements that support going green. This concept is fundamental to India from a business point of value proportions when electricity consumption is 40 percent more than the total produced in India[92].In India, Infosys, Dalmia Cement, Tata Motors Mahindra, and Mahindra have voluntarily allegiance to sourcing their energy generation from renewable energy and becoming part of RE100[65]. India's top cement manufacturing company "Dalmia" has announced the full-fledged use of renewable energy at the end of 2030. India and its corporate entities now see the opportunities in renewable, and they have started to work towards it. Likewise, Indian companies are doing exceptionally well under corporate PPAs (Power Purchase Agreements), and India ranks 2nd the most significant market in corporate PPAs (Power Purchase Agreements). Numerous Indian corporations such as Infosys, Vedanta, Havell's, Tata Motors, and Tech Mahindra also have a place in the Dow Jones Sustainability Index. Many cement manufacturing industries are producing energy from their waste, contributing towards reducing fuel and the carbon emissions produced through their manufacturing process[93, 94].

Nowadays, Indian companies are working tremendously on integrating ESG elements into their trade decision-making process. For example, Aditya Birla Sun Life, ICICI Prudential, and Axis Mutual funds have announced many ESG-integrated schemes[26], increasing the investment decision process and adopting ESG norms and regulations. According to the NSE 2020 study, 50 Indian companies have performed excellently in the area of policy disclosures as compared to ESG factors (Environment, Governance, social). This survey shows that environmental and social parameters are the least prioritized areas compared to the Governance area regarding policy disclosure. Figure 12 shows the Indian company's ESG scores for several ESG factors[95].

### *9.1.3. Challenges in ESG Investment*

ESG investing echoes ideal and futuristic but has decisive facts and numbers to gear it up because it has not been able to get complete power in the financial sector; same as another trait because it has several limitations and challenges demographic area wise to attain the target of sustainable future. According to SEBI (India regulatory Body), a minimum of 80 percent of ESG funds should go to the ESG subject, and the remaining funds can be used by fund managers where they can allocate investors' money. In Western countries, ESG fund investment is relatively easy. Still, in developing countries like India, India does not have sufficient clean technology such as electric vehicles and climate tech stocks to build a good portfolio like developed countries. Consequently, the listed companies (ESG and non-ESG) invest more money in adding names to investment portfolios. For instance, Reliance comes under 1000 listed companies for ESG discourser, its revenue generation stocks are highly dependent on fossil fuel, and it has made considerable investments in the clean energy business in the last few years. In 2021, Reliance declared an investment of 10 billion dollars, Rs.75 crores, for flourishing their green business over the next three years. This way, the reliance stock behaves like climate damage and tries to improve things. Until Reliance's green business does not come to the market, managers and investors must deal with the problem.

The genuineness of ESG data is another challenge in India. The data providers and funds are typically trusted in the company's annual reports and media reports for evaluating their ESG credentials, which are mostly only sometimes valid for assessment. Due to these problems, ESG investors need help to invest in ESG investing[24]. The next challenge in ESG investment is more costly than other average index funds because of excess research and broad workload. Good research is required for ESG funding that eats the maximum time investors earn because everyone wants to invest in ESG to remain in the sustainability race. The next challenge is the fees of financial firms. Financial firms charge higher fees for consulting and implementing ESG approaches. Another challenge is uncertainty in the selection process, which creates more insecurity for investors in ESG investment. The next challenge is the need for more quality data; many companies claim that they are ESG amenable but are not. In ESG investing, opaque data reporting cannot work; it should be transparent[96]. The next challenge is the need for a standard ESG measurement system. In the Indian market, many names around ESG investing, such as impact, socially responsible, responsible, sustainable, etc., confuse investors for investing. To attain good ESG investing growth in the future, there is a requirement to standardize ESG investing terms, data collection process, methodology, and ESG measurement standards during reporting. Another challenge in ESG investing in India is the old mindset of investors and fund managers. They thought that this type of investment would not give any ROI in the future and that money would be stuck; this type of investor thinking hinders the growth of ESG investing.

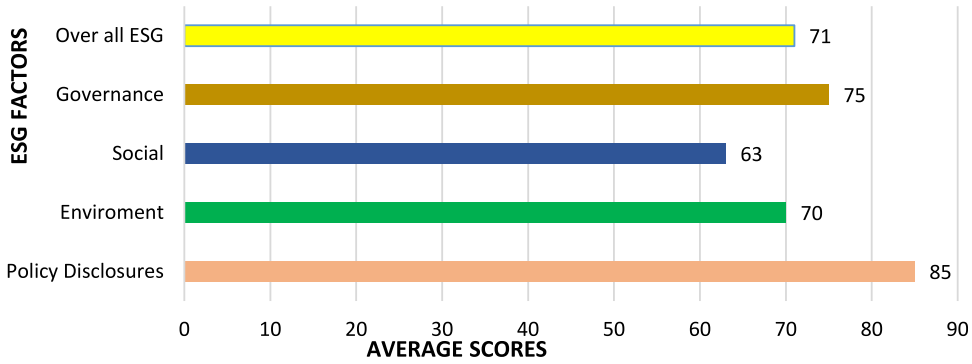



**Fig. 12** Indian companies' ESG scores for several ESG factors

Another challenge is the need to track the history of ESG funds records availability in the market. ESG funds started just 2 to 3 years back, putting investors in a dilemma to explore the investment this way. The last challenge, but not least, is the need for more investors' awareness. In the present day, when ESG investing is gaining popularity globally, investors still need to be made aware of every aspect of ESG, ESG funding, ESG investing, and many more. This challenge can be eliminated by developing proper awareness regarding the profit of ESG investing in more comprehensive terms[97]. Lastly, to fulfill the ESG investment goal, we all have to work on ESG investing strategy, and we need to invest in companies that are truly worried about social developments and ESG issues.

## 10. Limitation of Study

The present study is focused on a theoretical view of ESG, its framework, its need to report conditions in India with associated trends and challenges, its need in Indian organizations with its transformation landscape, its adoption preparedness and its challenges in India, the present status of reporting in India with different ESG disclosure practices trends, challenges and current status of India in ESG funds with investing followed by current challenges. In the future, the upcoming work will be more quantitative and more oriented towards ESG applications in distinguishing businesses with their reporting strategies by employing quantitative techniques to validate the results of the findings.

## 11. Future & Conclusion

The ESG is the primary framework for companies to follow in terms of sustainable development by integrating ESG into businesses, and by investment, the decision-making system becomes the international gateway for companies to practice sustainable development. The present study was aligned with the ESG concepts, reporting ratings, and many more in the context of India, as well as how it is essential to business and stakeholders with its pros and cons. In this study, the following conclusions can be exhibited:

- ESG practices become essential for attaining the long-term sustainability of any business/organization. Integrating ESG practices for the organization is crucial for fulfilling all its stakeholders' benefits. But unfortunately, in India, there are no fixed standard regulations concerning ESG. India has to take mandatory regulatory steps to treat the environment and stakeholders' priority.
- Recently, some steps toward the welfare of the environment and stakeholders have been taken by the Indian Government. On the other hand, in European countries, the government has taken many progressive steps to incorporate the ESG frameworks into the business, welfare of stakeholders, and many more.
- These steps are EU green bonds, ECO Labels, etc. Using these tools can help the Indian business but cannot be implemented directly; it needs to be modified by the Indian Government according to the Indian market scenario. For the smooth functioning of ESG in the Indian economy, India can use Green standards in existing ESG laws.
- In conclusion, metrology is fundamental to the accurate measurement and reporting of Scope 1, Scope 2, and Scope 3 emissions. By ensuring precise and reliable data, metrological practices enable organizations to manage their environmental impact effectively, comply with regulatory requirements, and achieve sustainability goals. As technology and standards continue to evolve, the role of metrology in emissions measurement will expand, supporting global efforts to mitigate climate change.

## References


[1] J. Kujala, S. Sachs, H. Leinonen, A. Heikkinen and D. Laude, Stakeholder engagement: past, present, and future. Bus. Soc., *61*(5) (2022) 1136–1196. https://doi.org/10.1177/00076503211066595.

[2] K. Miller, The triple bottom line: what it is & why it's important, Harvard Business School, 2020, Accessed on 23/1/2022

[3] T.T. Li, K. Wang, T. Sueyoshi and D.D. Wang, Esg: Research progress and future prospects. Sustainability (Switzerland), *13* (2021) 21. https://doi.org/10.3390/su132111663.

[4] S. Pareek and S. Pasumarti Subbarao, ESG reporting practices in India, UK and USA: an international comparison. Turkish J. Comput. Math. Edu., *12*(6) (2021) 2984–3006.

[5] D. Griggs et al., Policy: sustainable development goals for people and planet. Nature, *495*(7441) (2013) 305–307. https://doi.org/10.1038/495305a.

[6] T.C. Lindsey, Sustainable principles: common values for achieving sustainability. J. Clean Prod., *19*(5) (2011) 561–565. https://doi.org/10.1016/j.jclepro.2010.10.014.

[7] A. Boggia and C. Cortina, Measuring sustainable development using a multi-criteria model: A case study. J. Environ. Manage, *91*(11) (2010) 2301–2306. https://doi.org/10.1016/j.jenvman.2010.06.009.

[8] R.W. Kolb, Iso 26000. SAGE Encycl. Bus. Ethic. Soc. (2018). https://doi.org/10.4135/9781483381503.n648.

[9] T. Pri, T. Pri, T. Pri, U. Nations, and T. Pri, About the PRI What is the PRI? What are the six principles for responsible investment?, (2021), Accessed on 29/4/2022

[10] M.H. Schmiedeknecht, UN principles for responsible investments (PRI). Encycl. Sustain. Manag. (2022). https://doi.org/10.1007/978-3-030-02006-4_537-1.

[11] P. Matos. ESG and responsible institutional investing around the world: A critical review. (2020). https://doi.org/10.2139/ssrn.3668998.

[12] J. C. Stroehle, The social origins of ESG? An analysis of innovest and KLD 1 Robert G. Eccles," pp. 1–40,Accessed on 20/1/2022

[13] B. Berg et al., Harvard law school forum on corporate governance-introduction to ESG, *2020*(1) (2021) 1–6.

[14] Abhishek Dutta and Baitali Paul, Performance analysis of select ESG funds in India. Manag. J. Adv. Res., *3*(1) (2023) 1–7. https://doi.org/10.54741/mjar.3.1.1.

[15] A. Motwani. ESG Reporting: Environmental Dimension Disclosures by Large Energy Sector Companies in India (2023). https://doi.org/10.59324/ejtas.2023.1(2).11.

[16] R. Gupta and A. Motwani, Esg reporting in India: current scenario. Corp. Gov. Insight, *4*(2) (2022) 88–104. https://doi.org/10.58426/cgi.v4.i2.2022.88-104.

[17] R. Mandal, A. Murthy, and I. Stewardship, CSR in the post-pandemic era: The dual promise of ESG investment and investor stewardship, Accessed on 10/2/2023

[18] A. Afsharipour, Redefining corporate purpose: an international perspective. Seattle Univ. Law Rev., *40* (2017) 465–496.

[19] E. Gupta and A. Rangaswamy, The importance of 'Esg' and Its application in India. Lexforti Legal J., *II* (2021) IV.

[20] U. Varottil, The legal and regulatory impetus towards ESG in India: Developments and challenges. SSRN Electron. J. (2023). https://doi.org/10.2139/ssrn.4323313.

[21] U. C. Davis and U. C. Davis, University of California University of California no. December 2013, (2014) 6–10. Accessed on 23/9/2023

[22] I. MacNeil and I. marié Esser, From a financial to an entity model of ESG, 23(1). Springer International Publishing, (2022). https://doi.org/10.1007/s40804-021-00234-y

[23] B. Choudhury, Aligning corporate and community interests: from abominable to symbiotic. Brigh Young Univ. Law Rev., *2014*(2) (2014) 257–308.

[24] H. Vegan and D. Monte, ESG investments in India is complicated, (2022) Accessed on 10/2/2024

[25] M. Naniwadekar and U. Varottil, NUS Working Paper 2016 / 006 NUS Centre for Law & Business Working Paper 16 / 03 The Stakeholder Approach Towards Directors ' Duties Under Indian Company Law: A Comparative Analysis The Stakeholder Approach Towards Directors ' Duties Under Indian Company," no. January 2014, 2016, Accessed on 11/3/2023

[26] R. Sinha and M. Juneja, Environmental, social, governance financing goals in post-COVID-19 world. J. Resour. Energy Dev., *17*(2) (2022) 31–40. https://doi.org/10.3233/red-170203.

[27] R. Kaur, Business responsibility and sustainability reporting, (2023) 1–19, Accessed on 11/4/2023

[28] B. Lin, India transforms its ESG landscape to be future-ready, (2023) 1–11, Accessed on 14/3/2023.

[29] B. W. Tv, Four primary challenges hindering adoption of ESG in India, (2023) Accessed on 11/4/2023

[30] S. To, Z. Zhu and W. Zeng, Novel end-fly-cutting-servo system for deterministic generation of hierarchical micro-nanostructures. CIRP Annals, *64*(1) (2015) 133–136. https://doi.org/10.1016/j.cirp.2015.04.082.

[31] B. Menghnani, S. Babu, and F. Head, Business responsibility and sustainability reporting (BRSR)-an exploratory study, 1–28, Accessed on 4/4/2023

[32] European Comission, Sustainable finance: commission's action plan for a greener and cleaner economy, Press Release, (2018) Accessed on 24/4/2023

[33] K. Wöckener and M. Hauman, Building sustainability into the EU financial, (2019) Accessed on 4/4/2023

[34] Communication from The Commission to The European Parliament, The European Council, The Council, The European Central Bank, The European Economic and Social Committee and the Committee of the Regions Action Plan: Financing Sustainable Growth, 0203(504) (2017) 12008, Accessed on 14/5/2023

[35] T. Commission, T. Eu, and E. G. Deal, EU taxonomy for sustainable activities, (2023) Accessed on 4/4/2023

[36] EY, EU taxonomy regulation explained, (2020) Accessed on 12/6/2023

[37] Commission Européenne, Finance and the green deal, Europa, (2019) 1–6, Accessed on 4/4/2023

[38] Introduction to the climate bonds standards & certification scheme the climate bonds initiative: what we do, (2015) Accessed on 18/7/2023

[39] B. Home and S. Venkateshwaran, Reporting Partner in an increasingly volatile world - Relevance and evolution, (2023) Accessed on 23/9/2023

[40] The evolution of sustainability reporting in India. [Online]. Available: https://www.mca.gov.in/Ministry/pdf/BRR_110,Accessed on 4/4/2023

[41] ESG Reporting and its Framework in India, (2022) Accessed on 4/4/2024

[42] I. Revolution and S. Studies, World of change: global temperatures: feature articles world of change: global temperatures: feature articles, (2016) 2–4 Accessed on 4/4/2023

[43] U. N. C. Change, The paris agreement what is the Paris agreement?," Nrdc, (2021) 1–6, Accessed on 4/4/2023

[44] D. Byrne, "What is an ESG strategy?" Accessed: Apr. 05, 2024. [Online]. Available: https://www.thecorporategovernanceinstitute.com/insights/lexicon/what-is-an-esg-strategy/

[45] C. Stedman, Sustainability and ESG strategy and leadership, (2023) Accessed on 4/10/2023
[46] 10 types of noise measuring devices—noise measurement, Waqar Ali, 2023, Accessed: Nov. 25, 2024. [Online]. Available: https://www.hseblog.com/different-instruments-to-measure-noise-exposures,Accessed on 4/4/2023
[47] Guidance-Handbook-June-2019–120619.pdf, Accessed on 9/9/2023
[48] Calibration of NDIR CO2 Sensors. [Online]. Available: www.semeatech.com, Guidance_Handbook_2019, Accessed on 9/9/2023
[49] Stansel Rachel, Gas chromatography calibration: How to Calibrate GC, Guidance_Handbook_2019_FINAL", Accessed on 9/9/2023
[50] Gas chromatography—chemistry LibreTexts, Accessed: Jul. 11, 2024. [Online]. Available: https://chem.libretexts.org,
[51] List of essential test & measuring instruments for solar technical professionals. Accessed: Apr. 05, 2024. [Online]. Available: https://www.metravi.com/list-of-essential-test-measuring-instruments-for-solar-technical-professionals/
[52] UNDP, UNDP launches standards for bond issuers and private equity funds seeking SDG impact, (2020) Accessed on 9/9/2023
[53] TCFD, TCFD Recommendations, (2017) Accessed on 9/9/2023
[54] P. Study, Two-thirds of FTSE 100 mention TCFD in reporting, finds research, (2022) Accessed on 9/10/2023
[55] By Enviz, What is the global reporting initiative (GRI)?—IBM business operations blog, (2021) Accessed on 9/10/2023
[56] GRI, GRI 103: management approach 2016, Universal Standards 100 Series, (2018) 1–13, Accessed on 9/10/2023
[57] T. Sustainable and A. Standards, What is the SASB? What is the SASB? What are SASB Standards?, Accessed on 9/10/2023
[58] SASB, SASB Standards and other ESG frameworks, Sustainability and Accounting Standards Board, (2021) Accessed on 9/10/2023
[59] Sustainability in the spotlight: decoding India's ESG regulatory landscape—The Economic Times, Accessed on 16/10/2023
[60] G. R. I. S. Asia, S. Lanka, S. R. Trends, S. Asia, G. R. I. Standards, and G. R. I. S. Asia, Growing momentum for sustainability reporting in South Asia, (2021) 2021–2023, Accessed on 9/10/2023
[61] A. Goswami, A. Das, and A. Jain, An Introduction of ESG Disclosures in Indian Regulatory Space—Part 1 , India Corporate Law", (2021) 2019–2022, Accessed on 19/11/2023
[62] J. Anderson, S. Mortgage, and I. Trust, Anderson, ESG frameworks are profoundly damaging, (2021) 0–2, Accessed on 19/11/2023
[63] Green Business Bureau, ESG reporting: standards, frameworks, challenges and benefits, (2021) Accessed on 19/11/2023
[64] J. Niemoller, 5 ESG reporting challenges and how to address them, (2023) Accessed on 19/11/2023
[65] Niemoller and Jhon, ESG reporting frameworks: download now: climate risk management guide, (2021) Accessed on 19/11/2023
[66] A. Cookies, RobecoSAM celebrates 20 years in sustainability investing, (2015) Accessed on 19/1/2024
[67] H. Garz, C. Volk, and D. Morrow, The ESG risk ratings, Moving up the innovation, November 2018, Moving Up the Innovation Curve, Accessed on 19/1/2024
[68] B. Iss and E. S. G. Thought, A data-driven scoring and screening solution designed to measure and identify areas of environmental and social risk through company disclosure. Accessed on 19/1/2024
[69] RIO ESG, Making the CDP A-list: your guide to CDP Reporting, Rio, 2021, Rio Blog—Making the CDP A-List: your guide to CDP reporting, Accessed on 19/1/2024
[70] B. Freed, ESG reports and ratings: what they are, why they matter, 2022, ESG reports and ratings: what they are, why they—ProQuest, Accessed on 19/1/2023
[71] Dow Jones, Sustainability world index, Fact Sheet, (2016) 1–5, Dow Jones Sustainability World Index | S&P Dow Jones Indices, Accessed on 16/4/2024
[72] G. B. M. I. Shariah, S & P Global BMI, 2018,S&P Global BMI Shariah | S&P Dow Jones Indices ,Accessed on 16/4/2024
[73] B. Iss, ISS enhance investment decision making and engagement with ISS global non-financial data and analytics, Accessed on 16/4/2024
[74] H. Zurich, RepRisk, RepRisk | The need for independent providers of ESG data, (2020) Accessed on 16/4/2024
[75] T. Now and B. News, ESG investing part 3: MSCI exclusive—Dissecting the indices, Accessed on 18/5/2024
[76] G. Key, What is an MSCI ESG Rating? How does MSCI ESG Ratings work?, MSCI ESG Ratings Definition, Methodology, Example Accessed on 18/5/2024
[77] Capital Group, ESG Global Study 2022, (2022) 36 Accessed on 20/5/2023
[78] A. Brackley, E. Brock, and J. Nelson, Rating the raters yet again: six challenges for ESG ratings, Erm, 2022, Rating the raters yet again: six challenges for ESG ratings, Accessed on 22/5/2023
[79] E. S. G. Funds et al., What are ESG Funds? How are they different from other mutual funds? ESG Funds: an overview, (2022) Accessed on 24/6/2023.
[80] R. Seth, S. Gupta, and H. Gupta, ESG investing: a critical overview, Hans Shodh Sudha, 2(2), 2021 69–80, Accessed on 22/5/2023
[81] Morningstar, ESG funds available in India, Morningstar, p. 1, 2021,ESG funds available in India | Articles | Morningstar India, Accessed on 22/5/2023
[82] R. Mehta and F. Manager, Analysis is an extension of the security evaluation process. Why are ESG parameters critical for India?, ESGRisk.ai, Accessed on 29/6 /2023
[83] J. Gopani and A. J. Gopani, Jinesh Gopani, Fund Manager Axis Mutual Fund About Jinesh Gopani Jinesh Gopani is the Head of Equity at Axis AMC, (2016)
[84] N. Shetty, F. Manager, and V. Fund, Chirag mehta is the new CIO of quantum AMC, (2022) Accessed on 29/6 /2023
[85] ADECESG, What is ESG Investing? | ADEC ESG, ADEC Innovations, (2022) 2023, Accessed on 29/8 /2023
[86] PwC Report—Mind the Gap: the continued divide between investors and corporates on ESG, (2020) 2020 Accessed on 29/6 /2023
[87] N. Nemoto and P. J. Morgan, Environmental, social, and governance investments: opportunities and risks for Asia, Accessed on 29/6 /2023
[88] T. U. States and L. America, New and former signatories Growth in size and diversity of signatory base continues, (2022) Accessed on 30/6 /2023
[89] G. K. Sarangi, ADBI working paper series resurgence of ESG investment in India: towards a sustainable economy asian development bank institute, 1284, (2021),Accessed on 29/6 / 2023
[90] V.A.R. Barao, R.C. Coata, J.A. Shibli, M. Bertolini and J.G.S. Souza, Press information bureau government of India. Braz Dent. J., *33*(1) (2022) 1–12.
[91] D. Goals, Impact investment: catalyst for sustainable, 2021, Impact investment: catalyst for sustainable development goals, Accessed on 20/9 /2023
[92] G. K. Sarangi and F. Taghizadeh-Hesary, Rooftop solar development in India: measuring policies and mapping business models, (2021) 20, Accessed on 29/6 /2023
[93] M. Sector, Vedanta gets listed on The Dow Jones Sustainability World Index, (2022) Microsoft Word—Vedanta gets listed on

The Dow Jones Sustainability World Index.docx, Accessed on 29/9 /2022
[94] Dow Jones Sustainability World Index 2021: tech mahindra ranked top IT leader in the Dow Jones Sustainability World Index 2021—The Economic Times, Accessed on 30/9/2023
[95] Global prime news NSE releases a report titled—ESG Analysis on 50 Listed companies in India, (2023) Accessed on 29/6 /2023
[96] J. Major, What is ESG investing? | definition & fundamentals, (2021), Accessed on 29/6 /2023
[97] M. Funds and M. Funds, What is ESG investing?—List of ESG funds in india table of contents how did ESG investing evolve globally?, (2023) Accessed on 29/6 /2023

**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 Science and Engineering B 313 (2025) 117949

Contents lists available at ScienceDirect

Materials Science & Engineering B

journal homepage: www.elsevier.com/locate/mseb

# Strategically engineered defects in 2D-SnS for water splitting: Green energy solution for powering LEDs with hydroelectric cells

Hemant K. Arora [a], Jyoti Shah [c], R.K. Kotnala [c], A.S. Rao [d], Nitin K. Puri [a,b,*]

[a] Nanomaterials Research Laboratory (NRL), Department of Applied Physics, Delhi Technological University, Delhi 110042, India
[b] National Institute of Electronics and Information Technology (NIELIT) Patna, Amhara, Bihta, Bihar 8010116, India
[c] CSIR, National Physical Laboratory, Dr. K.S. Krishnan Road, New Delhi 110012, India
[d] Department of Applied Physics, Delhi Technological University, Delhi 110042, India

ARTICLE INFO

*Keywords:*
Solvothermal synthesis of SnS
Water splitting
Green energy production
Mesoporous
Powering LED light

ABSTRACT

The development of Hydroelectric cell (HEC) offers a promising solution to the current energy crises driven by the immense energy demand. Metal oxide and 2D nanosheets-based HECs are attracting significant attention for harvesting green electrical energy through water splitting, providing a sustainable power source for electronic devices. In this research, a SnS based HEC is fabricated to advance the environmentally friendly method of water splitting at room temperature, aimed at producing electricity. Mesoporous and highly crystalline 2D SnS nanosheets have been synthesized by rapid one-pot solvothermal method. Its morphology, structure and chemical composition were investigated employing various analytical techniques. X-ray diffraction (XRD), field-emission scanning electron microscopy (FESEM), high resolution transmission electron microscopy (HR-TEM) and Braumer emmer teller (BET) analysis have revealed the formation of orthorhombic, lamellar SnS nanosheets with high specific surface area. To fabricate the HEC, a zinc electrode was affixed to one side of the SnS pellet, and an inert silver electrode was painted on the opposite side. The fabricated SnS-based HEC, with an area of 4 $cm^2$, delivers a short circuit current of 15 mA, an open circuit voltage of 0.74 V, and an offload output power of 11.1 mW. Cyclic voltammetry revealed distinct cathodic and anodic peaks, corresponding to the redox reactions at the zinc and silver electrodes. The Nyquist curve of SnS-based HEC in its wet condition supports the ionic diffusion of dissociated $H_3O^+$ and $OH^-$ ions, demonstrating efficient electrochemistry.

## 1. Introduction

The development of green energy system has become essential for solving the issues posed by electricity demand and deterioration of the environment [1]. As the depletion of fossil fuels looms and the global population burgeons, there is a pressing need to meet rising energy demands while mitigating environmental pollution [2,3]. Also, the global warming and climate change have been compounded by rising greenhouse gas emission, which are driven by the need to meet the energy demands of a growing technology like gadgets, sensors, and etc. [4]. The scientific community has been making continuous efforts to obtain green energy through photocatalytic water splitting from available renewable energy sources, thus contributing to meet the world's energy needs [5,6]. Recently, hydroelectric cells (HEC) have garnered significant attention owing to their cost-effectiveness, environmentally friendly nature, and production of nontoxic by-products [7]. HEC is also an incredibly efficient method for generating hydrogen gas as a byproduct, which serves as a green fuel and can be utilized in fuel cells [8]. In this way, HEC because of its green energy source of electricity production by water splitting can be considered as an excellent substitute for the other energy sources such as solar system and hydropower [9,10]. Extensive research efforts have been directed towards exploring various ferrites, metal oxides, and their nanocomposites, with the aim of facilitating green energy generation through hydroelectric cells, eliminating the need for acids, electrolytes, light, or specific temperature conditions [10,11]. Key factors like porosity, lattice strain, and defects (including oxygen vacancies) are crucial in facilitating water dissociation on the material's surface. Electric current is generated through a redox reaction occurring at zinc and silver electrodes, which are affixed to opposite faces of the pellet fabricated as HEC [12,13]. Various materials, including $Mg_{0.8}Li_{0.2}Fe_2O_4$, $Al_2O_3$, $TiO_2$, and MgO, have been explored for hydroelectric cell (HEC) applications. In 2016, it was

* Corresponding author.
*E-mail address:* nitinkumarpuri@dtu.ac.in (N.K. Puri).

https://doi.org/10.1016/j.mseb.2024.117949
Received 4 August 2024; Received in revised form 12 December 2024; Accepted 16 December 2024
Available online 4 January 2025
0921-5107/© 2024 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

reported that $Mg_{0.8}Li_{0.2}Fe_2O_4$-based HECs produced a short-circuit current of 8 mA, marking a significant milestone [7]. Additionally, several other metal oxides have been investigated, with $SnO_2$ yielding a short-circuit current of 22.2 mA, $Al_2O_3$ producing 6.5 mA, $TiO_2$ generating 2.3 mA, and MgO yielding 1.5 mA [14]. Gupta et al. elaborated on the impact of doping on ZnO-based HEC output, demonstrating an enhancement of short-circuit current from 9 mA in pure ZnO to 25 mA for K-doped ZnO [15]. Furthermore, a Mg-doped $Al_2O_3$-based HEC exhibited a short-circuit current of approximately 16 mA [12].

Extensive exploration of various oxides, along with dopants and composites, has motivated further investigation into new materials with enhanced properties. The Metal oxide semiconductors (MOS) prepared for hydroelectric cell have certain limitation such that the defects can be induced at high temperature only [16,17]. Therefore, development of new material for HEC is required that can be synthesized at lower temperature and has more active sites for water splitting [18]. In recent years, 2D nanomaterials have garnered significant research interest due to their potential as foundational building blocks for the next generation of electronic devices [19]. These 2D semiconductor nanomaterials exhibit appealing electrical properties, including higher intrinsic mobilities compared to those found in organic electronics [20]. These materials exhibit enhanced electronic mobility, a high surface-to-volume ratio, and rapid response to surrounding environmental changes [21]. 2D nanomaterials, characterized by their layered structure, naturally form 2D nanomaterials due to stronger in-plane bonding relative to their out-of-plane van-der Waals interactions between layers [22,23].

Recently, transition metal chalcogenides (TMCs) have drawn significant attention for their promising applications in hydroelectric cells. Notably, layered SnS has garnered significant attention for its unique properties, finding applications in gas sensors, supercapacitors, photodetectors, resistive switching, hydrogen evolution reaction photocatalysis, solar cells, and more [24–28]. Significant efforts have been directed towards engineering layered SnS structures to increase the number of exposed sites [29]. SnS is a material that naturally exhibits p-type properties due to excess of sulphur (S) atoms. However, some authors have reported that the presence of sulphur (S) vacancies can impart n-type properties to SnS [30,31]. An alternative explanation for the n-type behavior observed in SnS is intrinsic doping, particularly through a disproportionally Sn-rich stoichiometry. Non-stoichiometry has been acknowledged as an effective strategy for altering majority carriers from holes to electrons in chalcogenides. This is achieved through the formation of anion vacancies and an excess of cations. Specially, an excess of Sn or a deficiency of S has been shown previously to induce n-type behavior in SnS nanostructures [32,33].

In the present investigation, we have synthesized 2D SnS nanosheets by one step solvothermal method for the fabrication of HEC. According to available research, no work has been carried out on SnS nanosheets for the green electricity generation by hydroelectric cell (HEC). The analysis of structure, morphology were done using XRD and FESEM. To explain the presence of sulphur vacancies and defects in the lattice of SnS nanosheets, the spectroscopic techniques such as Raman and PL are used. SnS based HEC demonstrates the remarkable output voltage and short circuit current values of 0.74 V and 15 mA, respectively. The as fabricated HEC produces the maximum power of 11.1 mW. The findings of this investigation are encouraging, indicating that power generation via hydroelectric cell has developed as a viable alternative to other renewable energy sources. Moreover, the utilization of SnS based HEC for water splitting opens avenues for meeting future hydrogen needs.

## 2. Material structural characterization

The X-ray diffraction (XRD) pattern was collected using a Bruker D-8 advanced instrument with CuKα radiation (λ = 0.154 nm) over a 2θ range of 20–80 degrees to determine the phase and crystallite size of the as-prepared SnS nanosheets. Raman spectra were obtained using a Raman spectrometer (α-Witek 300 RA) with a laser excitation wavelength of 532 nm. Photoluminescence (PL) spectroscopy was conducted to explore the defects of the samples at room temperature. PL spectra were acquired using a photoluminescence spectrometer (Perkin Elmer, LS 55) with a Xenon lamp as the excitation source, spanning the range of 200–800 nm with an excitation wavelength of 234 nm. Fourier-transform infrared (FTIR) spectroscopy was performed using a Perkin Elmer Spectrum II instrument across the wavenumber range of 400–4000 cm-1 to acquire FTIR spectra. The surface morphology of the as-prepared samples was examined using a ZEISS Gemini field emission scanning electron microscope (FE-SEM), and energy-dispersive X-ray spectroscopy (EDX) was performed to confirm the elemental composition. High-resolution transmission electron microscopy (HRTEM) was done using FEI Technai 200. XPS spectra were recorded on a Physical Electronics make PHI 5000 Versa Probe III spectrometer. Additionally, the BET (Brunauer-Emmett-Teller) and BJH (Barrett-Joyner-Halenda) technique was employed to analyze specific surface area, and pore volume, pore size. The voltage and current of the developed hydroelectric cells (HECs) of SnS were measured using a Keithley-2400 source meter. The Cyclic Voltammetry was done using Metrohm Autolab Potentiostat. Electrochemical impedance spectroscopy (EIS) was conducted to obtain Nyquist plots of wet cells, utilizing a Wayne-kerr 6500B instrument within a frequency range of 20 Hz-120 MHz.

## 3. Experimental section

### 3.1. Materials required

Stannous chloride dihydrate ($SnCl_2{\cdot}2H_2O$, 99 %), Citric acid ($C_6H_8O_7$, 99 %), and Ethylene Glycol ($C_6H_6O_2$, 99 %). All the chemicals were purchased from Fisher chemicals. The Thiourea ($CH_4N_2S$, 99 %) used in this study was purchased from Sigma Aldrich. All the chemicals belong to AR grade with highest purity of (99 %).

### 3.2. Synthesis of SnS nanosheets

The synthesis of 2D SnS nanosheets was conducted via the solvothermal method. The procedure commenced by dissolving 0.1 M of $SnCl_2{\cdot}2H_2O$ in 100 ml of ethylene glycol, followed by continuous stirring for 30 min. Subsequently, an appropriate quantity of citric acid ($C_6H_8O_7$) was introduced into the solution. Next, 3.33 g of thiourea ($CH_4N_2S$) was added, and the mixture was stirred continuously for 1 h. The resultant solution was then transferred to a 200 ml stainless steel autoclave, sealed, and maintained at 180 °C for 24 h before being allowed to cool to room temperature. The resultant black sample was washed several times with deionized (DI) water and anhydrous ethanol through centrifugation, followed by drying at 80 °C for 24 h. Fig. 1. shows the synthesized prepared SnS nanosheets using the solvothermal method.

### 3.3. Fabrication of hydroelectric cell

The as synthesized SnS nanosheets were processed into a square pellet (area = 4 $cm^2$), as illustrated in Fig. 1, using a hydraulic press by applying pressure of 5.5 tons. The pellet underwent annealing at 300 °C for 2hrs to induce suphur defects. A zinc sheet was attached to one of its faces of the pellet where as the silver paste was applied over the other face in a comb pattern. Zinc acts as the anode whereas the silver paste acts as the inert cathode. Low-resistance wires were soldered to both electrodes to fabricate a SnS based Hydroelectric cell. Fig. 1 shows the fabricated hydroelectric cell based on SnS nanosheets used for electricity generation.

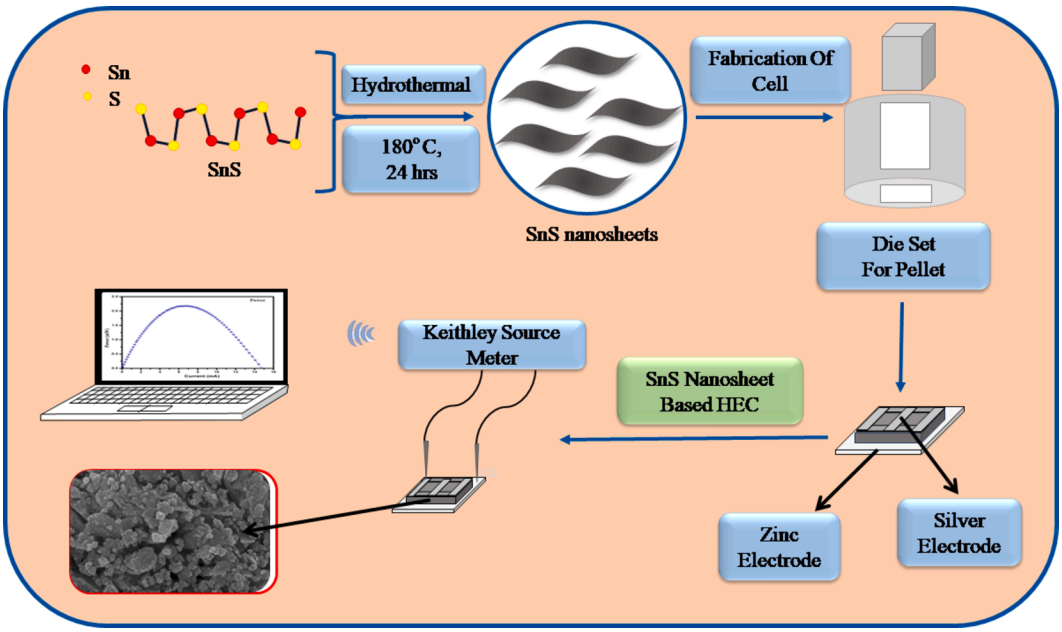



**Fig. 1.** Schematic diagram illustrating the preparation of SnS Nanosheets and fabrication of HEC.

## 4. Results and Discussion

### 4.1. X-ray diffraction (XRD)

Fig. 2 (a). displays the XRD analysis of the as-prepared SnS nanosheets, providing insight into the crystal structure of the synthesized sample through XRD technique. The XRD spectra exhibits distinct peaks corresponding to various crystalline planes of SnS including (101), (201), (210), (011), (111), (400), (311), (102), (411), (020), (112), (511), (221), (601), (420), (512) and (711) at angles of 21.9°, 26.1°, 27.4°, 30.5°, 31.6°, 32.1°, 39.1°, 42.7°, 44.8°, 45.4°, 48.7°, 51.2°, 52.9°, 54.2°, 56.6°, 64.29°, 66.48°, respectively. The diffraction patterns of SnS matches with the standard JCPDS card of orthorhombic SnS (JCPDS 39–0354) [34]. Remarkably, the peaks corresponding to the SnS phase exhibit exceptional sharp and strong intensity, suggesting a high level of crystallinity. Significantly, there were no impurity peaks, such as those attributed to $SnS_2$ and $SnO_2$, were detected in the XRD patterns. This suggests that the solvothermal process yielded highly phase-pure and crystalline SnS nanosheets. The interplanar spacing of the four highest intensity peaks was calculated using Bragg's relation: $n\lambda = 2d\ Sin\theta$. Where 'd' denotes the d-spacing values calculated for various high intensity planes of SnS nanosheets. The d-spacing values in Table 1 from the XRD analysis of 2D SnS nanosheets provide key insights into the crystalline structure and orientation. These values indicate the distance between atomic planes, confirming the layered structure characteristic of 2D materials and highlighting the nanosheets preferential orientation, which is essential for applications such as water splitting or optoelectronics where orientation impacts performance.

### 4.2. Raman spectroscopy

Raman spectroscopy is a robust characterization tool for detecting diverse material phases based on discrete vibrational modes that indicate the nature of the chemical bond order of the atomic arrangement. SnS possesses 21 optical active modes, comprising 12 are Raman active modes, 7 infrared active modes, and 2 inactive modes [8]. As per the literature, the reported Raman active modes for SnS include $A_g$, $B_{2g}$, $B_{3g}$, while $B_{2u}$ is identified as the infrared active mode for the SnS nanostructure [35]. Fig. 2(b). shows the Raman spectrum of the as prepared sample of SnS for HEC. The spectra exhibit two distinct peaks located at 99, 224. The peaks at 99 $cm^{-1}$ and 224 $cm^{-1}$ corresponded to the characteristic Ag mode of SnS originating from the interatomic vibration between the metal (Sn) and chalcogen S atoms [36–38]. Based on XRD and Raman spectra, It indicates that the as-prepared SnS nanosheets formed orthorhombic phase.

### 4.3. Defects study in SnS nanosheets

PL spectra are a very efficient, non destructive contactless approach for characterizing and evaluating the quality of surfaces and interfaces, as well as to probing defect levels inside material. The Fig. 2(c). shows the PL spectra of SnS recorded at an excitation wavelength of 234 nm in the range of 400–600 nm. It mainly consist of three peaks at 437 (UV emission), 449, 466 (blue emission) nm [39]. The broad emission spectra of the SnS may arise from various defects such as high density of Sn, Sulphur vacancies introduced during the process. Besides sulphur and Tin vacancies, various kinds of defects such as interstitial, stacking faults have also been induced in the nanosheets [18]. The emission peak observed at 437 nm corresponds to Sn interstitial ($I_{Sn}$) and dangling bonds present at the surface of the SnS pellet [40]. The emission at 466 nm is ascribed to indirect band to band transition and the peak at 449 nm corresponds to Sulphur vacancy ($V_S$) which plays crucial role for water splitting by HEC at room temperature [39].

### 4.4. Fourier transform Infra red (FTIR)

Fig. 2(d). shows the Fourier Transform Infra Red (FTIR) spectrum of the synthesized SnS sample. The sample shows a peak forming at approximately 536 $cm^{-1}$ and 873 $cm^{-1}$ which can be ascribed to SnS bonding [41,42]. There is also a band at 1047 $cm^{-1}$ that is due to $SO_3$ stretching [43]. The band occurring at 3480 $cm^{-1}$ and 1400 $cm^{-1}$ is typically attributed to the bending vibrations of water molecules (OH) [44]. The weak signal may instead result from surface-adsorbed moisture during sample preparation. The peak near 1750 $cm^{-1}$ could correspond to the C = O stretching vibration, potentially pointing to the presence of carbonyl groups or other oxygen-containing functional groups on the surface. The characteristic peaks of SnS in the FTIR study and XRD result ($2\theta = 31.6°$), confirms the existence of SnS nanostrucure. Fig. 3(a). Shows the structure of the 2D SnS nanostructure and Fig. 3(b). shows the induced sulphur vacancies in the SnS nanostructure.

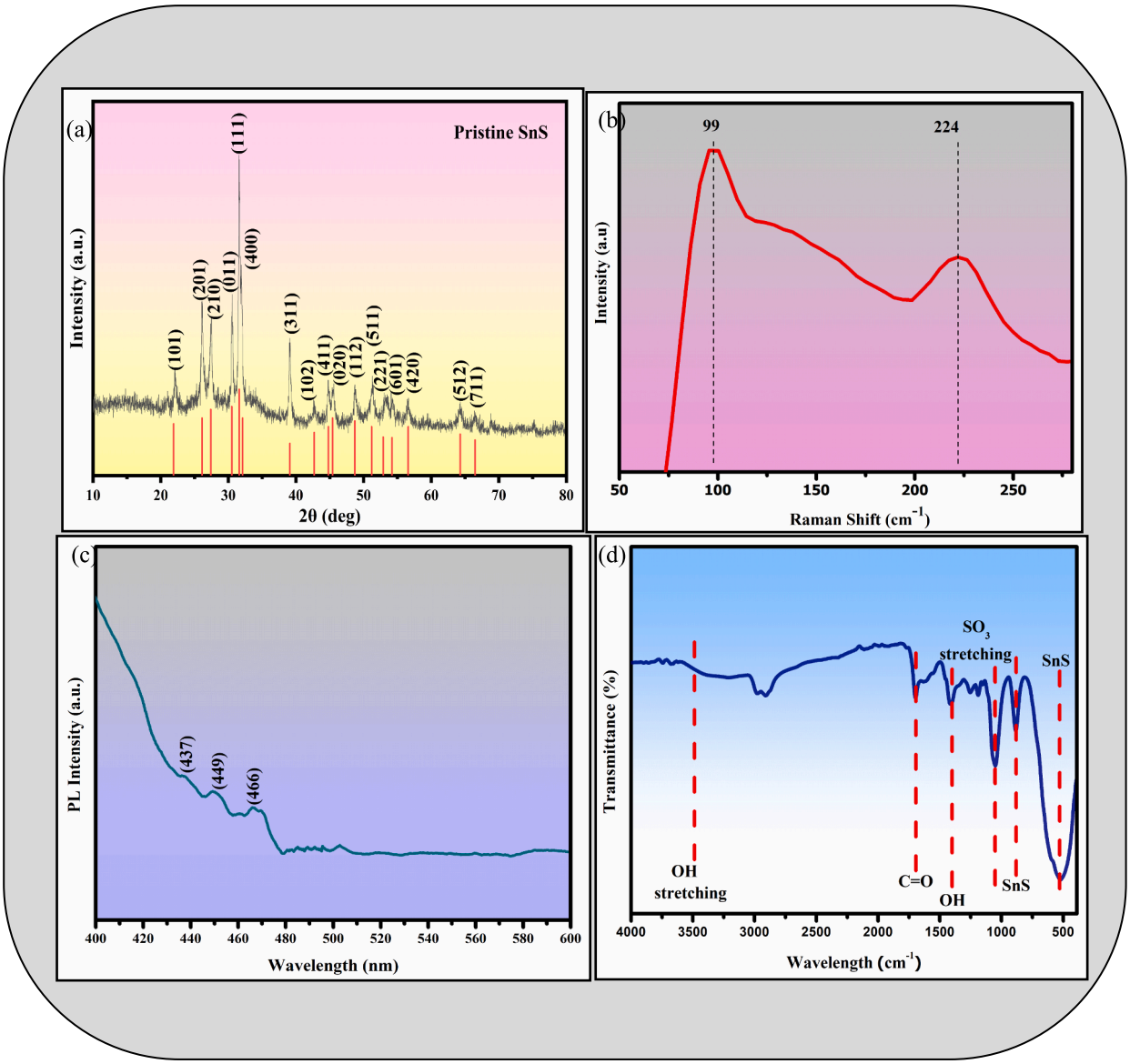



**Fig. 2.** (a) X-ray diffraction pattern, (b) Raman spectra, (c) Photoluminescence (PL) spectra and (d) FTIR spectra of the as synthesized SnS nanosheets.

**Table 1**
Provides d-spacing values calculated for various planes of SnS nanoflakes.

| Sample | 2θ (degree) | hkl planes | d spacing (nm) |
|---|---|---|---|
| SnS | 26.01 | (201) | 0.34 |
| | 30.5 | (011) | 0.29 |
| | 31.6 | (111) | 0.28 |
| | 39.1 | (311) | 0.23 |

## 5. Surface morphology and porosity analysis

### *5.1. Field emission scanning microscopy (FESEM) & electron dispersive X-ray spectroscopy (EDX)*

Recognizing the critical role of material microstructure in hydroelectric cell applications, Field Emission Scanning Microscopy (FESEM) serves as a vital tool for investigating the surface morphology of the synthesized sample. The morphology and structures of the SnS sample intended for HEC applications have been examined through Field Emission Scanning Microscopy (FESEM) images, showcased in Fig. 4(a). These images highlight the lamellar sheets of different thickness (100–150 nm). The sheets are stacked over one another, demonstrating the occurrence of agglomeration, which may be interpreted as a result of nanosheets attempting to minimize surface energy. The emergence of sheets showcases anisotropic growth and shaped by the unique surface energies inherent to distinct planes. The formation of nanostructures in sheet-like morphology reflects a competitive growth along different crystalline directions. Furthermore, the morphology reveals the presence of a porous network within the sample. Subsequent analysis of this porous network has been conducted using the BET technique. The sample's elemental composition was determined utilizing Energy Dispersive X-ray Spectroscopy (EDX), as shown in Fig. 4(e). The EDX spectra of sample indicated the presence of Sn and S without the presence of any other elements. Elemental mapping images of as prepared-sample further confirmed the uniform distribution of Sn and S in SnS. The Sn/S ratio of SnS nanosheets was determined to be 1.14.

### *5.2. High resolution transmission electron microscopy (HR-TEM)*

High Resolution Transmission Electron Microscopy (HR-TEM) is conducted by first sonicating a dispersion of SnS sample in ethanol and then loading a single drop of this dispersion onto a TEM copper grid.

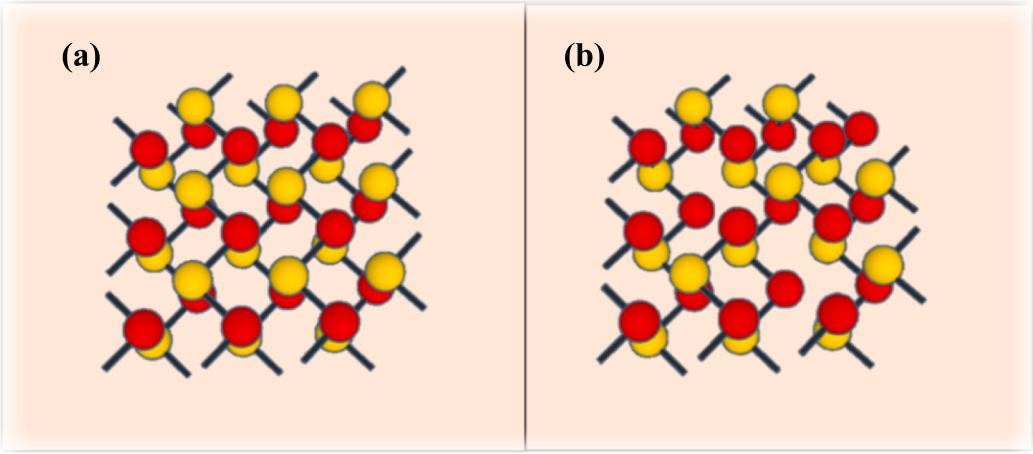



**Fig. 3.** Structure of SnS and S defective SnS (SnS + $V_S$) where S and Sn atoms are shown as yellow and red balls, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

sample's HR-TEM image was acquired to determine the nanostructures shape and size. Fig. 4(b). is the bright-field TEM micrograph of as prepared sample for HEC. The HR-TEM image reveals the sheet like morphology of the nanostructures. As shown in Fig. 4(c) these sheets are in contact with each other and also show some pin hall defects on its surface that promotes the flow of ions through hydroelectric cell. The SnS sheets produced here have small lateral length, grow randomly and contribute to an overall increase in the stacking length of layers. From Fig. 4(c), the measured inter planer distances are 0.28 nm, 0.34 nm, 0.29 nm, and 0.23 nm, corresponding to the (111), (201), (011), and (311) crystallographic planes of the orthorhombic SnS phase (PDF No. 39–0354). These nanosheets form a continuous network that facilitates ion flow in the HEC. In Fig. 4(d), the selected area electron diffraction (SAED) pattern further confirms that the measured inter planer distances correspond to the (201), (111), and (311) diffraction planes of the orthorhombic SnS phase (PDF No. 39–0354). The HR-TEM and FESEM results are in fair agreement with each other.

## 6. X-ray Photoelectron Spetroscopy (XPS)

XPS, a surface sensitive and non-destructive spectroscopic technique, has been used to explore the oxygen states and surface elements present in SnS nanosheets annealed at 300 °C. The XPS profile was recorded, and the findings are illustrated in Fig. 5. The survey spectra for SnS sample exhibited characteristic binding energy levels corresponding to Sn and S elements. Importantly, no peaks indicative of contaminants were detected, as shown in Fig. 5 (a).The high-resolution Sn 3d spectrum displayed two prominent peaks at 494.9 eV and 486.8 eV, corresponding to the Sn $3d_{3/2}$ and Sn $3d_{5/2}$ energy states, respectively, which are characteristic of $Sn^{2+}$ in SnS [55,56]. Additionally, the S 2p spectrum showed peaks at 168.96 eV (S $2p_{1/2}$) and 161.8 eV (S $2p_{3/2}$), confirming the presence of $S^{2-}$ in SnS [55,56]. It is noteworthy that the detection of $Sn^{2+}$ signals conclusively confirms that the $Sn^{2+}$ oxidation state is the dominant and exclusive oxidation state present in SnS nanosheets annealed at 300 °C, indicating the stability and purity of the material under these conditions.

Neither technique detected any peaks characteristic of the $SnS_2$ phase, further corroborating the purity of the SnS phase and the absence of $SnS_2$ under the given annealing conditions. This comprehensive analysis underscores the stability of the SnS phase at annealing temperature at 300 °C, validating the structural and chemical integrity of the sample for potential applications.

## 7. BET and BJH analysis

Surface area and pore size are the important parameter have significant role in water splitting. To get the specific surface area distribution in the as-prepared sample Bruanauer-Emmett-Teller (BET) analysis was done. The nitrogen ($N_2$) adsorption–desorption isotherm and Barret-Joyner-Halenda (BJH) derived pore size distribution of SnS nanosheets presented in Fig. 6(a). The SnS nanosheets exhibited type IV isotherms with distinct hysteresis at $P/P_o$ values ranging from 0.1 to 0.9, indicating the presence of mesoporous structures [45]. As depicted in Fig. 6(b), a wide distribution of pore sizes in the range of 50–1050 Å was observed in the SnS nanosheets. Analyses revealed that SnS posseses a mesoporous structure with an average pore diameter 17.07 nm and pore volume 4.9 X $10^{-2}$ $cm^3g^{-1}$, along with a BET surface area 6.45 $m^2$ $g^{-1}$. The high porosity observed in the SnS nanosheets is attributed to the removal of S at certain specific sites, providing more active sites for trapping the hydronium ions for further water splitting process.

## 8. Working performance of SnS based HEC

It is well known that metal oxides and metal sulfides have unsaturated or uncoordinated cations $M^{\delta+}$ and vacancies as $V^{\delta-}$ on their surfaces, which serve as active defect pair sites $M^{\delta+} - V^{\delta-}$ for the adsorption and chemidissociation of polar water molecules [16]. In the case of SnS, $Sn^{2+}$ cations and vacancies on the surface attract the oxygen atoms of water molecule. These sulphur vacancies trap the electrons and strongly attract $OH^-$ ions from the $H_2O$ molecule. As water reaches the defect pairs of electronegative surface cations and sulphur vacancies, it chemidissociates into hydronium and hydroxide ions. Upon dissociation, $OH^-$ ions water molecule gets physisorbed on the surface, while protons hop via hydrogen-bonded water molecules. The dissociated hydroxide and hydronium ions diffuse through SnS pellet and migrate towards their respective electrodes. Hydroxide ions react with Zinc electrode, forming zinc hydroxide and releasing electrons. These electrons migrate towards the silver electrode via external circuit. Meanwhile, hydronium ions diffuse towards the silver electrode and evolve hydrogen gas [46,47].

The occurrence of redox reactions on both electrodes is responsible for the generation of current and voltage in the cell. These are the electrochemical half-reactions that should occur on the cell's surface and its electrodes.

At the surface of SnS HEC.

$4H_2O \rightarrow 2H_3O^+ + OH^-$.

At the anode (At the zinc electrode).

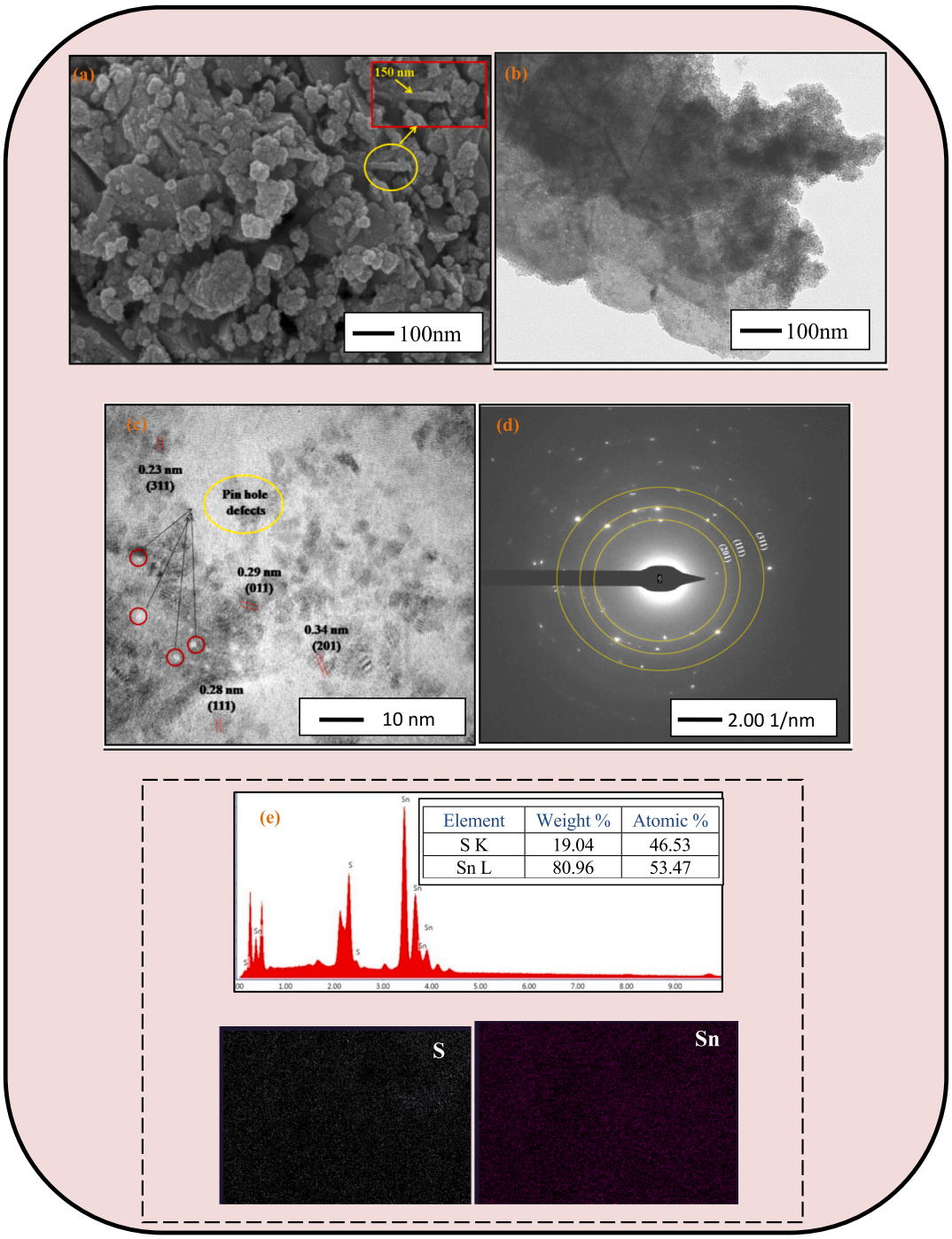



**Fig. 4.** (a) FESEM, (b) HRTEM images, (c) HRTEM with defects and d-spacing, (d) SAED pattern, and (e) EDX spectra and elemental mapping of SnS nanosheets.

$Zn + 2OH^- \rightarrow Zn(OH)_2 + 2e^-$ $E_{oxd}$ = -0.76 V.
At the cathode (At the silver electrode).
$2H_3O^+ + 2e^- \rightarrow H_2(g) + 2H_2O$ $E_{red}$ = + 0.22 V.

### 8.1. V-I polarization curve

The performance of SnS based HEC was assessed using the voltage-current polarization (V-I) curve. The three regions of V-I polarization curves of electrochemical cells (a) Low Current Density region (LCD) (b) Moderate Current Density Region (MCD) (c) High Current Density Region (HCD).

In the LCD region a rapid voltage drop is observed which is ascribed to as activation loss region. The cause of this loss is the potential required to overcome the energy barrier of the electrochemical reaction. In the MCD region the linear sweep is observed that represents the Ohmic losses in the HEC are also responsible for the cell's resistance to ionic flow. In the HCD region a sudden drop in voltage is observed due to abundant aggregation of ions on the surface of electrodes [48].This type of loss is termed as mass transport loss or concentration loss. All the above mentioned losses occurring in HEC are shown in Fig. 7 (a). represented by LM, MN, NO [49]. The SnS HEC delivers a higher offload current of 15 mA and potential produced is 0.74 V as shown in Fig. 7 (b). The power is calculated by using formula P = V*I, therefore the calculated power is 11.1 mW.

In order to give a comprehensive comparison, the prepared SnS based HEC is compared with other state of the art HECs. Comparison of

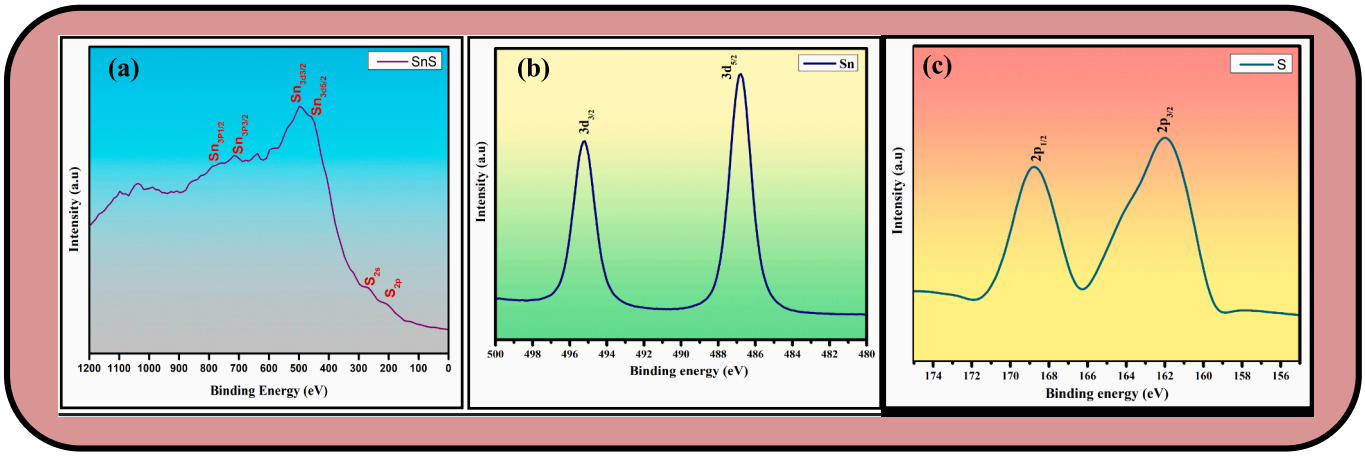



**Fig. 5.** Survey (a) and high-resolution ((b) Sn 3d and (c) S 2p) XPS profiles of SnS nanosheets.

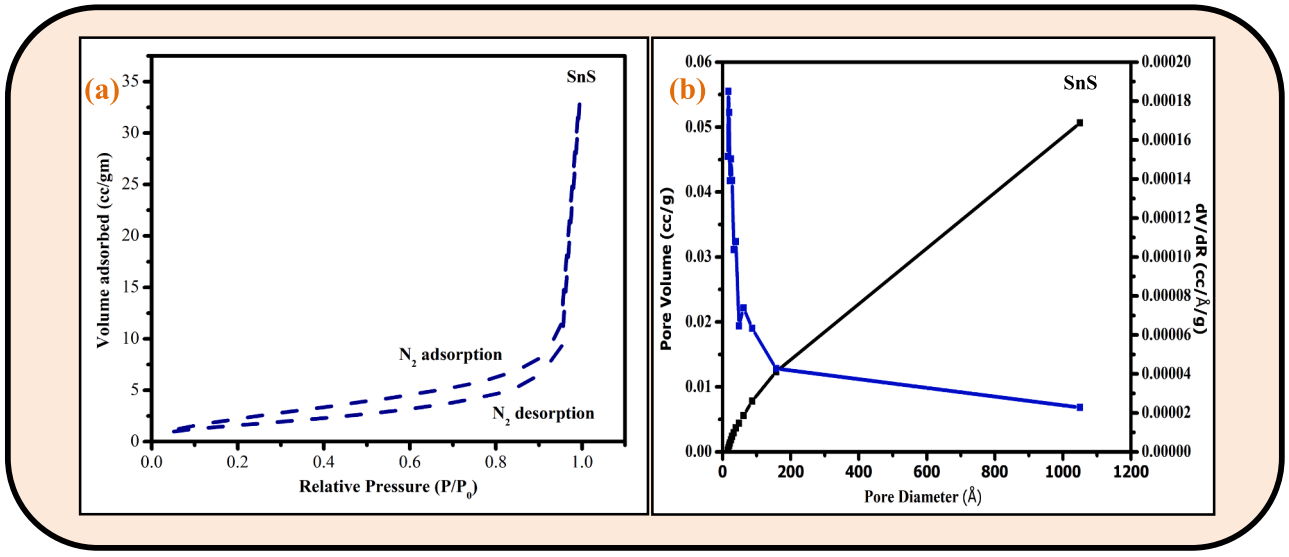



**Fig. 6.** (a) The $N_2$ adsorption and desorption isotherm and (b) pore size distribution of SnS.

results exhibiting relative value of current and voltage are reported in Table 2.

### 8.2. Stability, repeatability and application of SnS based HEC

In addition to the above mentioned parameters (i.e. current, voltage and power) a reliable HEC should demonstrate prolonged stability. The long term repeatability of SnS based HEC is illustrated in Fig. 8 (a). highlighting its consistency. In assessing the real time application of the SnS based HEC, we observed remarkable repeatability for 30 days with slightly drop in current with time. The HEC was employed to power 4 LED diodes by using few drops of water. This was achieved by configuring the 3 cells in series, while the LED diodes were arranged in parallel, as illustrated in Fig. 8 (b).

### 8.3. Cyclic voltammetry (CV)

The electrochemical behavior of SnS was further elucidated through cyclic voltammetry (CV) analysis of the as-prepared sample conducted at room temperature as shown in Fig. 9 (a). This analytical technique delivers crucial insights into the electrochemical reactions occurring at the surface of electrode's. In this study, SnS was employed as the working electrode, while platinum served as the counter electrode, and Ag/AgCl was utilized as the reference electrode. The experiments were carried out in a 0.3 M KOH electrolyte solution. During scanning, a potential window ranging from −0.6 to 1.5 V was applied to SnS, with a scan rate of 20 mV/s. The resulting CV curve exhibited a peak anodic current of 3.2 mA and a peak cathodic current of 5.3 mA. Notably, the peak anodic potential and peak cathodic potential of SnS were measured at 0.36 V and 0.29 V, respectively. The observed behavior of the CV curves indicates a faradic capacitive mechanism, characterized by reversible and rapid desorption/adsorption of ions at the electrode surface without inducing any phase change in the electrode material [50].

### 8.4. Electrochemical impedance spectroscopy (EIS)

Electrochemical impedance spectrosocopy was adopted to investigate the resistance of the SnS nanosheets prepared for HEC. The charge transport mechanism within HEC has been investigated through the Nyquist impedance response of the fabricated HEC. Fig. 9 (b). depicts the Nyquist plot of wet SnS HEC cell. The Nyquist plot of the SnS based HEC was obtained at an open circuit potential. As illustrated in the Nyquist plot, the diameter of the semicircle corresponds to the interfacial charge transfer resistance ($R_{ct}$), representing the resistance of the electrochemical reactions [51]. To identify the roles of the electrode,

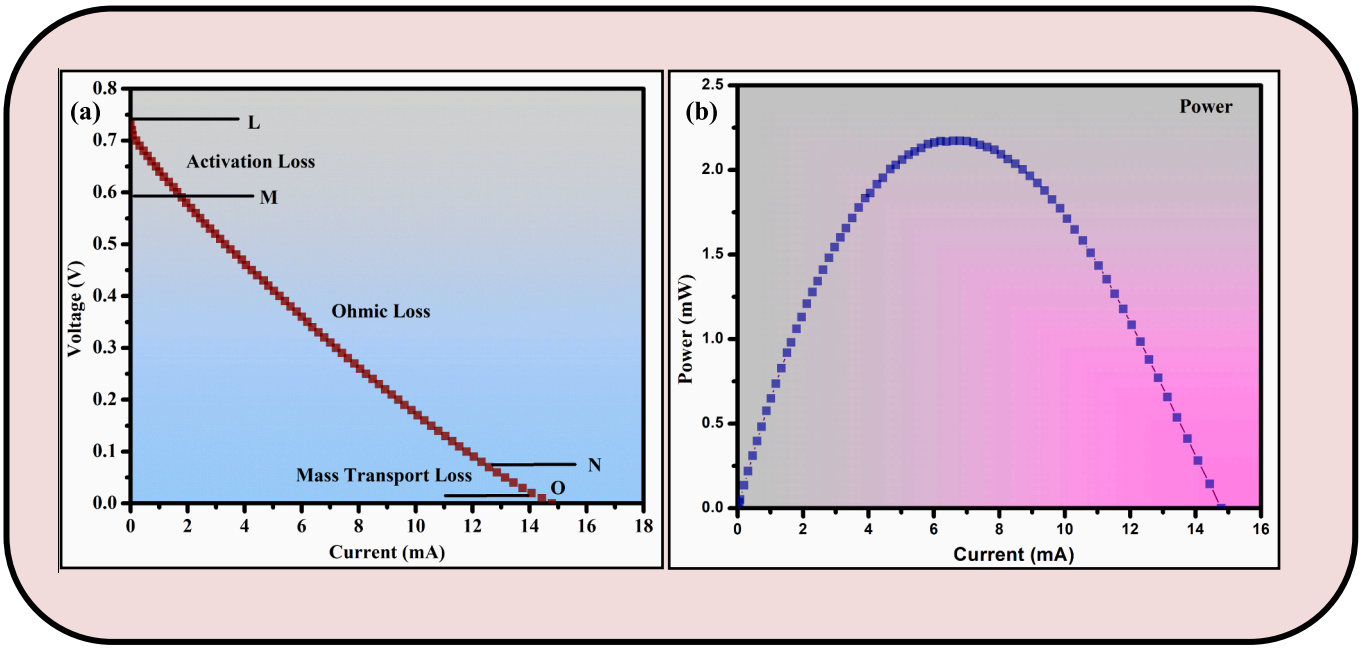



**Fig. 7.** (a) Room Temperature V-I polarization curve and (b) power of as fabricated SnS based HEC.

**Table 2**
Comparison of voltage and current produced by various fabricated Hydroelectric cells.

| Hydroelectric Cell | Max. Current (mA) | Voltage (V) | Reference |
|---|---|---|---|
| $MgGd_{0.02}Fe_{1.98}O_4$ | 6.8 | 0.85 | [52] |
| NiO | 12 | 1.03 | [53] |
| C doped $TiO_2$ | 7 | 0.74 | [54] |
| $SiO_2$ | 1.1 | 0.96 | [14] |
| MgO | 1.5 | 0.94 | [14] |
| $TiO_2$ | 2.3 | 0.90 | [14] |
| ZnO | 5.0 | 0.90 | [14] |
| SnS | 15 | 0.74 | Present work |

charge transfer, and ionic diffusion in SnS-based HEC, the experimental Nyquist data were fitted with an equivalent circuit, shown in the inset of Fig. 9 (b). The intercepts with the real axis at high frequencies represent the overall ohmic series resistance (Rs) of the cell. The Rs of SnS was 17.3 Ω, indicating a lower series resistance of SnS nanomaterials that leads to higher Jsc of 15.9 mA $cm^{-2}$. There are two well defined semi-circles in the middle in and low frequency ranges for the plot of SnS. The left arc represents the charge transfer occurring at the Zinc electrode and pellet interface. The $R_{ct1}$ can be calculated by fitting the left arc, and found to be 5.27 Ω indicating good electrochemical activity of SnS. The right arc indicates the finite layer Nernst mass diffusion impedance within the pellet. SnS demonstrates a higher $R_{ct2}$ value (7.32 Ω), suggesting deficient diffusion, thereby impending ion transport and elevates electrode resistance.

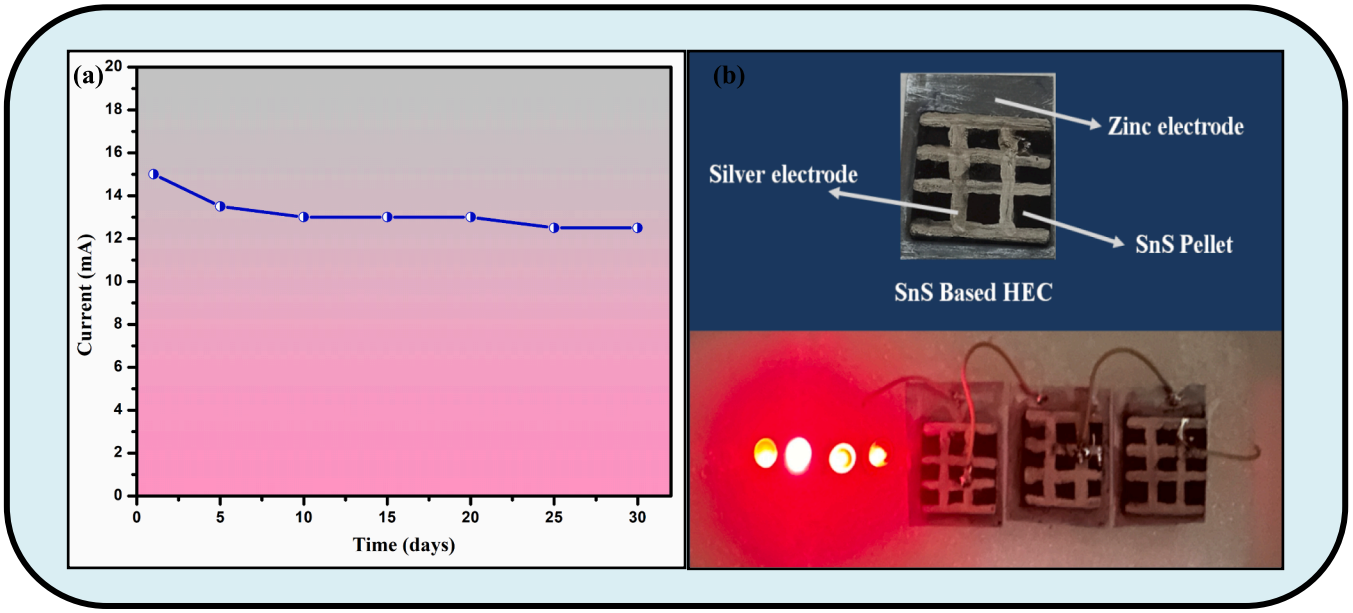



**Fig. 8.** (a). Represents repeatability and stability, and (b) shows the application of SnS based HEC device for powering LED lights.

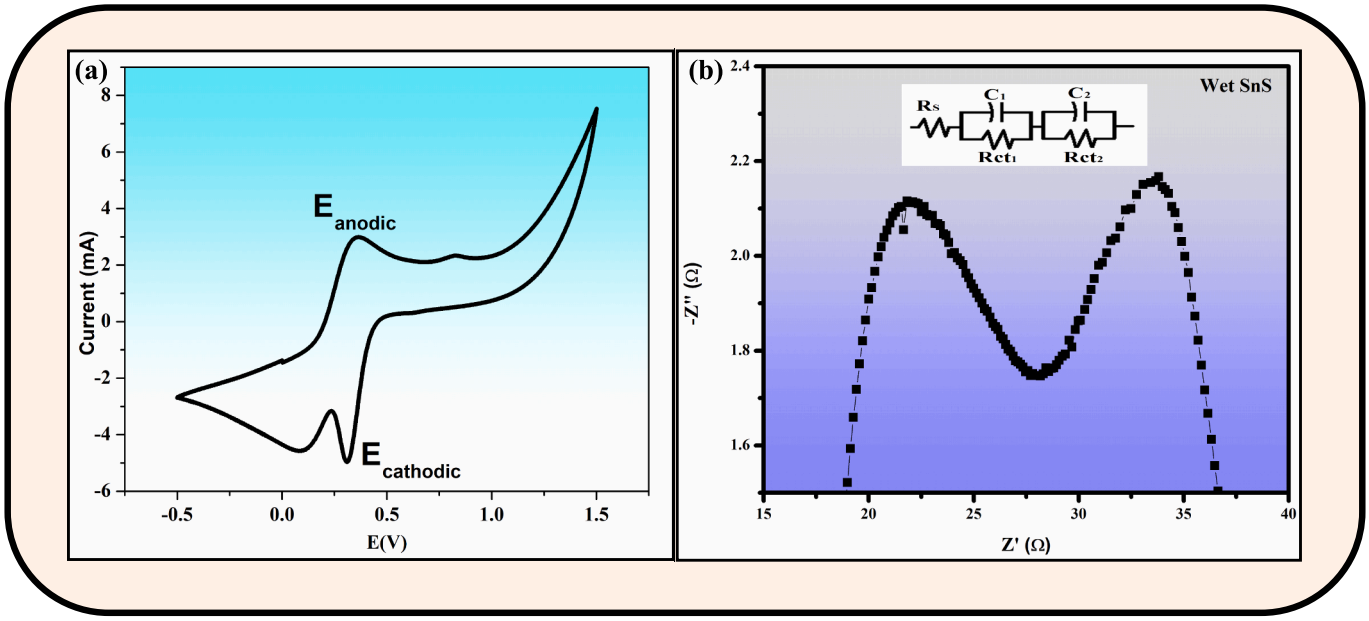



**Fig. 9.** (a) CV curve of SnS electrode in 0.3 M KOH solution and (b). Nyquist curve of wet SnS HEC.

## 9. Conclusion

In summary, SnS based hydroelectric cell was prepared by low temperature solvothermal synthesis of 2D SnS. The nanostructure having orthorhombic phase which was confirmed by X-ray Diffraction pattern. The specific surface area obtained by BET analysis for SnS is 6.45 $m^2 g^{-1}$ with pore diameter being approximately 17.07 nm. Nyquist curve for SnS based HEC in wet state exhibit ionic diffusion in the cell and also shows the high reduction in the impedance of the cell under the wet conditions. The fabricated HEC of cross section area 4 $cm^2$ exhibited an off-load power output of 11.1 mW. The HEC showcases an impressive electric current output of 15 mA, demonstrating its capability to sustain stable performance with just a few drops of water. Furthermore, its remarkable persists consistently for 30 days, affirming its reliability and durability in various operational conditions. The power output produced via water splitting by HEC is encouraging and proves to be green energy source at low cost. These findings may pave the way for its potential deployment in innovative applications. Electricity generation by HEC device does not emit any toxic gases therefore it can be the best strategy to save the environment and to achieve mission of net carbon zero. These ground-breaking results emphasize the practicality of HEC device as alternative to solar cells and fuel cells.

**Data and code avaibility:** The data presented in this article will be made available on reasonable request.

## CRediT authorship contribution statement

**Hemant K. Arora:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jyoti Shah:** Writing – review & editing, Investigation, Formal analysis, Data curation, Conceptualization. **R.K. Kotnala:** Writing – review & editing, Validation, Resources, Investigation, Data curation, Conceptualization. **A.S. Rao:** Data curation, Resources, Validation, Visualization. **Nitin K. Puri:** .

## 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.

## Acknowledgements

The authors express their gratitude to Delhi Technological University, Delhi, for providing research facilities. They also extend their sincere thanks to the CSIR National Physical Laboratory-Delhi for providing instrumental facilities, and to JNU- Delhi for the Raman facility.

### *Author Contributions*

Nitin k. Puri: Conceptualization, Investigation, Methodology, Resources, Supervision, Writing – review & editing. Hemant K. Arora: Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. Jyoti Shah: Conceptualization, Data curation, review & editing, Resources. R.K. Kotnala: Conceptualization, Data curation, review & editing, Resources, Technical discussions.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mseb.2024.117949.

## Data availability

Data will be made available on request.

## References:

[1] S. Chu, A. Majumdar, Opportunities and challenges for a sustainable energy future, *Nature* 488 (7411) (2012) 294–303.
[2] P.M. Cox, R.A. Betts, C.D. Jones, S.A. Spall, I.J. Totterdell, Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model, *Nature* 408 (6809) (2000) 184–187.
[3] J.A. Turner, A realizable renewable energy future, *Science* 285 (5428) (1999) 687–689.
[4] J.A. Turner, Sustainable hydrogen production, *Science* 305 (5686) (2004) 972–974.
[5] W. Wang, Xu. Xiaomin, W. Zhou, Z. Shao, Recent progress in metal-organic frameworks for applications in electrocatalytic and photocatalytic water splitting, *Adv. Sci.* 4 (4) (2017) 1600371.
[6] X. Du, T. Zhao, Z. Xiu, Z. Xing, Z. Li, K. Pan, S. Yang, W. Zhou, $BiVO_4$@ $ZnIn_2S_4$/$Ti_3C_2$ MXene quantum dots assembly all-solid-state direct Z-Scheme photocatalysts for efficient visible-light-driven overall water splitting, *Appl. Mater. Today* 20 (2020) 100719.

[7] R.K. Kotnala, J. Shah, Green hydroelectrical energy source based on water dissociation by nanoporous ferrite, *Int. J. Energy Res.* 40 (12) (2016) 1652–1661.

[8] J. Shah, S. Jain, A. Shukla, R. Gupta, R.K. Kotnala, A facile non-photocatalytic technique for hydrogen gas production by hydroelectric cell, *Int. J. Hydrogen Energy* 42 (52) (2017) 30584–30590.

[9] A. Gaur, Jyoti Shah, R.K. Kotnala, D. Kumar, Development of Mg-doped hematite ($\alpha$-$Fe_2O_3$)-based hydroelectric cell to generate green electricity, *New J. Chem.* 46 (44) (2022) 21158–21166.

[10] S. Saini, Jyoti Shah, R.K. Kotnala, K.L. Yadav, Nickel substituted oxygen deficient nanoporous lithium ferrite based green energy device hydroelectric cell, *J. Alloy. Compd.* 827 (2020) 154334.

[11] Manash, Aniket, Rakesh Kumar Singh, Vivek Kumar, Jyoti Shah, Shashank Bhushan Das, Singh Sonu Kumar, Nishant Kumar, and R. K. Kotnala. "Studies on structural and magnetic properties of nanoporous Li+ substituted $MgFe_2O_4$ nanomaterials for its application in hydroelectric cell with other areas of science & technology." *Materials Today: Proceedings* 80 (2023): 1002-1013.

[12] R. Gupta, J. Shah, R. Das, S. Saini, R.K. Kotnala, Defect-mediated ionic hopping and green electricity generation in $Al_{2-x}Mg_xO_3$-based hydroelectric cell, *J. Mater. Sci.* 56 (2021) 1600–1611.

[13] P. Kumar, S. Vashishth, I. Sharma, V. Verma, Porous $SnO_2$ ceramic-based hydroelectric cells for green power generation, *J. Mater. Sci. Mater. Electron.* 32 (1) (2021) 1052–1060.

[14] R.K. Kotnala, R. Gupta, A. Shukla, S. Jain, A. Gaur, J. Shah, Metal oxide based hydroelectric cell for electricity generation by water molecule dissociation without electrolyte/acid, *J. Phys. Chem. C* 122 (33) (2018) 18841–18849.

[15] R. Gupta, J. Shah, R. Singh, R.K. Kotnala, Nonphotocatalytic water splitting process to generate green electricity in alkali doped zinc oxide based hydroelectric cell, *Energy Fuel* 35 (11) (2021) 9714–9726.

[16] U. Bhakar, A. Agarwal, S. Sanghi, J. Shah, R.K. Kotnala, Production of green electricity from strained $BaTiO_3$ and $TiO_2$ ceramics based hydroelectric cells, *Mater. Chem. Phys.* 262 (2021) 124277.

[17] J. Shah, K.C. Verma, A. Agarwal, R.K. Kotnala, Novel application of multiferroic compound for green electricity generation fabricated as hydroelectric cell, *Mater. Chem. Phys.* 239 (2020) 122068.

[18] Mohan, Visakh V., K. C. Akshaya, and K. P. Vijayakumar. "Fabrication of p-type and n-type SnS thin films through vacuum-free deposition techniques." *Materials Today: Proceedings* (2023).

[19] P. Li, X. Wang, S. Yang, Preparation of core-shell NiCo-LDH@ NiCo-hydroxysulfide by two-step electrodeposition for high-performance supercapacitors, *J. Storage Mater.* 97 (2024) 112838.

[20] L.i. Lv, L. Lei, Q.-W. Chen, C.-L. Yin, H. Fan, J.-P. Zhou, Oxygen vacancies-modified S-scheme heterojunction of Bi-doped La2Ti2O7 and La-doped Bi4Ti3O12 to improve the NO gas removal avoiding NO2 product, *Appl Catal B* 343 (2024) 123464.

[21] S.-L. Li, K. Tsukagoshi, E. Orgiu, P. Samorì, Charge transport and mobility engineering in two-dimensional transition metal chalcogenide semiconductors, *Chem. Soc. Rev.* 45 (1) (2016) 118–151.

[22] Costa, Beatriz NL, Filipe Camarneiro, Ana Marote, Catarina Barbosa, Carlos Vedor, Diogo Tomé, Filipa J. Costa et al. "Functionalized Nanodiamonds for Targeted Neuronal Electromagnetic Signal Detection." *ACS Applied Materials & Interfaces* 16, no. 44 (2024): 60828-60841.

[23] S.F. Wang, W.K. Fong, W. Wang, C. Surya, Growth of highly textured SnS on mica using an SnSe buffer layer, *Thin Solid Films* 564 (2014) 206–212.

[24] M.F. Afsar, M.A. Rafiq, A.I.Y. Tok, Two-dimensional SnS nanoflakes: synthesis and application to acetone and alcohol sensors, *RSC Adv.* 7 (35) (2017) 21556–21566.

[25] T. Li, W. Li, W. Long, H. Zhang, H. Tan, Y. Zhan, Z. Ren, Flake-like SnS anchored on the carbon layer for supercapacitor electrode with photo-responsive capacitance enhancement, *J. Alloy. Compd.* 976 (2024) 173377.

[26] L.-L. Wang, H.-Y. He, Q. Li, Z.-Q. Ren, Photocatalytic activities of SnS-reduced graphene oxide by the photodegradation of malachite green in water, *Mater. Res. Innov.* 20 (6) (2016) 458–464.

[27] Lu, Xiu Fang, Yishu Zhang, Naizhou Wang, Sheng Luo, Kunling Peng, Lin Wang, Hao Chen et al. "Exploring low power and ultrafast memristor on p-type van der Waals SnS." *Nano letters* 21, no. 20 (2021): 8800-8807.

[28] U. Chalapathi, B. Poornaprakash, S.-H. Park, Chemically deposited cubic SnS thin films for solar cell applications, *Sol. Energy* 139 (2016) 238–248.

[29] F. Jamali-Sheini, M. Cheraghizade, R. Yousefi, SnS nanosheet films deposited via thermal evaporation: the effects of buffer layers on photovoltaic performance, *Sol. Energy Mater. Sol. Cells* 154 (2016) 49–56.

[30] A.J. Biacchi, S.T. Le, B.G. Alberding, J.A. Hagmann, S.J. Pookpanratana, E. J. Heilweil, C.A. Richter, A.R. Hight, Walker, Contact and noncontact measurement of electronic transport in individual 2d SnS colloidal semiconductor nanocrystals, *ACS Nano* 12 (10) (2018) 10045–10060.

[31] M. Ristov, G.J. Sinadinovski, I. Grozdanov, M. Mitreski, Chemical deposition of tin (II) sulphide thin films, *Thin Solid Films* 173 (1) (1989) 53–58.

[32] T.H. Sajeesh, A.R. Warrier, C. Sudha Kartha, K.P. Vijayakumar, Optimization of parameters of chemical spray pyrolysis technique to get n and p-type layers of SnS, *Thin Solid Films* 518 (15) (2010) 4370–4374.

[33] Manh Hung, Nguyen, Chuong V. Nguyen, Vinaya Kumar Arepalli, Jeha Kim, Nguyen Duc Chinh, Tien Dai Nguyen, Dong-Bum Seo, Eui-Tae Kim, Chunjoong Kim, and Dojin Kim. "Defect-induced gas-sensing properties of a flexible SnS sensor under UV illumination at room temperature." *Sensors* 20, no. 19 (2020): 5701.

[34] M.B. Zaman, R. Poolla, S.A. Khandy, A. Modi, R.K. Tiwari, Thioglycolic acid assisted hydrothermal growth of SnS 2D nanosheets as catalysts for photodegradation of industrial dyes, *Nanotechnology* 32 (24) (2021) 245706.

[35] B.H. Baby, D. Bharathi Mohan, Synthesis and Raman Analysis of SnS Nanoparticles Synthesized by PVP Assisted Polyol Method vol. 1665, no. 1 (2015).

[36] H.R. Chandrasekhar, R.G. Humphreys, U. Zwick, M. Cardona, Infrared and Raman spectra of the IV-VI compounds SnS and SnSe, *Phys. Rev. B* 15 (4) (1977) 2177.

[37] S.H. Chaki, M.D. Chaudhary, M.P. Deshpande, Synthesis and characterization of different morphological SnS nanomaterials, *Adv. Nat. Sci. Nanosci. Nanotechnol.* 5 (4) (2014) 045010.

[38] I.P. Parkin, L.S. Price, T.G. Hibbert, K.C. Molloy, The first single source deposition of tin sulfide coatings on glass: aerosol-assisted chemical vapour deposition using [Sn (SCH2CH2S) 2], *J. Mater. Chem.* 11 (5) (2001) 1486–1490.

[39] B.H. Baby, D. Bharathi Mohan, Phase Formation Study of SnS Nanoparticles Synthesized through PVP Assisted Polyol Method 360 (2018) 012003.

[40] E. Kharatzadeh, M. Khademalrasool, GO nanosheets decorated with SnS nanoparticles: excellent photocatalytic performance under visible-light irradiation, *PCCP* 25 (25) (2023) 17043–17057.

[41] R. Ariappan, T. Mahalingam, V. Ponnuswamy, Preparation and characterization of electrodeposited SnS thin films, *Optik* 122 (24) (2011) 2216–2219.

[42] M. Kul, Electrodeposited SnS film for photovoltaic applications, *Vacuum* 107 (2014) 213–218.

[43] J. Henry, K. Mohanraj, S. Kannan, S. Barathan, G. Sivakumar, Structural and optical properties of SnS nanoparticles and electron-beam-evaporated SnS thin films, *J. Exp. Nanosci.* 10 (2) (2015) 78–85.

[44] M.A. Dar, D. Govindarajan, G.N. Dar, Facile synthesis of SnS nanostructures with different morphologies for supercapacitor and dye-sensitized solar cell applications, *J. Mater. Sci. Mater. Electron.* 32 (15) (2021) 20394–20409.

[45] S. Nawaz, Y. Khan, S.AM. Abdelmohsen, S. Khalid, E.M. Björk, M.A. Rasheed, M. Siddiq, Polyaniline inside the pores of high surface area mesoporous silicon as composite electrode material for supercapacitors, *RSC Adv.* 12 (27) (2022) 17228–17236.

[46] J. Shah, A. Shukla, R.K. Kotnala, Highly accelerated, sustainable, abundant water splitting at room temperature generating green electricity by Sb-doped SnO2 hydroelectric cell, *ACS Sustain. Chem. Eng.* 9 (45) (2021) 15229–15238.

[47] S. Badola, J. Shah, A. Gaur, D.S. Satish Khasa, T.K. Rawal, A.K. Mandal, Srivastava, R.K. Kotnala, "Strategic enhancement of oxygen defects in ZnO from ZnS for water splitting to generate green electricity by hydroelectric cell." *Applied, Mater. Today* 34 (2023) 101904.

[48] H. Abdi, B. Mohammadi-ivatloo, S. Javadi, A.R. Khodaei, E. Dehnavi, Energy Storage Systems Distributed Generation Systems: Design, Operation and Grid Integration (2017) 00007–00008.

[49] Frano Barbir. PEM Fuel Cells: Theory and Practice; Dorf, R. C., Ed.; Elsevier Academic Press, 2012.

[50] L.M. Johny, R. Sriram, N.S. Nirmala Jothi, P. Sagayaraj, Single crystalline SnO2 nanorods functionalized with TiO2 nanospheres and their electrochemical properties, *Ceram. Int.* 44 (12) (2018) 14471–14479.

[51] F. Zhang, X. Yang, H. Wang, M. Cheng, J. Zhao, L. Sun, Structure engineering of hole–conductor free perovskite-based solar cells with low-temperature-processed commercial carbon paste as cathode, *ACS Appl. Mater. Interfaces* 6 (18) (2014) 16140–16146.

[52] I.A. Parray, S.A. Ali, J. Shah, R.K. Kotnala, Green energy generation via water splitting by non-photocatalytic process based on Gd doped magnesium ferrite hydroelectric cells, *J. Alloy. Compd.* 999 (2024) 174951.

[53] A. Saini, R. Kashyap, J. Rani, R.K. Sharma, Electrical energy of the order of 24 W/m2 from water splitting using nickel oxide: effect of synthesis process on power output, *Results Chem.* 5 (2023) 100911.

[54] S.B. Kajal, A. Shukla, J. Shah, Deepika, R. Kumar, Increased Green Electricity Generated in Carbon-Doped TiO2-Based Hydroelectric Cell, *J. Electron. Mater.* (2024) 1–10.

[55] Chastain, J., J. F. Moulder, and W. F. Stickle. "Handbook of x-ray photoelectron spectroscopy: a reference book of standard spectra for identification and interpretation of XPS data. 1992." *There is no corresponding record for this reference* (1995).

[56] S. Cheng, G. Conibeer, Physical properties of very thin SnS films deposited by thermal evaporation, *Thin Solid Films* 520 (2) (2011) 837–841.

# Supposititious Sarcasm Detection and Sentiment Analysis Coping Hindi Language in Social Networks Harnessing Zipf- Mandelbrot Probabilistic Optimisation and Perplexity Entropy Learning

HIMANI POKHRIYAL[1]

Applied Mathematics, Delhi Technological University, new delhi, new delhi, India, himpokh1698@gmail.com

GOONJAN JAIN[2]

Applied Mathematics, Delhi Technological University, Delhi, delhi, India, goonjan_jain@hotmail.com

Sarcasm is used to convey the contempt via poking and ridiculing the other person. It is frequently used to make fun of other people by saying unpleasant things. The presence of informal language patterns, modified vocabulary, and speak-text utilized in online communities makes sarcasm detection a challenging task. Due to the occupancy of these terms in the conversation it is challenging to understand and identify sarcasm. Further, sarcasm detection in Hindi, a native Indian language is itself a challenging task to perform. This study proposes a novel method for sarcasm detection in Hindi comments taken from social networking sites using probabilistic non-linear optimization technique. The encapsulations of word embedding algorithms infused in an optimization problem is used for sarcasm detection. The Hindi datasets i.e., News Headlines and Twitter comments are utilized for sarcasm detection. The outcomes of proposed method is assessed using a variety of learning metrics such as accuracy, precision, recall, and F1 score surpassing the existing state-of-the-art methods. Performing ablation study and statistical validation tests within real-world social network confirms the efficacy and reliability of the proposed method. The proposed method gives an outstanding F1 scores of 0.9316% and 0.9359% on Dataset-1 and 2.

**CCS Concepts** • Computing methodologies → Artificial intelligence ~ Natural language processing • Mathematics of computing → Mathematical analysis ~ Mathematical optimization ~ Continuous optimization

**Additional Keywords and phrases**: Sarcasm detection, HindiSentiWordNet, Natural language processing, Optimisation technique, Social media

## 1 Introduction

Sarcasm is a figure of speech used to indirectly contempt or mock another person. Sarcasm conveys the presence of interconnection between positive and negative intent in the sentence. Analysing the implicit intent of target text and sentiment cues with contradictory patterns helps in sarcasm detection. For example, consider a sarcastic sentence "*I love working at the office on holidays*". This sentence appears to be carrying a positive implicit intent demonstrated by the target word "*love*". However, the hidden negative intent is conveyed by the phrase "*working at the office on holidays*". The sentence shows that the pairs "*love*" - "*office*" and "*holidays*" - "*working*" act as contrary word links. To support accurate sarcasm, understanding the intonation of sarcasm in the user's tone also plays an essential role. Subsequently, the implicit intent of expression with intonation gives rise to sarcasm detection as it often relies on subtle contextual cues and indirect meanings rather than explicit sense. Sarcasm detection helps to prevent conflicts and misinterpretations. Understanding the tone and dynamics of sarcastic conversation provides valuable insights for researchers, marketers, and policymakers. The interactions, conversations, and trends on social media are analyzed more effectively with accurate sarcasm detection.

India's official language is Hindi and a sizable portion of the population in various states speaks it in their native tongue [1]. Despite the prevalence of the Hindi language in the Indian subcontinent it follows the free-word order arrangement of the words. Fig. 1 shows three sentences that depict similar meanings with different free word order rearrangements. The existence of varied tonal traits with complex features in sentence led to difficulty in understanding the sarcasm in Hindi content [2]. Most of the work in sarcasm is

[1] First Author
[2] Corresponding author

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM 2375-4699/2025/01-ART
http://dx.doi.org/10.1145/3712061



Just Accepted

done in English [3] [4]. However when it comes to Hindi there are not as many resources available to fully capture the aspects of the figure of speech [5]. So investigations and experiments have been somewhat restricted for sarcasm detection in Hindi [6]. The field of sarcasm detection in Hindi has been limited by the absence of categorized datasets, lesser sentiment ratings, and certain absence of analysis tools such as part-of-speech (POS) taggers [7] [8].

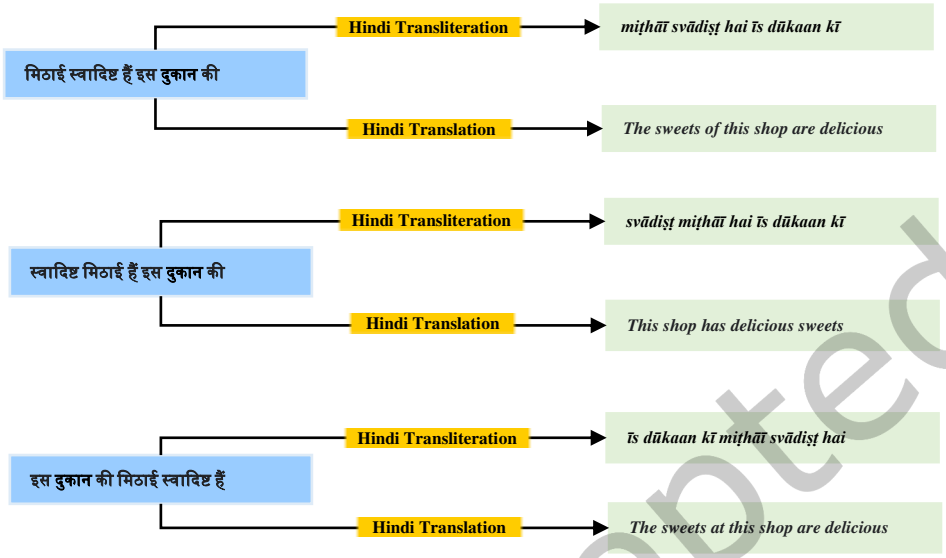



**Fig. 1.** Example of Hindi comments as in transliterated and translated text

The goal of this study is to perform sarcasm detection using mathematical optimization technique in Hindi text. Fig. 2 and 3 are taken from Twitter[3] demonstrates two examples of Hindi comments, showcasing the presence of sarcasm.




**Fig. 2.** Snapshot of Hindi comment taken from Twitter

Various approaches for sarcasm detection have been proposed in the literature, involving methods using supervised learning [9] ensemble approaches [10] and deep learning networks [11]. However, these methods do not show affective indicators of sarcasm in Hindi with missing interconnections between intricate sentimental cues and the implicit intent of the target text. Due to this, they led to give low recall. Certain studies with supervised methods [11] [6] [12] provide sarcasm detection but are unable to detect the presence of implicit intent of expression with intonation affecting sarcasm. Meanwhile, sarcasm detection using an ensemble supervised model [13] neglects the sentiment cues of sentence with a low accuracy score.

[3] https://twitter.com/

Just Accepted

This paper investigates the use of mathematical optimization technique in combination with word embedding for sarcasm detection. The proposed method follows three core steps to perform sarcasm detection. In the first step, we calculate four sentence scores and average embedding score using word embedding. In the second step, we use HindiSentiWordNet (HSWN) to perform the sentiment analysis. Additionally, we identify the triggered indicator of sarcasm using the key-phrase extraction and procure highest key-phrase value. In the final step, we formulate a non-linear optimization problem infusing all the calculated scores in preceding steps into it. The final optimal score of the defined optimization problem is used for sarcasm detection in Hindi text.




**Fig. 3.** Snapshot of Hindi comment taken from Twitter

***Motivation to perform sarcasm detection in Hindi:*** In the proposed method, we build an effective and reliable method for sarcasm detection in a poorly resourced Hindi text using a mathematical optimization technique.

The contributions of the proposed method are as follows:

1. We propose an automated sarcasm detection method titled "Simulated Probabilistic Optimisation Zipf-Mandelbrot Distribution via Entropy Paradigm for Sarcasm Detection - (SPOZ-ESD)". This proposed method is used for sarcasm detection in Hindi comments taken from social media.
2. The results of the SPOZ-ESD are validated using four statistical validation tests. This shows the consistency in results and confirms the efficacy of the proposed method. The ablation study is conducted to assess the effectiveness of defined hyperparameters used in the proposed study.
3. The reliability of the proposed method (SPOZ-ESD) is confirmed by the promising accuracy, precision, recall and F1 scores on the two datasets. The performance of the proposed method outperforms the existing state-of-the-art methods in terms of all evaluation metrics.

The subsequent sections of this paper are structured as follows: *Section 2* gives an overview of the related work conducted in the field of sarcasm detection. *Section 3* delves into a detailed explanation of the proposed method for sarcasm detection in Hindi. *Section 4* focuses on the employed experimental settings and the subsequent results obtained for the proposed method. The paper concludes in *Section 5* with defining the future work.

## 2 Related Work

We have seen the abundance used of sarcasm detection in English. Riloff et al. [14] introduced a bootstrapping technique that helped to recognize text with positive and negative situation phrases from sarcastic tweets. Ptáček et al. [15] evaluated long short-term memory (LSTM) and support vector machine (SVM) classifiers with different feature combinations on multi-lingual datasets in Czech and English. Tan et al. [16] provided the use of the deep neural network (DNN) for the task of sarcasm detection. Kalchbrenner et al. [17] utilized the convolutional neural network (CNN) with embedded knowledge to detect sarcasm. Shen et al. [18] suggested a method for sarcasm detection using a dynamic convolutional neural network (DCNN) with max-pooling. Zhang et al. [19] depicted sarcasm detection using encoder-decoder architecture M2Seq2Seq with

Just Accepted

inter and intra-modality attention. Sharma et al. [20] conveyed the use of fuzzy logic applied on word2vec, glove and bi-directional encoder representations from transformer (BERT) model for sarcasm identification. Vitman et al. [21] provided dual-channel network (DC-Net) which detect sarcasm with context emotion traits using DNN and CNN composition with transformer model. Their major objective included noticing the recently derived disparity characteristic and incorporating it supervised models with transformer. All these methods subsequently detect sarcasm but ignore the showcasing of sentiment cue features with tonal traits in sentences. Zhang et al. [22] gave a self-adaptive representational model for sarcasm detection. Punetha and Jain [23] demonstrated optimal model-based sentiment classification via the optimization technique.

Various multimodal-based sarcasm detection methods have been defined in past studies. They signified the use of the MUStARD dataset to detect sarcasm with the help of historical background comments from TV shows [24] [25]. Sarcasm detection has been given use for graph neural networks (GNN), a graph-structured data structure given by Wu et al. [26]. Past studies revealed that English has served as the language of choice for a majority of investigations on sarcasm detection. Several studies are available in Arabic, Dutch, Urdu, and Chinese. Galal et al. [27] gave sarcasm detection in Arabic using an enhanced machine-learning algorithm. Hassan et al. [28] provided sarcasm detection in Urdu by employing a hybrid approach with a transformer-based technique. To identify sarcasm in Arabic posts, Alayba et al. [29] fused LSTM and CNN models. Liu et al. [30] gave a multi-strategy grouping methodology (MSELA) for handling unbalanced Chinese dataset features for sarcasm detection. Similar research has been carried out in linguistics, like Indonesian [31] and Spanish [32] languages, for depicting sarcasm using supervised-based approaches.

Research on code-mix social networking material, has recently been conducted in relation to the low-resource language Hindi. There had a dearth of research on texts written exclusively in Hindi. One of the first studies for sarcasm detection in Hindi has published by Desai and Dave [7]. They created a collection of Hindi statements that are sarcastic and trained an SVM based classifier using different semantic attributes, such as punctuation marks and polarisation. The model divided the statements into five groups according to how sarcastic the sentences are but failed to demonstrate the presence of implicit intonation and sentimental cues. Bharti et al. [8] employed a contextualized method of part-of-speech tagging with the hidden Markov model (POS-HMM) to ascertain whether a specific tweet is considered as sarcastic or not. It included counting the amount of positive and negative keywords in the news and comments. It employed an established set of Hindi words with orientation values. Another pattern-based structure has been put forward by Bharti et al. [33], who used the discrepancies inherent in the equivalent remarks and time-based facts to anticipate sarcasm in Hindi tweets. Additionally, Katyayan et al. [34] investigated the detection of sarcasm in 1,000 Hindi statements that had been taken from social media sites like Facebook, Instagram, and Twitter. Their work utilized POS tagger and BOW techniques for feature extraction. It explored three basic machine learning techniques, namely naive Bayes, SVM, and decision tree for sarcasm classification, but failed to give the interpretation with poor reliability. Bharti et al. [35] used another parsing-based context technique (PBCT) to count the number of positive and negative keywords in sentences. This is used a predefined list of Hindi words, with the polarity value, to determine if the given tweet was considered sarcastic or not. Further supervised learning-based method TANA using LSTM-SVM with neural network architecture [36] depicted sarcasm. Kumar et al. [37] provided the sarcasm detection method using a CNN-LSTM device. Bharti, Babu, and Raman [38] provided a context-based pattern for sarcasm detection in Hindi tweets. They utilized the rule-based method of the context-based parsing system (CBPSD). This defined approach utilized Hindi news as the context of a tweet with a lexicon tool but contradicted what was specified as a sarcastic sentence. Multi lingual sarcasm detection in Algerian given by Bousmaha et al. [39] used deep learning method with EasyOCR system. Whereas, Marreddy et al. [40] provided BERT based method for Telugu based sentiment analysis and Ahmed et al. [41] used unsupervised lexicon model with embeddings to learn NLP patterns of satire contexts. Deep learning based on multi-head graph convolutional network (MH-GCN) and LSTM-CNN have been utilised for sentiment evaluation [42][43].

From the above discussion of the past studies, we have inculcated that there is a dearth of research in the area of mathematical optimization technique for sarcasm detection in Hindi. Also, these studies detect sarcasm but neglect the showcasing of interconnections between sentiment cues and the inherent intonation

Just Accepted

of sentences. Thus, the primary effort of the proposed work is to use a mathematical optimization technique with detection of sentiment cues and intonation presence in sentences for sarcasm detection in Hindi.

## 3 Proposed Methodology

### 3.1 Problem Definition

The goal of this research is to detect sarcasm based on a mathematical optimization technique with the assistance of word embedding in Hindi comments. The first task is to calculate *innate sentence values* and average embedding score using word embedding. Then, determine the implicit intent of the sentence with tonal trait scores based on sentiment analysis and key-phrase extraction. The final task utilizes the probabilistic non-linear optimization problem to perform sarcasm detection.

### 3.2 SPOZ-ESD- Proposed Method

The proposed method aims to differentiate between a sarcastic and non-sarcastic sentence in Hindi taken from social media comments. The following are the steps of the proposed method to detect sarcasm in Hindi:

Step 1: Compute *Innate Sentence Values* (*ISV*) and Average Embedding Score (ω).
Step 2: Determine the Optimal Sentiment Score (OSS) and Highest Triggered Sarcasm Indicator Score ($\rho$)
Step 3: Detect Sarcasm in Hindi Text.

The architectural flow of the proposed method for sarcasm detection in Hindi comments is shown in Fig. 4. Table 1 gives the notations used in the proposed method. Fig. 5 depicts the internal architectural working of the proposed method (SPOZ-ESD) expressed in terms of terminologies. The detailed explanations of steps for the proposed method for sarcasm detection are given in subsequent subsections.

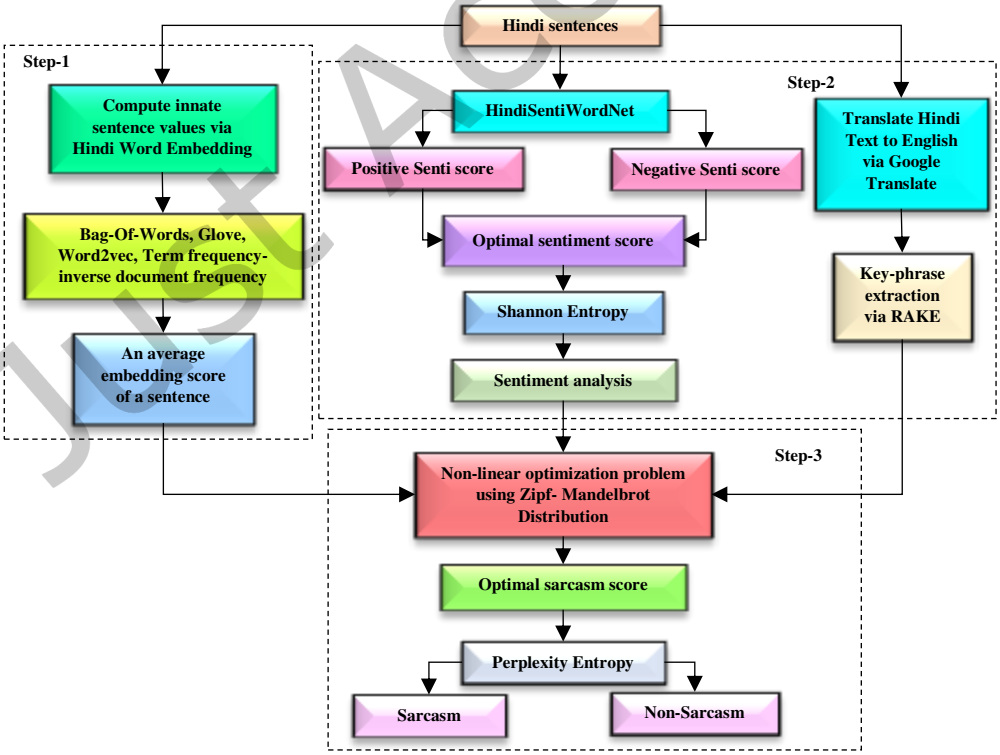



**Fig. 4.** Architectural flow of the proposed method (SPOZ-ESD)

Just Accepted

**Table 1** Notations used in the SPOZ-ESD method

| Notation | Description |
|---|---|
| γ | Calculating innate sentence value using Bag-of-Words |
| μ | Calculating innate sentence value using Tf-Idf |
| $\tau$ | Calculating innate sentence value using word2vec |
| $\theta$ | Calculating innate sentence value using glove |
| $f_j$ | Occurrences of each word $1 \leq j \leq n, n \in N$ |
| $l_j$ | No. of occurrences of each word $1 \leq j \leq n, n \in N$ |
| ω | The average taken of four ISV's obtained from bag-of-words, Tf-Idf, glove and word2vec is termed as average embedding score |
| σ | Sentiment threshold value calculated from Shannon entropy |
| δ | The optimal sarcasm score of the sentence obtained from a formulated probabilistic non-linear optimization problem using Zipf- Mandelbrot distribution |
| ρ | Highest Triggered Sarcasm Indicator Score using key-phrase extraction model RAKE |
| φ | Sarcasm detection threshold value calculated from Perplexity entropy |

Just Accepted

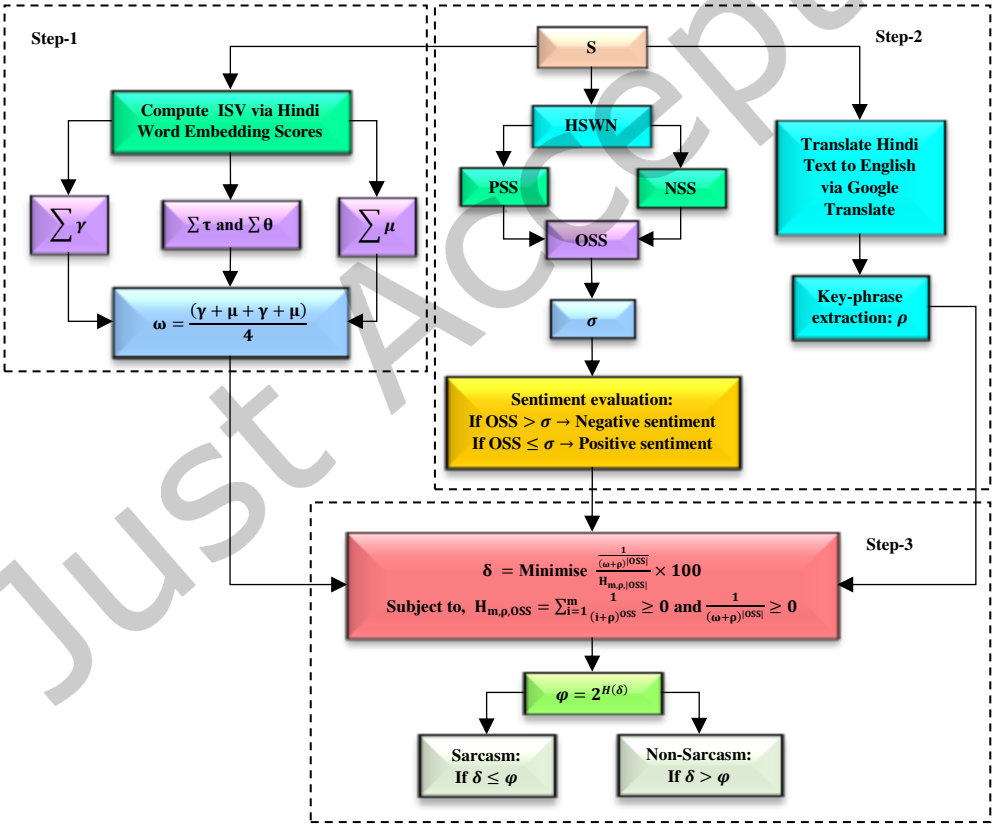



**Fig. 5.** Internal architectural working of the proposed method (SPOZ-ESD)

### *3.2.1 Compute Innate Sentence Values (ISV) and Average Embedding Score (ω)*

The first step is the calculation of the numeric values for the sentence incorporating four-word embedding algorithms viz., Bag-of-Word (BOW), Term Frequency-Inverse Document Frequency (Tf-Idf), word2vec and glove. These values obtained from word embedding provide low-dimensional persistent vectors depicting individual variables with complex features. We generate four values - the sum of BOW ($\gamma$), Tf-Idf ($\mu$),

word2vec ($\tau$) and glove ($\theta$). *Innate sentence value*s (ISV) is the set of four individual sentence scores, such that ISV = $\{\gamma, \mu, \tau, \theta\}$. *Equation (1)- (4)* depicts the calculation of $\gamma$, $\mu$, $\tau$ and $\theta$ respectively.

$$\gamma = \sum_{i=1}^{i=j} \gamma_{w_i}, \text{ total summation value from BOW} \quad (1)$$

$$\mu = \sum_{i=1}^{i=j} \mu_{w_i}, \text{ total summation value from Tf-Idf} \quad (2)$$

$$\tau = \sum_{i=1}^{i=j} \tau_{w_i}, \text{ total summation value from word2vec} \quad (3)$$

$$\theta = \sum_{i=1}^{i=j} \theta_{w_i}, \text{ total summation value from glove} \quad (4)$$

*Algorithms 1, 2, 3* and *4* give the computation of $\gamma$, $\mu$, $\tau$ and $\theta$ using four word embedding algorithms to calculate the *innate sentence values*.

ALGORITHM 1: Calculate the *innate sentence value* using BOW

*Input:* $\boldsymbol{W = (w_1, w_2, w_3, \ldots, w_j)}$ *where* $\boldsymbol{w_j}$ *represents the* $\boldsymbol{j^{th}}$ $\boldsymbol{(1 \leq j \leq n)}$ *word in S.*
*Output: Innate sentence value {*$\boldsymbol{\gamma}$*} using BOW.*

1. *For each unique* $\boldsymbol{w_j}$, *count the no. of each word* $\boldsymbol{(f_j)}$ *in sentence* $\boldsymbol{S_i}$.
2. *Initialize* $\boldsymbol{\gamma = 0}$
3. *for each* $\boldsymbol{w_j}$
   $\boldsymbol{\gamma = \gamma + f_j}$
4. *End*

ALGORITHM 2: Calculate the *innate sentence value* using Tf-Idf

*Input:* $\boldsymbol{W = (w_1, w_2, w_3, \ldots, w_j)}$ *where* $\boldsymbol{w_j}$ *represents the* $\boldsymbol{j^{th}}$ $\boldsymbol{(1 \leq j \leq n)}$ *word in S.*
*Output: Innate sentence value {*$\boldsymbol{\mu}$*} using Tf-Idf.*

1. *For each unique* $\boldsymbol{w_j}$, *count the no. of occurrences of each word* $\boldsymbol{(l_j)}$ *in the sentence* $\boldsymbol{S_i}$.
2. *Initialize* $\boldsymbol{\mu = 0}$
3. *for each* $\boldsymbol{w_j}$
   $\boldsymbol{\mu = \mu + l_j}$
4. *End*

ALGORITHM 3: Calculate the *innate sentence value* using word2vec

*Input:* $\boldsymbol{W = (w_1, w_2, w_3, \ldots, w_j)}$ *where* $\boldsymbol{w_j}$ *represents the* $\boldsymbol{j^{th}}$ $\boldsymbol{(1 \leq j \leq n)}$ *word in S.*
*Output: Innate sentence value {*$\boldsymbol{\tau}$*} using word2vec.*

1. *From genism. models import KeyedVectors*
2. *Calculate sentence score (sentence, word2vec):*
   *embedding = word2vec.embed_sentence(tokens)*
   *sentence_embedding = mean(embedding)*
   *sentence_vector = sum(sentence_embedding)*
3. *For each embedding of the set of words* $\boldsymbol{w_j}$,
   *Sentence embeddings,* $\boldsymbol{u_e = \sum_{j=1}^{j=n}(sentence_vector)}$.
4. *Initialize* $\boldsymbol{\tau = 0}$
   *for each* $\boldsymbol{w_j}$
   $\boldsymbol{\tau = \tau + u_{e_j}}$
5. *End*

Just Accepted

ALGORITHM 4: Calculate the *innate sentence value* using the glove

*Input:* $\boldsymbol{W} = (\boldsymbol{w_1}, \boldsymbol{w_2}, \boldsymbol{w_3}, \ldots, \boldsymbol{w_{j,}})$ *where* $\boldsymbol{w_j}$ *represents the* $\boldsymbol{j^{th}}$ $(\boldsymbol{1 \leq j \leq n})$ *word in S.*
*Output: Innate sentence value* $\{\boldsymbol{\theta}\}$ *using glove.*

1. *Load pre-trained glove embedding vector → glove_model.*
2. *Tokenize and retrieve the vectors for the sentence.*
3. *Sum the vectors to create a sentence vector embedding* $\boldsymbol{v_j}$*:*
   *Sum vector of sentence* $\boldsymbol{F} = \sum_{j=1}^{j=n} \boldsymbol{v_j}$.
4. *Initialize* $\boldsymbol{\theta} = 0$
   *for each* $\boldsymbol{w}$
   $\boldsymbol{\theta} = \boldsymbol{\theta} + \boldsymbol{F_j}$
5. *End*

After calculating the ISV values of a sentence, we further calculate the average embedding score ($\omega$) by taking an average of four ISV values. *Equation (5)* gives the formula to calculate the average embedding score ($\omega$).

$$\omega = \frac{(\gamma+\mu+\tau+\theta)}{4} \tag{5}$$

### *3.2.2 Determine the Optimal Sentiment Score (OSS) and Highest Triggered Sarcasm Indicator Score ($\boldsymbol{\rho}$)*

In the second step, we employ the Hindi SentiWordNet (HSWN) to determine the overall sentiment of Hindi text. The sentiment of the sentence is determined in terms of two polarities, i.e., positive and negative sentiment. The two scores obtained from HSWN are the positive senti-score (PSS) and the negative senti-score (NSS). We compute the optimal sentiment score (OSS) by considering the average absolute difference between PSS and NSS, as shown in *equation (6)*.

$$\text{OSS} = \frac{|PSS-NSS|}{2} \tag{6}$$

The evaluation of sentiment analysis involves the use of sentiment threshold value ($\sigma$). The calculated $\sigma$ is 0.3, obtained using Shannon entropy as shown in Table 2. The decision of sentiment analysis is made based on the comparison between OSS and $\sigma$. If OSS $\leq \sigma$, then the sentence is tagged as "*positive sentiment*" and if OSS $> \sigma$, then the sentence is tagged as "*negative sentiment*". *Algorithm 3* gives the procedure of sentiment analysis based on the utilization of the HSWN model and entropy mechanism.

ALGORITHM 3: Sentiment Analysis Based on HSWN

*Input:* $W$ *– set of words in sentences, HSWN,* sentiment threshold value $(\sigma)$
*Output: OSS, positive sentiment or negative sentiment.*

1. *Let* $W = \{ w_1, w_2, w_3, \ldots, w_h \}$ *where* $w_i$ *represents the* $i^{th}$ $1 \leq i \leq k$ *words in the input sentence.*
2. *For all* $w_i$ in the sentence
   *If* $w_i \in$ *HSWN,*
   $PSS = \frac{positive\ sentiment\ value\ of\ w_i}{k}$ *and*
   $NSS = \frac{neutral\ sentiment\ value\ of\ w_i}{k}$*, where* $k$ *is the total number of words in* $S$
3. *Compute* $OSS = \frac{|PSS-NSS|}{2}$
4. *If* $OSS \leq \sigma \rightarrow$ *"positive sentiment"*
   *Else "negative sentiment"*
5. *End*

After getting the sentiment of the sentence, the Rapid Automatic Keyword Extraction (RAKE) model is utilized to calculate the triggered context score of the sentence that acts as an affective indicator of sarcasm.

Just Accepted

RAKE automatically finds and eliminates irrelevant phrases or terms from the sentence. We first translate the Hindi sentence to English sentence and then apply the RAKE model to get the tuple of words with respective numeral scores. Thus, the set of structures of affected keywords of sentence with scores is procured. The highest keyword value is fetched to further interpolate it in the formulated optimization problem for sarcasm detection. The highest keyword value is termed as *Highest Triggered Sarcasm Indicator Score (ρ).*

#### *3.2.3 Detect Sarcasm in Hindi*

In the final step, sarcasm detection is performed using a probabilistic non-linear optimization problem. The optimization uses Zipf- Mandelbrot distribution as the minimized objective function. The Zipf-Mandelbrot distribution used as an objective function helps to reduce errors by approximating the target function in sarcasm. It gives a more accurate representation of word frequencies, enhancing the robustness of handling the sparsity of adjusting frequency distribution. The *equation (7)* shows the formulated non-linear optimization problem consisting of the Zipf- Mandelbrot distribution.

$$\delta = \text{Minimise } \frac{\frac{1}{(\omega+\rho)^{|OSS|}}}{H_{m,\rho,|OSS|}} \times 100$$

$$\text{Subject to, } H_{m,\rho,OSS} = \sum_{i=1}^{m} \frac{1}{(i+\rho)^{OSS}} \geq 0 \text{ and } \frac{1}{(\omega+\rho)^{|OSS|}} \geq 0 \quad (7)$$

The aim of the formulated non-linear optimization problem is to minimize the objective function with respect to constraints and get the optimal sarcasm score ($\delta$) of the sentence. The optimal sarcasm score is then compared with the computed sarcasm detection threshold value (φ). It is calculated as 0.5 using perplexity entropy structure shown in Table 2. The decision of sarcasm detection is made based on the comparison between $\delta$ and $\varphi$. If $\delta \leq \varphi$, then the sentence is tagged as "*sarcastic*" and if $\delta > \varphi$, then the sentence is tagged as "*non-sarcastic*". *Algorithm 4* defines the procedure of sarcasm detection based on the proposed method SPOZ-ESD.

ALGORITHM 4: Sarcasm Detection using the Proposed Method SPOZ-ESD

*Input:* $\boldsymbol{\omega}$ *(average score of sentence),* $\boldsymbol{\rho}$ *(highest keyword value),* $\boldsymbol{m}$ *(total number of sentences), OSS (optimal sentiment score), sarcasm detection threshold value (*$\boldsymbol{\varphi}$*)*
*Output: sarcasm and non-sarcasm detected based on optimal sarcasm score* ($\boldsymbol{\delta}$)

1. *Compute* $\boldsymbol{\delta}$ *by formulating the non-linear optimization problem:*

   $\boldsymbol{\delta} = \text{Minimise } \frac{\frac{1}{(\omega+\rho)^{|OSS|}}}{H_{m,\rho,|OSS|}} \times 100$

   *Subject to,* $H_{m,\rho,OSS} = \sum_{i=1}^{m} \frac{1}{(i+\rho)^{OSS}} \geq 0$ *and* $\frac{1}{(\omega+\rho)^{|OSS|}} \geq 0$
2. *If* $\boldsymbol{\delta} \leq \boldsymbol{\varphi} \rightarrow$ *"sarcasm"*
   *Else "non-sarcasm"*
3. *End*

### 3.3 Illustrative Examples

The two examples are taken from the two datasets to demonstrate the implementation of the proposed method (SPOZ-ESD) for sarcasm detection.

SENTNCE 1- Hindi text: अन्य लोगों की समस्या सुनने के लिए अच्छा है .. क्योंकि आपको एहसास होगा कि आपकी समस्याएं उनकी तुलना में छोटी हैं

Hindi Transliterated text: Anya logoan kī samasyā sunane ke lie achchhā hai .. Kyoanki āpako ehasās hogā ki āpakī samasyāean unakī tulanā main chhoṭī haian

Hindi Translated text: Is good to hear other people's problem.. Because you will realise that your problems are small compared to them

SENTENCE-2: Hindi text: कर्म एक कुतिया है? अच्छी तरह से बस यह सुनिश्चित करें कि कुतिया सुंदर है ।।

Just Accepted

Hindi Transliterated text: Karma ek kutiyā hai? Achchhī tarah se bas yah sunishchit karean ki kutiyā suandar hai ||

Hindi Translated text: Karma is a bitch? Well just make sure that bitch is beautiful.

Fig. 6 and 7 depict the presence of non-sarcasm and sarcasm in the given contexts.

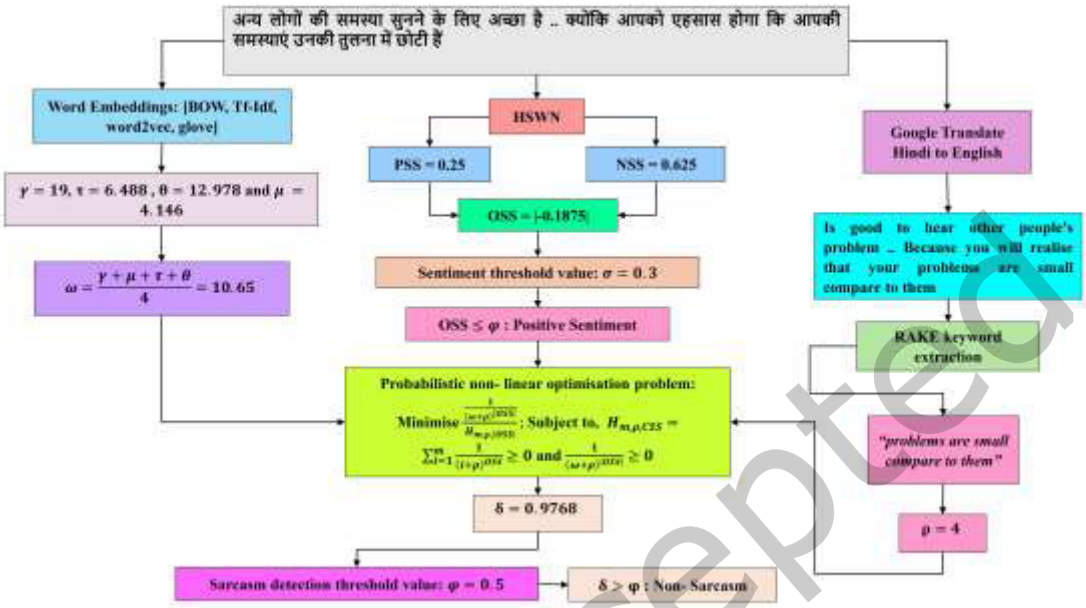



**Fig. 6.** Illustration of proposed SPOZ-ESD method depicting absence of sarcasm

## 4 Experimental and Result Analysis

### 4.1 Hyperparameter Setting

In this section, we give information about thresholds used in SPOZ-ESD. The proposed framework consisted of two essential thresholds i.e., the sentiment threshold value ($\sigma$) and the sarcasm detection threshold value ($\varphi$).

#### *4.1.1 Sentiment Threshold Value ($\boldsymbol{\sigma}$) and Sarcasm Detection Threshold Value ($\boldsymbol{\varphi}$)*

The $\sigma$ is calculated from the absolute Shannon entropy. It measures the typical level of knowledge and unexpectedness related to an arbitrary variable's possible results. *Equation (8)* shows the formulation of the sentiment threshold value ($\sigma$) computed using the Shannon entropy paradigm.

$$\sigma = -\sum_{i=1}^{m} p(OSS_i) \log_2(OSS_i) \tag{8}$$

The $p(OSS)$ is the probability distribution of the OSS of the sentence. The value of $\sigma$ is computed as 0.3.

The perplexity entropy [20] of distribution in the proposed method is utilized to calculate the sarcasm detection threshold value ($\varphi$), as shown in *equation (9).*

$$\varphi = 2^{H(\delta)} \tag{9}$$

$$H(\delta) = \frac{\sum_{i=1}^{m} p(\delta_i) \log_2(\delta_i)}{100} \tag{10}$$

The modified Shannon entropy ($H(\delta)$) is shown in *equation (10)*, computed based on optimal sarcasm score ($\delta$). The $p(\delta)$ is the probability distribution of $\delta$. The computed value of $\varphi$ is 0.5. Table 2 demonstrates the threshold value scales used in the proposed method SPOZ-ESD.

Just Accepted

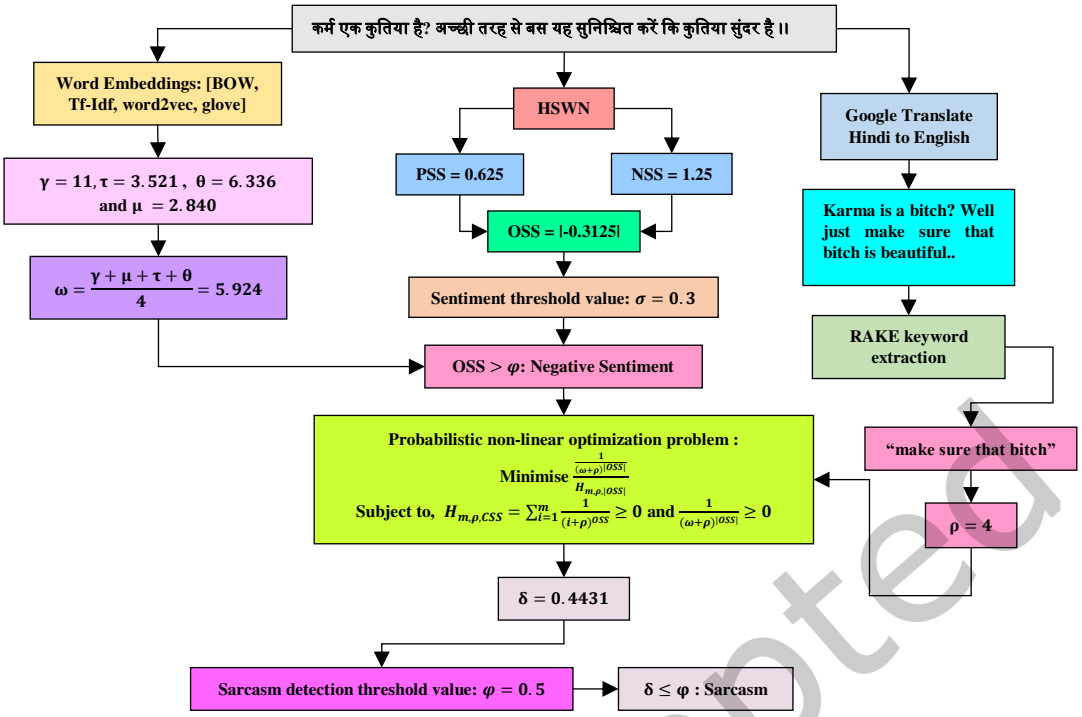



**Fig. 7.** Illustration of the proposed method (SPOZ-ESD) depicting presence of sarcasm

**Table 2** Parameters used in SPOZ-ESD

| Parameter | Values |
|---|---|
| Sentiment threshold value ($\boldsymbol{\sigma}$) | (Dataset-1): $\sigma(\mathrm{OSS}) = -\sum_{i=1}^{6656} p(\mathrm{OSS_i}) \log_2(\mathrm{OSS_i}) \approx 0.3$<br>(Dataset-2): $\sigma(\mathrm{OSS}) = -\sum_{i=1}^{7169} p(\mathrm{OSS_i}) \log_2(\mathrm{OSS_i}) \approx 0.3$ |
| Sarcasm detection threshold value ($\boldsymbol{\varphi}$) | (Dataset-1): $\varphi = 2^{H(\delta)} = 2^{-1.10469} = 0.4657 \approx 0.5$<br>$H_{\text{Dataset 1}}(\delta) = \frac{\sum_{i=1}^{6656} p(\delta_i) \log_2(\delta_i)}{100} = -1.1046$<br>(Dataset-2): $\varphi = 2^{\sigma} = 2^{-1.0688} = 0.4773 \approx 0.5$<br>$H_{\text{Dataset 2}}(\delta) = \frac{\sum_{i=1}^{7169} p(\delta_i) \log_2(\delta_i)}{100} = -1.0668$ |

## 4.2 Dataset Description

The proposed method is tested on two benchmark datasets for sarcasm detection in Hindi. The Dataset-1 and Dataset-2 consisted of News Headlines [44] and comments from Twitter [45] [46] in Hindi respectively. Table 3 gives the statistics of datasets used in the proposed method. Fig. 8 (a) and (b) depicts the word cloud associated with two datasets.

**Table 3** Statistics of datasets used in the proposed method (SPOZ-ESD)

| Datasets | Sarcasm sentences | Non-sarcasm sentences | Total number of sentences |
|---|---|---|---|
| Dataset-1 (news headlines) | 3,554 | 3,102 | 6,656 |
| Dataset-2 (Twitter comments) | 4,407 | 2,762 | 7,169 |

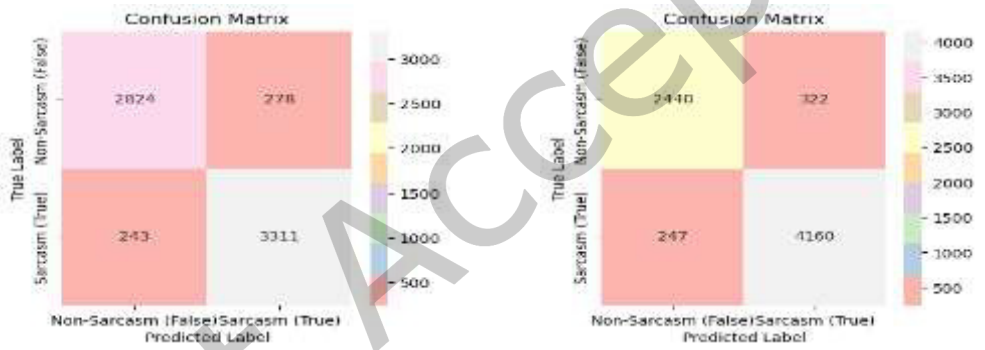
Just Accepted



(a). Word-cloud for Dataset-1 (b). Word-cloud for Dataset-2

**Fig. 8.** Word cloud for the two datasets

Fig. 9 (a) and (b) represent the confusion matrices of two datasets. These confusion matrices comprises of true positives, true negatives, false positives and false negatives. The true positives tell the number of sentences that are predicted by the proposed method as "*sarcasm*" amongst the actual sarcastic sentences. The true negatives tell the number of sentences that are predicted by the proposed method as "*non-sarcasm*" amongst the actual non-sarcastic sentences. The false negatives tell the number of sentences that are predicted by the proposed method as "*sarcasm*" amongst the actual non-sarcastic sentences. The false positives tell the number of sentences that are predicted by the proposed method as "*non-sarcasm*" amongst the actual sarcastic sentences.




(a). Confusion matrix for Dataset-1 (b). Confusion matrix for Dataset-2

**Fig. 9.** Confusion matrices generated for Dataset-1 and 2

Just Accepted

### 4.3 Performance Evaluation

Fig. 10 and 11 shows the snapshots of sentiment analysis using HSWN of sample sentences taken from Dataset-1 and Dataset-2, respectively.

| sentences | senti_score | positive_sentiscore | negative_sentiscore |
|---|---|---|---|
| एक "कर की विरासत और उधार लेने की विरासत दो साल पहले विरासत में मिली अर्थव्यवस्था को अपंग कर दिया।" | -0.500 | 0.250 | 0.75 |
| एक "हमारे अधिकांश छात्र" अंडरपरफॉर्मिंग स्कूलों में फंस गए हैं। | -0.500 | 0.250 | 0.75 |
| डिस्कॉन्सिन के लिए $ 300 मिलियन के प्रोत्साहन आवंटन ने "इस राज्य के हर कोने में हजारों नौकरियों का निर्माण किया है।" | 0.125 | 0.125 | 0.00 |
| 2006 के एक ने प्राइड वोट "भेदभाव से कोई लेना -देना नहीं था।" | 0.875 | 1.125 | 0.25 |
| पेपर बैग पर 5-प्रतिशत का शुल्क एक कर है। | -0.750 | 0.000 | 0.75 |

**Fig. 10.** Sentiment analysis of sample sentences from Dataset-1

| headlines | senti_score | positive_sentiscore | negative_sentiscore |
|---|---|---|---|
| एक क्रांतिकारी युद्ध कॉड लंबे समय से अतिदेय है। | 0.500 | 0.750 | 0.25 |
| जैसे डिफेंस ने हर गेम को अंकों की एक टन-टन छोड़ दिया। | -0.330 | 0.000 | 0.33 |
| लेकिन यह सिर्फ पृथ्वी का प्राकृतिक चक्र है। | 0.295 | 0.625 | 0.33 |
| लोग मैक पर भी खेलते हैं। | -0.330 | 0.000 | 0.33 |
| मुझे लगता है कि अभी भी एक मौका है कि वह जीवित है। | 0.670 | 1.000 | 0.33 |

**Fig. 11.** Sentiment analysis of sample sentences from Dataset-2

The performance of the proposed method is evaluated using the evaluation metrics shown in *equation (11)-(20)*. The evaluation metrics utilized are accuracy, precision, recall, F1 score, specificity, Matthews correlation coefficient (MCC), misclassification rate, false-positive rate, false-negative rate, and false discovery rate.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{11}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{12}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{13}$$

$$\text{F1- score} = \frac{2*Precison*Recall}{Precision+Recall} \tag{14}$$

$$\text{Specificity} = \frac{TN}{TN+FP} \tag{15}$$

$$\text{Matthews correlation coefficient (MCC)} = \frac{(TP\times TN)-(FP\times FN)}{\sqrt{(TP+FP)\times(TP+FN)\times(TN+FP)\times(TN+FN)}} \tag{16}$$

$$\text{Misclassification rate (MR)} = \frac{FP+FN}{TP+FP+TN+FN} \tag{17}$$

$$\text{False positive rate (FPR)} = \frac{FP}{FP+TN} \tag{18}$$

$$\text{False negative rate (FNR)} = \frac{FN}{TP+FN} \tag{19}$$

$$\text{False discovery rate (FDR)} = \frac{FP}{TP+FP} \tag{20}$$

Fig. 12 shows the performance of the proposed method across two datasets. The accuracy scores for Dataset-1 and Dataset-2 achieved by the proposed method are *0.9217* and *0.9206*, respectively. The precision values are obtained as *0.9270* and *0.9281* for the two datasets. The recall metric, which measured the correct identification of relevant responses, achieved scores of *0.9225* and *0.9439* for Dataset-1 and Dataset-2, respectively. The F1-score, which is a balanced combination of precision and recall, is computed as *0.93162* and *0.9359* for Dataset-1 and Dataset-2, respectively. Meanwhile, the MCC is defined as a metric assessing the performance of learning methods in dual and multi-class categorization. It yielded values of *0.8426* and *0.8397* for Dataset-1 and 2, respectively. The false positive rate indicating the ratio of false positive predictions possessed values of *0.0905* and *0.0958* for Dataset-1 and Dataset-2 respectively.

Meanwhile, the false negative reflects the rate of incorrectly identifying false-negative assertions is computed as *0.05711* for Dataset-1 and *0.04962* for Dataset-2. Additionally, the false discovery rate which measures the false positive predictions relative to true positives stands at *0.08709* (Dataset-1) and *0.08746* (Dataset-2). The misclassification rate, also known as the classification error revealed the percentage of inaccurately labelled assumptions calculated as *0.08020* for Dataset- 1 and *0.08567* for Dataset- 2. From the above evaluation, getting higher values in terms of accuracy, precision, recall, F1 score, and MCC showed the robust reliability of the proposed method for sarcasm detection. Meanwhile, the lower scores of false positive,

Just Accepted

false negative, false discovery and misclassification rates showed a minor error percentage in the proposed method for sarcasm detection.

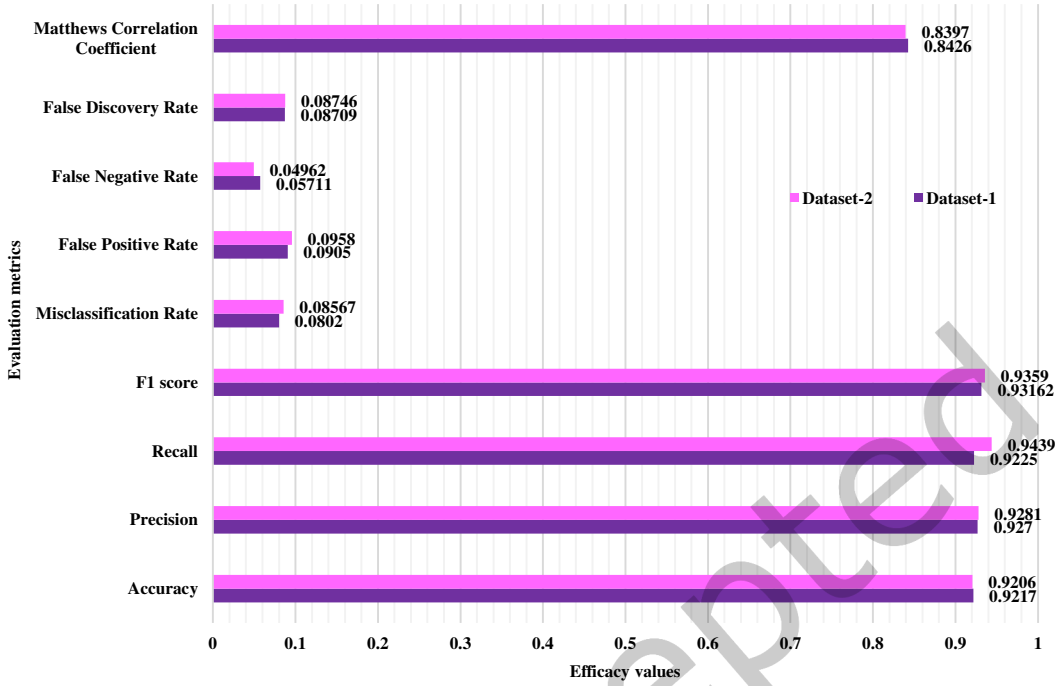



**Fig. 12.** Performance evaluation of SPOZ-ESD over two datasets

Overall, the results indicated that the proposed method (SPOZ-ESD) exhibited improved classification with enhanced sarcasm detection in the given contexts.

### *4.3.1 Macro and Micro Evaluation*

Macro and micro statistics determine how a specific methodology functions with varied databases. The overall macro and micro evaluation of the proposed paradigm on the two datasets is illustrated in Fig. 13.

The cumulative micro and macro precisions of the two datasets are calculated as *0.9253* and *0.9275*, respectively. The micro and macro recall scores across the two datasets are computed as *0.9377* and *0.9332*, respectively. On the two datasets, the macro and micro F1 scores are calculated as *0.9337* and *0.9314*, respectively. Scores close to 1.00 indicated consistent findings. These values demonstrated that the proposed method produced consistent results with confirmed efficacy when applied to varied datasets.

## 4.4 Comparison Analysis of SPOZ-ESD with Existing State-of-The-Art Methods

Table 4 reports the performance evaluation of the proposed method (SPOZ-ESD) in comparison with existing state-of-the-art methods on Dataset-1 and Dataset-2.

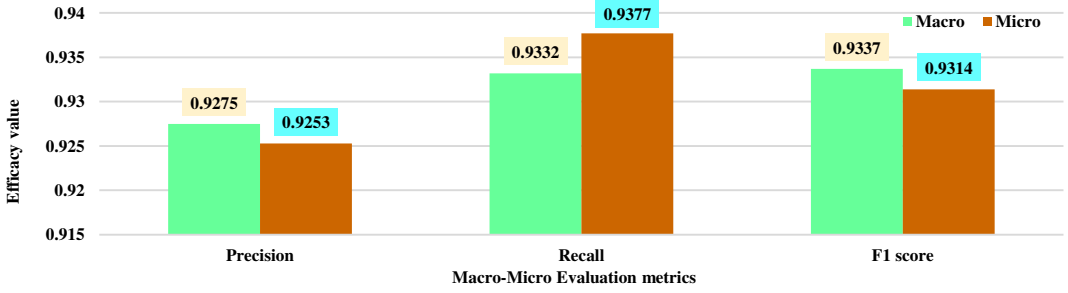



**Fig. 13.** Macro-micro evaluation of SPOZ-ESD on Dataset-1 and Dataset-2

Just Accepted

**Table 4** Comparison analysis of SPOZ-ESD with existing state-of-the-art methods on Dataset-1 and Dataset-2

| | Dataset-1 | | | | Dataset-2 | | | |
|---|---|---|---|---|---|---|---|---|
| Methods | Accuracy | Precision | Recall | F1 score | Accuracy | Precision | Recall | F1 score |
| SVM [7] | 0.7182 | 0.7071 | 0.6882 | 0.6975 | 0.7004 | 0.7189 | 0.6993 | 0.7089 |
| PBCT [35] | 0.6882 | 0.6071 | 0.6937 | 0.6475 | 0.6701 | 0.6258 | 0.6946 | 0.6584 |
| POS-HMM [8] | 0.7624 | 0.7229 | 0.7482 | 0.7353 | 0.7763 | 0.7305 | 0.7580 | 0.7439 |
| TANA [36] | 0.8915 | 0.8834 | 0.8713 | 0.8773 | 0.9016 | 0.8946 | 0.8825 | 0.8885 |
| CBPSD [38] | 0.7539 | 0.7538 | 0.7710 | 0.7620 | 0.7350 | 0.7472 | 0.7583 | 0.7527 |
| CNN-LSTM [37] | 0.8993 | 0.8930 | 0.8847 | 0.8888 | 0.8841 | 0.8859 | 0.8940 | 0.8899 |
| **SPOZ-ESD** | **0.9217** | **0.9270** | **0.9225** | **0.9316** | **0.9206** | **0.9281** | **0.9439** | **0.9359** |

**Dataset-1:** While comparing the performance of SPOZ-ESD with existing state-of-the-art methods on Dataset-1, we observed that the proposed method achieved enhanced results in terms of accuracy, precision, recall and F1 score for sarcasm detection. PBCT [35] and CNN-LSTM [37] methods used embedded parsing with supervised learning techniques but ignored the contextual indicator interpretation of sarcasm. The proposed method (SPOZ-ESD) accomplished a better accuracy rate, surpassing the PBCT [35] and CNN-LSTM [37] methods by 23.35% and 2.24%, respectively. According to the experimental findings, the SPOZ-ESD performed better than the most robust model TANA [36] by 5.12% in terms of recall. Additionally, SPOZ-ESD accomplished a better recall score surpassing the CBPSD [38] by 15.15%. Additionally, over this dataset, the proposed method (SPOZ-ESD) accomplished a better F1 score by 23.41% and 19.63% surpassing SVM [7] and POS-HMM [8] methods respectively. Fig. 14 depicts the performance measure of SPOZ-ESD compared with other existing state-of-the-art methods on Dataset-1.

**Dataset-2:** The SPOZ-ESD exhibited a 25.05% and 3.65% gain in accuracy score in comparison with PBCT [35] and CNN-LSTM [37], respectively. According to the experimental findings the SPOZ-ESD performed better than the SVM [7] method by 20.92% in terms of the precision score. It performed poorly due to the inconsistent specific discrepancies with neglected polarity in sentences. Additionally, over this dataset, the SPOZ-ESD accomplished a better recall rate surpassing POS-HMM [8] and TANA [36] by 18.59% and 6.14%, respectively. The decrement in performance of these methods occurred due to ignorance of utterance-level collaboration of language and knowledge into it. The proposed method performed well in terms of F1 rating, showing an 18.32% increment in its performance compared to CBPSD [38]. Fig. 15 depicts the performance measure of SPOZ-ESD compared with other existing state-of-the-art methods on Dataset-2. The incremental performance of SPOZ-ESD occurred due to the presence of the implicit intent expressed in the text using mathematical optimization technique. This indicated that when contrasted with alternative approaches, i.e., the existing state-of-the-art methods the SPOZ-ESD yielded better results.

The Fig. 16 gives the receiver operating characteristic (ROC) curves for the two datasets. The given curves aid in illustrating the curve between false positive and true positive rates. It represents a visual aid of showing how well the proposed method (SPOZ-ESD) worked based on binary classification performed across various bias limits with good classification results.

Just Accepted

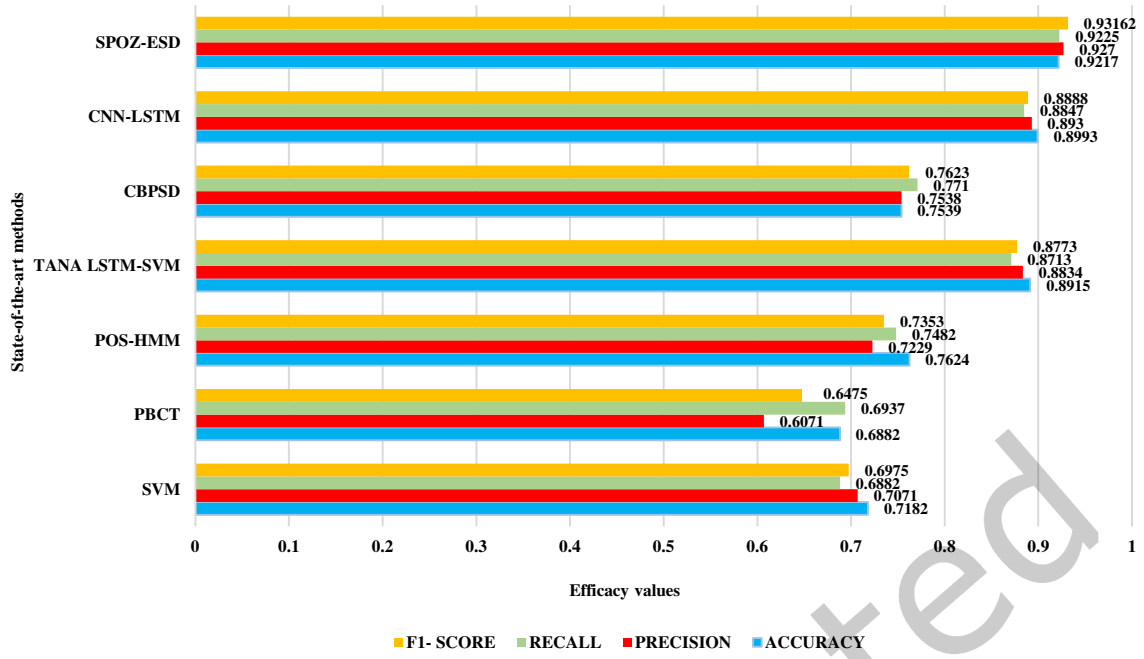



**Fig. 14.** Comparison analysis of SPOZ-ESD with existing state-of-the-art methods on Dataset-1

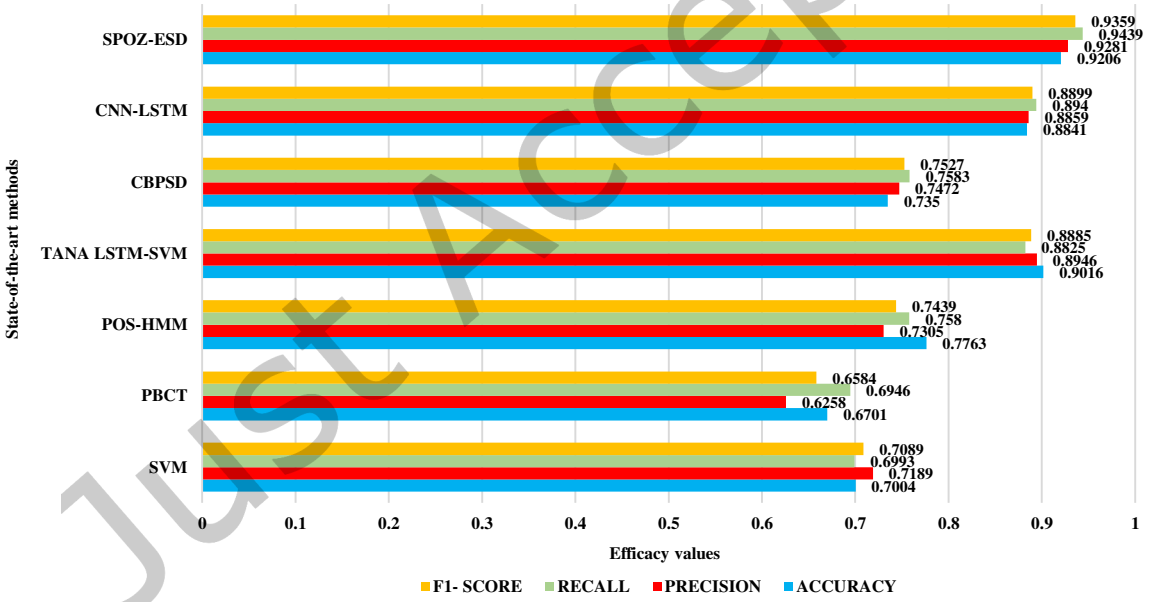
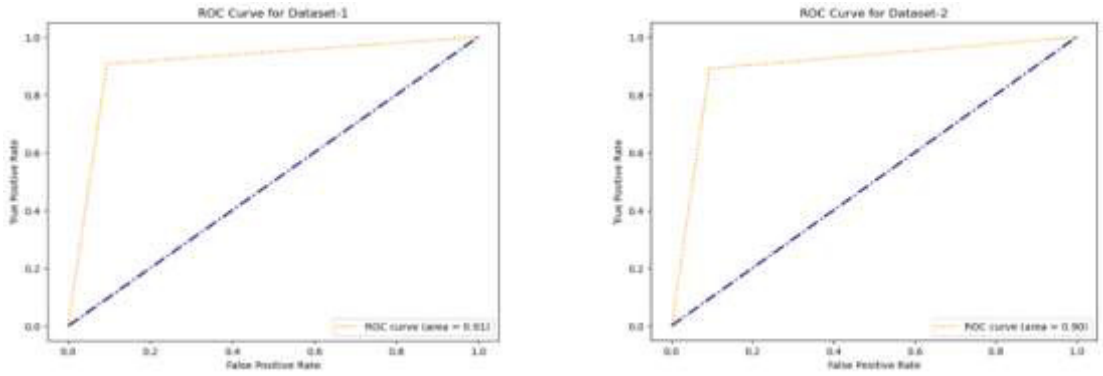



**Fig. 15.** Comparison analysis of SPOZ-ESD existing state-of-the-art methods on Dataset-2

Just Accepted



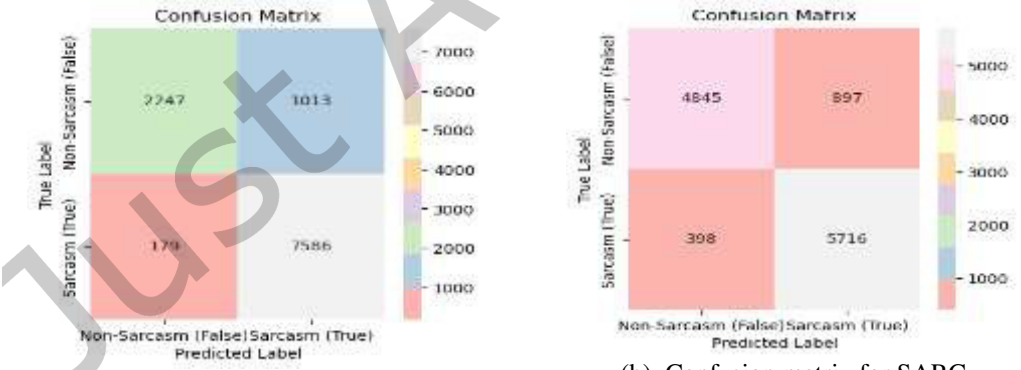

ROC curve of Dataset-1 ROC curve of Dataset-2

**Fig. 16.** ROC curve for two datasets

Additionally, the proposed method (SPOZ-ESD) has been implemented on two English datasets, i.e., Semeval- 2017 [47] and self-annotated Reddit corpus (SARC) [48]. These datasets were utilized to validate the performance of the proposed method for sarcasm detection in English. The performance obtained from these datasets showed the reliability of SPOZ-ESD in English datasets. Table 5 shows the results obtained for SPOZ-ESD performed on English datasets. Fig. 17 shows the confusion matrix generated for the two English datasets.

**Table 5** Results of SPOZ-ESD implemented on English datasets

| Datasets | Accuracy | Precision | Recall | F1 score |
|---|---|---|---|---|
| Semeval-2017 [47] | 0.8907 | 0.8643 | 0.9349 | 0.8982 |
| SARC [48] | 0.8918 | 0.8821 | 0.9769 | 0.9271 |




(a). Confusion matrix for Semeval-2017 (b). Confusion matrix for SARC

**Fig. 17** Confusion matrices generated for Semeval-2017and SARC datasets

## 4.5 Statistical Validation of SPOZ-ESD- Proposed Method Using Four Statistical Tests

### *4.5.1 Statistical Validation of SPOZ-ESD based on z-Test*

We conducted a separate two-proportion z-test using random data taken from the two datasets. In sample 1, a total of 999 random sentences ($n_1$) were considered out of which 801 ($X_1$) sentences were accurately labelled as sarcasm. The computed proportion of correctly labelled sarcastic sentences ($p_1$) obtained as 0.8018. In sample 2, a set of 573 random sentences ($n_2$) has analysed, out of which 462 ($X_2$) sentences were correctly tagged as sarcastic. The corresponding proportion of correctly identified sarcastic sentences ($p_2$) for this sample is computed as 0.8062.

Just Accepted

For two population proportions ($p_1$ and $p_2$), null hypotheses ($H_0$) and alternative hypotheses ($H_a$) are defined as follows:

$$H_0: p_1 = p_2, \text{ i.e., the accuracy score of sample 1} = \text{accuracy score of sample 2}$$

$$H_a: p_1 \neq p_2, \text{ i.e., the accuracy score of sample 1} \neq \text{, the accuracy score of sample 2}$$

In the proposed method the condition of null hypothesis is accepted when the critical range z – value at $\alpha = 0.05$ in *equation (21)* gives the value of the given two sum proportions of two samples is lesser than the computed z score also called as z-statistics in *equation (22).*

$$P = \frac{X_1+X_2}{N_1+N_2} = \frac{801+462}{999+573} = 0.8034 \tag{21}$$

We computed a two-tailed test called z-statistics on two population proportions using *equation (22)* as,

$$z = \frac{|p_1-p_2|}{\sqrt{P(1-P)(\frac{1}{n_1}+\frac{1}{n_2})}} = \frac{|0.8018-0.8062|}{\sqrt{0.8034\times(1-0.8034)\times(\frac{1}{999}+\frac{1}{573})}} = \frac{0.0044}{0.02082} = 0.2113 \tag{22}$$

*Equation (22)* shows that there is less evidence to reject the null hypothesis ($H_0$) which acclaimed that the computed z-statistic as 0.2113 is less than the critical region z-value at $\alpha = 0.05$, which is 0.5040. This shows that available evidence does not substantiate the claim that the population proportion $p_1$ has been distinct from $p_2$ at the significance level of $\alpha = 0.05$. In Fig. 18, the critical region of the hypothesis is depicted graphically. This finding suggested that the consistency of the proposed method's accuracy score holds across diverse sample sizes derived from various datasets. This shows that outcomes of the proposed mathematical optimization technique for sarcasm detection are reliable and consistent.

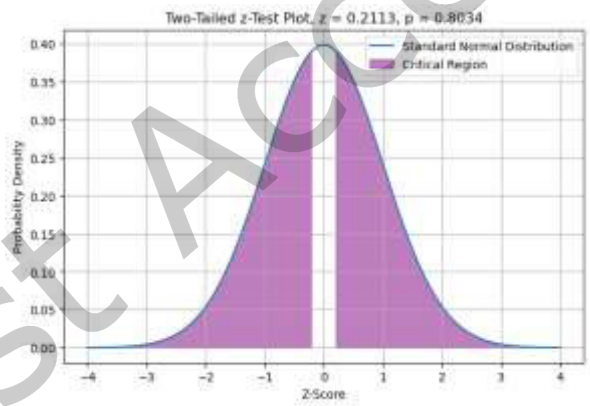



**Fig. 18.** Statistical validation of the SPOZ-ESD using z-test

Just Accepted

*4.5.2 Statistical Validation of SPOZ-ESD based on Kolmogorov-Smirnov (KS) Test*

In the Kolmogorov-Smirnov (KS) test, we compared two independent samples taken from the two datasets. The aim is to determine whether the two samples have similar distributions without assuming any change in the particular distribution of the data. It specifically deals with the similarity between the empirical cumulative distribution function (ECDF) of the given two samples.

In sample 1, we took a total of ($N_1$) 1,133 random sentences out of which ($m_1$) 1,011 sentences were correctly classified as sarcastic. The ECDF value is calculated as ($s_1$) 0.8923. In sample 2, we took a total of ($N_2$) 496 random sentences out of which ($m_2$) 444 were correctly classified as sarcastic sentences. Thus, the computed ECDF value of sample 2 is ($s_2$) 0.8951.

For two population proportions ($s_1$ and $s_2$), null hypotheses ($H_0$) and alternative hypotheses ($H_a$) are defined as follows:

$H_0$: $F_{N_1}(s_1) = F_{N_2}(s_2)$, i.e., accuracy of sample 1 = accuracy of sample 2

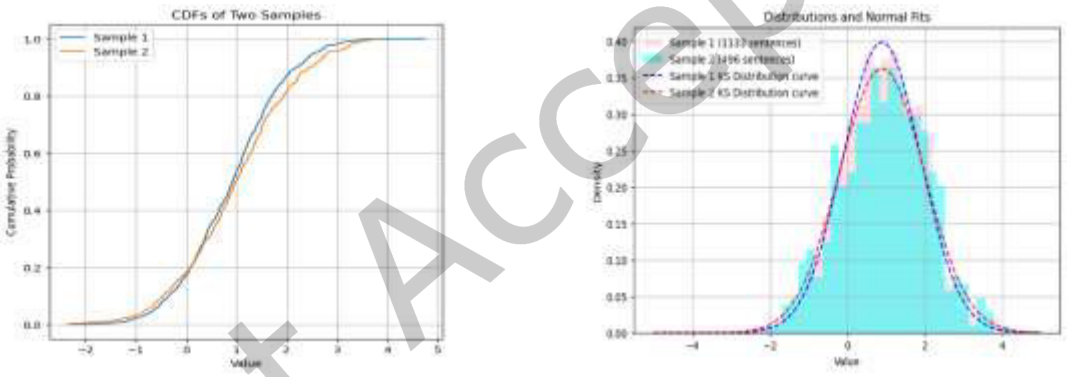
$H_a$: $F_{N_1}(s_1) \neq F_{N_2}(s_2)$, i.e., the accuracy of sample 1 ≠ accuracy of sample 2

According to the Kolmogorov-Smirnov test for two samples, the condition to be satisfied is, $D$-value ($D_{N_1, N_2}$) (called as KS- statistics) must be less than $D_\alpha$, i.e., $D_{N_1, N_2} < D_\alpha$ ($\alpha$ is significance level). If it follows this condition, then the two samples hereby accept the null hypothesis. In the case of the proposed method, if it follows the condition of the KS test then the given null hypothesis is accepted else rejected. We compute the given KS- statistics $D_{N_1, N_2}$, as shown in *equation (23)*.

$$D_{N_1,N_2} = \max\left(|F_{1133}(s_1) - F_{496}(s_2)|\right) = 0.8889 \tag{23}$$

$$D_\alpha = \frac{D(\alpha)\times\left(\sqrt{\frac{1}{N_1}+\frac{1}{N_2}}\right)\times D_{N_1,N_2}\times(1-D_{N_1,N_2})}{10} = \frac{D(0.05)\times\left(\sqrt{\frac{1}{1133}+\frac{1}{496}}\right)\times 0.8889\times(1-0.8889)}{10} = 0.8901 \tag{24}$$

Computing the given values in *equation (24)* demonstrated that the given null hypothesis has been accepted without any failure following the given conditions of the Kolmogorov-Smirnov test. It demonstrated that the computed $D_{N_1, N_2}$ of two samples $\approx 0.8889$ is less than $D_\alpha = 0.8901$, i.e., $D_{N_1, N_2} < D_\alpha$. This shows the given null hypothesis is accepted. Fig. 19 (a) represents the cumulative distribution curve for two samples meeting at the $D_\alpha$ point. Fig. 19 (b) depicts the given $D_\alpha$ value calculated as 0.8901 at the significance level $\alpha = 0.05$ in *equation (24)*, where the two curves' peaks meet. This implied that the two samples have identical accuracy scores in terms of distribution i.e., having consistent distribution for varied databases.




(a). Cumulative distribution frequency curve for two samples

(b). Graphical representation using Kolmogorov Smirnov test

**Fig. 19.** Statistical validation of the SPOZ-ESD using Kolmogorov-Smirnov test

Just Accepted

#### *4.5.3 Statistical Validation of SPOZ-ESD based on Mann Whitney U Test*

The Mann-Whitney U test is used to figure out if two distinct samples originating from different populations have the same distribution or not. On the basis of the provided sample data, we constructed a distribution by employing the mean and standard deviation coming from the number of appropriately categorized sentences in each of the samples.

In sample 1, we took a total ($N_1$) 1556 random sentences out of which ($m_1$) 1222 sentences were correctly classified as sarcastic. The proportion of correctly labelled sentence ($U_1$) is calculated as 0.7853. In sample 2, we took a total of ($N_2$) 847 random sentences out of which ($m_2$) 666 sentences were correctly classified as sarcastic. Similarly, the proportion of correctly labelled sentences ($U_2$) is calculated as 0.7863.

For two population proportions ($U_1$ and $U_2$), null hypotheses ($H_0$) and alternative hypotheses ($H_a$) are defined as follows:

$H_0$: $U_1 = U_2$, i.e., accuracy distribution of sample 1 = accuracy distribution of sample 2

$H_a$: $U_1 \neq U_2$, i.e., accuracy distribution of sample 1 ≠ , accuracy distribution of sample 2

We computed the given $U$- statistics as shown in *equation (25)*. The formulated expected mean value and the standard deviation of the given samples are shown in *equations (26)* and *(27)*, respectively.

$$U = \min(U_1, U_2) = \min(0.7853, 0.7863) = 0.7853 \tag{25}$$

$$\mu_{m_1,m_2} = \left(\frac{U_1 \times U_2}{2}\right) = \left(\frac{0.7853 \times 0.7863}{2}\right) = 0.3087 \tag{26}$$

$$\sigma_{m_1,m_2} = \sqrt{\frac{U_1 \times U_2 \times (U_1 + U_2 + 1)}{12}} = \sqrt{\frac{0.7853 \times 0.7863 \times (0.7853 + 0.7863 + 1)}{12}} = 0.6001 \tag{27}$$

According to this test condition, the $z_B$-value must be less than the given critical $z$-value at significance level $\alpha = 0.05$. If this condition is satisfied then the null hypothesis is accepted based on the given $z_B$-value. *Equation (28)* shows the computation of the $z_B$-value.

$$z_B = \frac{U - \mu_{N_1,N_2}}{\sigma_{N_1,N_2}} = \frac{0.7853 - 0.3087}{0.6001} = 0.7941 \tag{28}$$

The computed $z_B$ value is $0.794$. This represents that the computed $z_B$ value is less than the z-value for the right-tailed as $z_{0.05} = 1.96$. The stated condition of the Mann Whitney U test is satisfied, which shows that there is less evidence to reject the null hypothesis ($H_0$). The available evidence does not substantiate the claim that the population proportion $U_1$ is considered as distinct from $U_2$ at the significance level of $\alpha = 0.05$. Fig. 20 shows the critical region of the hypothesis $z_B$ value depicted graphically. This finding showed the proposed method SPOZ-ESD holds a firm accuracy score across diverse sample sizes derived from various datasets.

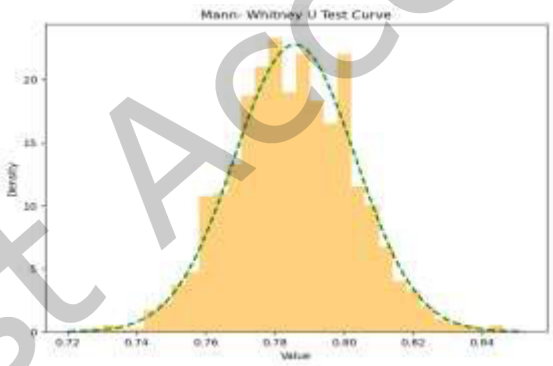



**Fig. 20.** Statistical validation of SPOZ-ESD using Mann- Whitney U test

Just Accepted

#### *4.5.4 Statistical Validation of SPOZ-ESD based on Kruskal- Wallis Test (H- Test)*

A non-parametric statistical procedure called the Kruskal- Walli's test was employed to find if the medians of three or more distinct groups are considered to be distinct in any statistically meaningful way. It is shown with an expansion of the t-test for distinct samples beyond two sets using a non-parametric substitute. We conducted this test using random sentences taken from the two datasets used in the proposed method.

Three samples were taken for testing. In sample 1, a total of 1,442 random sentences ($n_1$) were taken out of which 1,233 ($J_1$) sentences accurately labelled as sarcasm. The proportion rank sum of correctly labelled sentences ($R_1$) for sample 1 is computed as 0.8550. In sample 2, a set of 936 random sentences ($n_2$) were taken with 800 ($J_2$) sentences correctly tagged as sarcastic. The corresponding proportion of the rank sum for sample 2 of correctly identified sarcastic sentences ($R_2$) is computed as 0.8547. In sample 3, we took a total of 588 ($n_3$) random sentences out of which 500 ($J_3$) sentences were correctly classified as sarcastic. The proportion rank sum of correctly labelled sarcastic sentences for sample 3 is computed as ($R_3$) 0.8503.

For three population proportions, null hypotheses ($H_0$) and alternative hypotheses ($H_a$) are defined as follows:

$H_0$: accuracies of sample 1, sample 2 and sample 3 are equivalent

$H_a$: The accuracies of sample 1, sample 2 and sample 3 are not equivalent

According to the Kruskal- Wallis test, if $H_{n_1,n_2,n_3} < \chi^2_\alpha$ (in this case, we have taken significance level $\alpha = 0.05$), then the samples accept the null hypothesis else reject it. In the case of the proposed method (SPOZ-ESD), the null hypothesis was formulated as having a consistent accuracy score drawn from the same distribution and vice-versa for the alternative hypothesis.

We computed the given $H_{n_1,n_2,n_3}$- value as shown in *equation (29)*.

$$H_{n_1,n_2,n_3} = \frac{12}{N\times(N+1)}\left[\sum_{i=1}^{k}\frac{R_i^2}{n_i} - 3\times(N+1)\right] \tag{29}$$

where $N$ denotes the total number of samples taken in observation, $k$ is the number of groups, $R_i$ is considered as the sum of the ranks of the $i^{th}$ group and $n_i$ is the number of observations in the $i^{th}$ group. The total number of samples is computed as $N = n_1 + n_2 + n_3 = 1442 + 936 + 588 = 2966$. Using *equation (29)*, we determined $H_{n_1,n_2,n_3}$ as shown in *equation (30)*.

$$H_{n_1,n_2,n_3} = \frac{12}{N\times(N+1)}\left[\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \frac{R_3^2}{n_3} - 3\times(N+1)\right]$$

$$H_{n_1,n_2,n_3} = \frac{12}{2966\times(2966+1)}\left[\frac{(0.8550)^2}{1442} + \frac{(0.8547)^2}{936} + \frac{(0.8503)^2}{588} - 3\times(3263+1)\right] = 0.8356 \tag{30}$$

The degree of freedom is calculated as $k - 1 = 3 - 1 = 2$. Therefore, at significance level $\alpha = 0.05$ with a degree of freedom as 2, the $\chi^2_\alpha$ value is obtained as 5.99, i.e., $\chi^2_{0.05} = 5.99$. According to the defined test condition, the given null hypothesis has been accepted as $H_{n_1,n_2,n_3} < \chi^2_\alpha$. Fig. 21 depicts the dotted line of the obtained H-value of the three samples taken. This value demonstrates that the three sample values have consistent accuracy scores without showing any change in them.

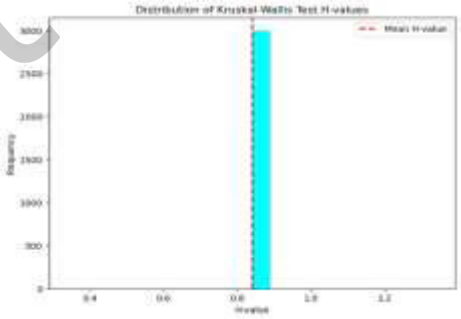


**Fig. 21**. Statistical validation of SPOZ-ESD using the Kruskal Wallis test

## 4.6 Ablation Study

The ablation test has been carried out to see how the change in important parameter affected the performance of the proposed method. The sarcasm detection threshold value (φ) considered as a crucial parameter, which affect the performance of the SPOZ-ESD. The study proceeded by taking changed values of φ as, $\varphi > 0.5$ and $\varphi < 0.5$ to check the change in performance of the SPOZ-ESD. Table 6 shows the results of the ablation study.

Just Accepted

**Table 6** Ablation study of SPOZ-ESD

| Dataset | Removed module | Accuracy | Precision | Recall | F1- score |
|---|---|---|---|---|---|
| Dataset-1 | $\varphi < 0.5$ i.e., $\varphi = \{0, 0.1, 0.2, 0.3, 0.4\}$ | 0.7306 | 0.8977 | 0.5008 | 0.6429 |
| | $\varphi > 0.5$ i.e., $\varphi = \{0.6, 0.7, 0.8, 0.9, 1\}$ | 0.8149 | 0.6891 | 0.9144 | 0.7859 |
| | **SPOZ-ESD with $\varphi = 0.5$** | **0.9217** | **0.9270** | **0.9225** | **0.9316** |
| Dataset-2 | $\varphi < 0.5$ i.e., $\varphi = \{0, 0.1, 0.2, 0.3, 0.4\}$ | 0.6952 | 0.8624 | 0.5256 | 0.6531 |
| | $\varphi > 0.5$ i.e., $\varphi = \{0.6, 0.7, 0.8, 0.9, 1\}$ | 0.8576 | 0.7152 | 0.9319 | 0.8092 |
| | **SPOZ-ESD with $\varphi = 0.5$** | **0.9206** | **0.9281** | **0.9439** | **0.9359** |

It is noteworthy that $\varphi < 0.5$ exhibited lower performance more dominantly on accuracy and recall scores. The reason for low accuracy and recall rates is due to the increased number of false positives in the datasets. False negatives specifically occurred when the proposed method predicted a negative outcome, whereas the actual label is shown as positive. In this case, a number of sentences that were actually labelled as sarcasm were detected as non-sarcasm by the proposed method due to decrement value of $\varphi$ ($< 0.5$). The average value of accuracy, precision, recall and F1 score was calculated corresponding to $\varphi$ in the range $\{0, 0.1, 0.2, 0.3, 0.4\}$ as shown in Table 6. Furthermore, when compared to the SPOZ-ESD outcomes with the correct measurement $\varphi = 0.5$, $\varphi > 0.5$ displayed lower precision and accuracy measures on the two datasets. These lower precision and accuracy rates occurred due to an increased number of false positives in the datasets. False positives specifically occurred when the proposed method predicted a positive outcome where the actual label is considered negative. In this case, a number of sentences that were actually labelled as non-sarcasm were detected as sarcastic by the proposed method due to the increment in the value of $\varphi$ ($> 0.5$) value. Similarly, the average value of accuracy, precision, recall and F1 score was calculated corresponding to $\varphi$ in the range $\{0.6, 0.7, 0.8, 0.9, 1\}$ as shown in Table 6.

The aforementioned analysis showed the computational efficiency of the proposed method has been affected by changing the value of $\varphi$. These values showed a notable escalation in false positive and false negative rates with low evaluation measures. One important finding from this investigation is that using $\varphi = 0.5$ in SPOZ-ESD showed an efficient enhancement in sarcastic declarations with fewer false positive and false negative rates.

### 4.7 Case Study

A case study has been conducted between the proposed method SPOZ-ESD and the existing state-of-the-art method i.e., CNN-LSTM. This case study showed the effectiveness of sarcasm detection which is greatly impacted by the different characteristics of the semantic relatedness component.

The case study showed that the SPOZ-ESD (proposed method) correctly detected sarcasm with the presence of inherent features of implicit expression compared to CNN-LSTM method. It revealed that the proposed method effectively captured the indicators of sentiment cues and implicit intent of tonal traits affecting sarcasm. The proposed method (SPOZ-ESD) specifically uprooted the triggered term with the highest score, which contrasted with the understanding of sarcastic signs. This extraction helped to get the correct sarcastic signal. Additionally, SPOZ-ESD helped learners retain psychological information about sentences in the form of practical details, which is advantageous when dealing with tonal characteristics. Moreover, the proposed method integrated aspects of language to improve the contrastive retention of information-based learning. As a result, it determined the relationship between the statement and the understanding of its surroundings, making it easier to retrieve the right answer. This has been accomplished by combining a basic comprehension of sarcasm indicators. A summary of the case study results is given in Fig. 22 and 23.

### 4.8 Computation Cost of SPOZ-ESD

The time complexity to determine sentiment on getting the optimal sentiment score computed as $O(m)$ where $m$ is the total number of sentences in dataset. The time complexity for determining the optimal sarcasm score is $O(m)$. The computed time complexity of the probabilistic non-linear optimization problem using Zipf-

Just Accepted

Mandelbrot distribution is $O(m)$. Therefore, the final time complexity of the proposed method SPOZ-ESD algorithm for sarcasm detection is approximated as $O(m)$.

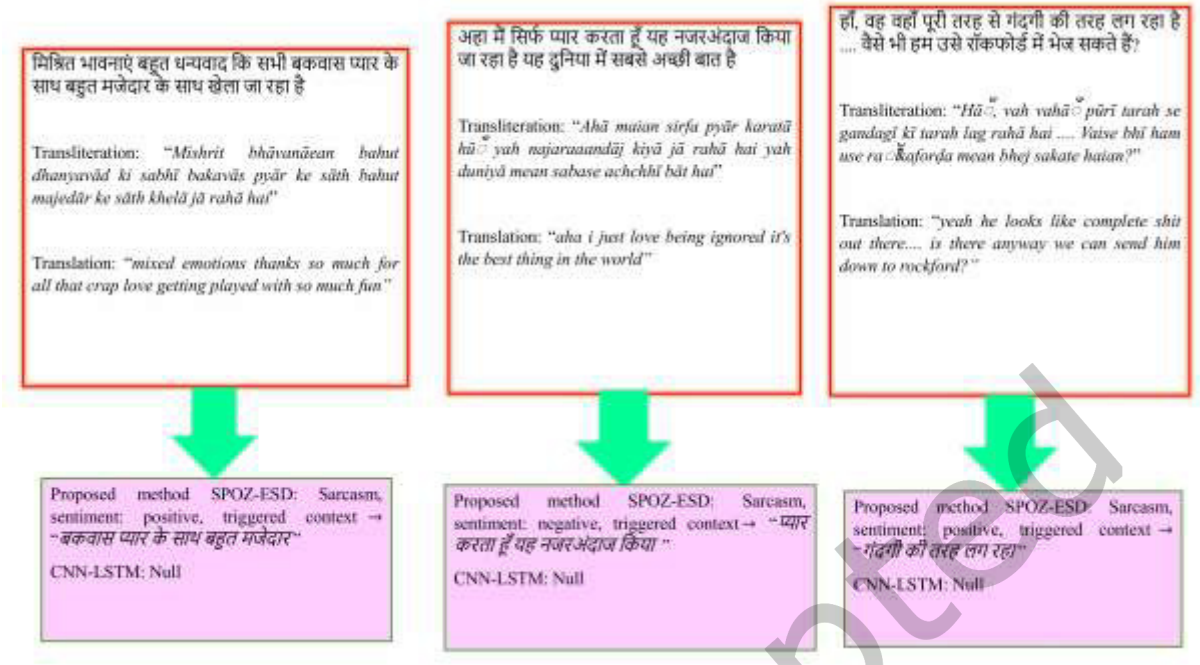



**Fig. 22**. Case study visualisation of sample sentences showing better results by SPOZ-ESD in comparison to CNN-LSTM

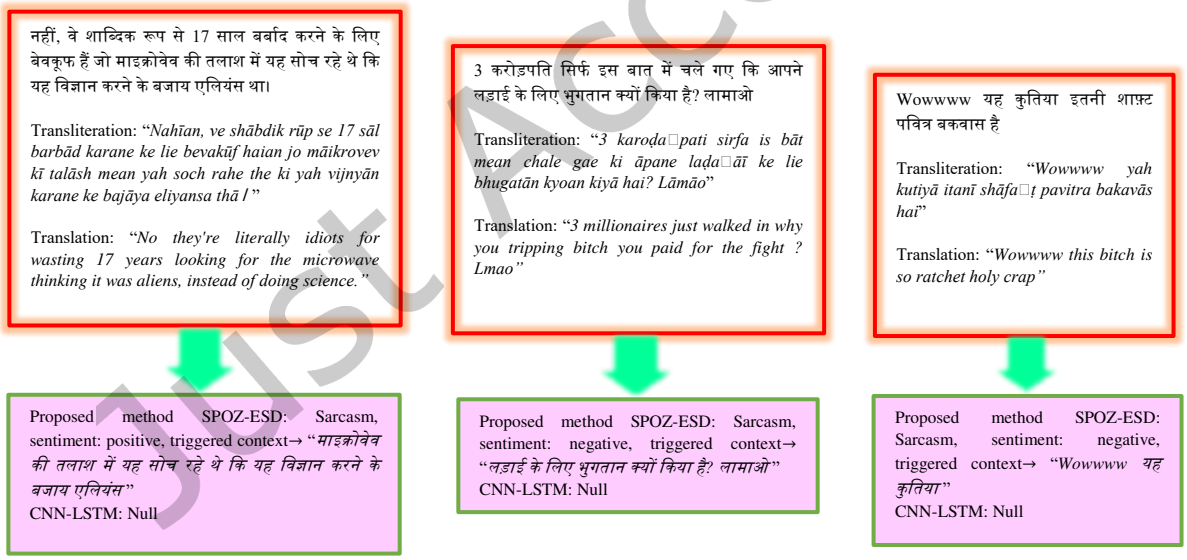



**Fig. 23.** Case study visualization of sample sentences showing better results by SPOZ-ESD in comparison to CNN-LSTM

## 4.9 Challenges of SPOZ-ESD

Despite receiving promising results, the proposed approach is not wary of its share of challenges. A few of the challenges that lead to lower performance are listed below. Fig. 24 shows few sample sentences that the proposed method mislabelled.

1. **Ethnic uniqueness**: The proposed method is unable to correctly classify sentences with significant ethnic and contextual uniqueness. Since, sarcasm is customized to each person, situation, and relationship between the speakers, the lack of such contextual and background knowledge imposed restrictions on the performance of the proposed method.

Just Accepted

2. **Veiled over polite sarcasm:** Sometimes, being respectful could be a pretext for sarcasm and highly complex to identify sarcasm. Giving someone a highly polite compliment is considered as a common indication of sarcasm, which the proposed method did not capture.

## 5 Conclusion

This study proposes a novel mathematical optimization technique-based method for sarcasm detection in social media comments. This proposed method helps to learn the inherent intonations with sentiment cues of sarcastic sentences. Furthermore, it is the first method to the best of our knowledge which addresses the tonal nature of the impacted sarcasm indicator in Hindi using mathematical optimization technique. The proposed method represents a significant advancement in sarcasm detection in Hindi compared with existing state-of-the-art methods. The proposed method performed superiorly with 0.9217, 0.9206 accuracy; 0.9270, 0.9281 precision; 0.9225, 0.9439 recall; and 0.9316, 0.9359 F1 score on Dataset-1 and Dataset-2 respectively. The results of the proposed method are superlative in comparison to existing state-of-the-art methods. The proposed method is reliable, depicting it as statistically validated using four tests in terms of consistent accuracy scores. An ablation study is implemented to evaluate the effectiveness of proposed method for sarcasm detection. It assess the contribution of each component and understand how critical each feature is affecting in overall performance.




**Fig. 24.** Examples of misclassified sentences by SPOZ-ESD from the datasets

Additionally, it has expanded the scope of studies in the domain of mathematical optimization problems for sarcasm recognition and its indicators. In future research, we intend to examine the connections between irony, oxymoron, humblebrag, and sarcastic statements by combining the proposed method SPOZ-ESD.

Datasets: Dataset-1: https://github.com/sbharti1984/Hindi-News

Dataset-2: https://github.com/deepanshu1995/Irony-DetectionHindi-English-Code-Mixed and https://github.com/sahilswami96/SarcasmDetection_CodeMixed

Just Accepted

## References


[1] R. D. Parshad, S. Bhowmick, V. Chand, N. Kumari, and N. Sinha, "What is India speaking? Exploring the 'Hinglish' invasion," *Phys. A Stat. Mech. its Appl.*, vol. 449, pp. 375–389, 2016.

[2] A. Kumar and G. Garg, "The multifaceted concept of context in sentiment analysis," in *Cognitive Informatics and Soft Computing: Proceedings of CISC 2019*, Springer, vol. 1040, pp. 413–421, 2020.

[3] S. Poria, I. Chaturvedi, E. Cambria, and A. Hussain, "Convolutional MKL based multimodal emotion recognition and sentiment analysis," in *Proceedings of the IEEE Sixteenth International Conference on Data Mining (ICDM)*, pp. 439–448, 2016.

[4] N. Majumder, S. Poria, H. Peng, N. Chhaya, E. Cambria, and A. Gelbukh, "Sentiment and sarcasm classification with multitask learning," *IEEE Intell. Syst.*, vol. 34, no. 3, pp. 38–43, 2019.

[5] A. Kumar and V. H. C. Albuquerque, "Sentiment analysis using XLM-R transformer and zero-shot transfer learning on resource-poor indian language," *Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 20, no. 5, pp. 1–13, 2021.

[6] D. Ghosh, W. Guo, and S. Muresan, "Sarcastic or not: Word embeddings to predict the literal or sarcastic meaning of words," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 1003–1012, 2015.

[7] N. Desai and A. D. Dave, "Sarcasm detection in hindi sentences using support vector machine," *Int. J.*, vol. 4, no. 7, pp. 8–15, 2016.

[8] S. K. Bharti, K. Sathya Babu, and S. K. Jena, "Harnessing online news for sarcasm detection in hindi tweets," in *Proceedings of the International Conference on Pattern Recognition and Machine Intelligence*, pp. 679–686, 2017.

[9] Potamias, Rolandos Alexandros, Georgios Siolas, and Andreas-Georgios Stafylopatis. "A transformer-based approach to irony and sarcasm detection." *Neural Computing and Applications* vol. 32, no. 23 pp. 17309-17320, 2020.

[10] M. S. Akhtar, A. Ekbal, and E. Cambria, "How intense are you? Predicting intensities of emotions and sentiments using stacked ensemble [application notes]," *IEEE Comput. Intell. Mag.*, vol. 15, no. 1, pp. 64–75, 2020.

[11] A. Kumar and G. Garg, "Empirical study of shallow and deep learning models for sarcasm detection using context in benchmark datasets," *J. Ambient Intell. Humaniz. Comput.*, vol. 14, no. 5, pp. 5327–5342, 2023.

[12] D. Davidov, O. Tsur, and A. Rappoport, "Semi-supervised recognition of sarcasm in Twitter and Amazon," in *Proceedings of the Fourteenth Conference on Computational Natural Language Learning*, pp. 107–116, 2010.

[13] A. Kumar, S. Dikshit, and V. H. C. Albuquerque, "Explainable artificial intelligence for sarcasm detection in dialogues," *Wirel. Commun. Mob. Comput.*, vol. 2021, pp. 1–13, 2021.

[14] E. Riloff, A. Qadir, P. Surve, L. De Silva, N. Gilbert, and R. Huang, "Sarcasm as contrast between a positive sentiment and negative situation," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 704–714, 2013.

[15] T. Ptáček, I. Habernal, and J. Hong, "Sarcasm detection on czech and english twitter," in *Proceedings of the Twenty-Fifth International Conference on Computational Linguistics Technical papers*, pp. 213–223, 2014.

[16] Y. Y. Tan, C.-O. Chow, J. Kanesan, J. H. Chuah, and Y. Lim, "Sentiment Analysis and Sarcasm Detection using Deep Multi-Task Learning," *Wirel. Pers. Commun.*, vol. 129, no. 3, pp. 2213–2237, 2023.

[17] N. Kalchbrenner, E. Grefenstette, and P. Blunsom, "A convolutional neural network for modelling sentences," in *Proceedings of the Fifty-Second Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 655-665, 2014.

[18] T. Shen, T. Zhou, G. Long, J. Jiang, S. Pan, and C. Zhang, "Disan: Directional self-attention network for rnn/cnn-free language understanding," in *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, pp. 1-10, 2018.

[19] Y. Zhang *et al.*, "A Multitask learning model for multimodal sarcasm, sentiment and emotion recognition in conversations," *Inf. Fusion*, vol. 93, pp. 282-301, 2023.

[20] D. K. Sharma, B. Singh, S. Agarwal, N. Pachauri, A. A. Alhussan, and H. A. Abdallah, "Sarcasm Detection over Social Media Platforms Using Hybrid Ensemble Model with Fuzzy Logic," *Electronics*, vol. 12, no. 4, pp. 937-957, 2023.

[21] Vitman, Oxana, Yevhen Kostiuk, Grigori Sidorov, and Alexander Gelbukh. "Sarcasm detection framework using context, emotion and sentiment features. " *Expert Systems with Applications*, vol. 234, pp. 121068, 2023.

[22] Y. Zhang, Y. Yu, M. Wang, M. Huang, and M. S. Hossain, "Self-Adaptive Representation Learning Model for Multi-Modal Sentiment and Sarcasm Joint Analysis," *ACM Trans. Multimed. Comput. Commun. Appl.*, vol. 20, no. 5, pp. 1–17, 2024.

[23] N. Punetha and G. Jain, "Integrated shannon entropy and COPRAS optimal model-based recommendation framework," *Evol. Intell.*, vol. 17, no. 1, pp. 385–397, 2024.

[24] A.-C. Băroiu and Ștefan Trăușan-Matu, "Comparison of Deep Learning Models for Automatic Detection of Sarcasm Context on the MUStARD Dataset," *Electronics*, vol. 12, no. 3, pp. 666-670, 2023.

[25] A. K. Mishra, M. S. Choudhry, and M. Kumar, "Underwater image enhancement using multiscale decomposition and gamma correction," *Multimed. Tools Appl.*, vol. 82, no. 10, pp. 15715–15733, 2023.

[26] L. Wu *et al.*, "Graph neural networks for natural language processing: A survey," *Found Trends in Mach. Learn.*, vol. 16, no. 2, pp. 119–328, 2023.


Just Accepted

[27] M. A. Galal, A. H. Yousef, H. H. Zayed, and W. Medhat, "Arabic sarcasm detection: An enhanced fine-tuned language model approach," *Ain Shams Eng. J.*, vol. 15, no. 6, pp. 102736–102749, 2024.

[28] M. E. Hassan, M. Hussain, I. Maab, U. Habib, M. A. Khan, and A. Masood, "Detection of Sarcasm in Urdu Tweets using Deep Learning and Transformer based Hybrid Approaches," *IEEE Access*, vol. 12, pp. 61542-61555, 2024.

[29] A. M. Alayba, V. Palade, M. England, and R. Iqbal, "A combined CNN and LSTM model for Arabic sentiment analysis," in *Proceedings of the International Cross-Domain Conference on Machine Learning and Knowledge Extraction*, pp. 179–191, 2018.

[30] P. Liu, W. Chen, G. Ou, T. Wang, D. Yang, and K. Lei, "Sarcasm detection in social media based on imbalanced classification," in *Proceedings of the Fifteenth International Conference on Web-Age Information Management*, pp. 459–471, 2014.

[31] E. Lunando and A. Purwarianti, "Indonesian social media sentiment analysis with sarcasm detection," in *Proceedings of the International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, pp. 195–198, 2013.

[32] R. Justo Blanco, J. M. Alcaide Salinas, M. I. Torres Barañano, and M. Walker, "Detection of Sarcasm and Nastiness: New Resources for Spanish Language," *Cognt. Comput.*, vol. 10, pp. 1135-1151, 2018.

[33] S. K. Bharti and K. S. Babu, "Sarcasm as a contradiction between a tweet and its temporal facts: a pattern-based approach," *Int. J. Nat. Lang. Comput.*, vol. 7, no. 5, pp. 1-13, 2018.

[34] P. Katyayan and N. Joshi, "Performance evaluation of machine learning algorithms for detecting Hindi sarcasm," *Bharatiya Vaigyanik evam Audyogik Anusandhan Patrika*, vol. 29, no. 1, pp. 43–48, 2022.

[35] S. K. Bharti, K. S. Babu, and S. K. Jena, "Parsing-based sarcasm sentiment recognition in twitter data," in *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pp. 1373–1380, 2015.

[36] D. K. Jain, A. Kumar, and S. R. Sangwan, "TANA: The amalgam neural architecture for sarcasm detection in indian indigenous language combining LSTM and SVM with word-emoji embeddings," *Pattern Recognit. Lett.*, vol. 160, pp. 11–18, 2022.

[37] A. Kumar, S. R. Sangwan, A. K. Singh, and G. Wadhwa, "Hybrid deep learning model for sarcasm detection in Indian indigenous language using word-emoji embeddings," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 22, no. 5, pp. 1–20, 2023.

[38] S. K. Bharti, K. S. Babu, and R. Raman, "Context-based sarcasm detection in hindi tweets," in *Proceedings of the Ninth International Conference on Advances in Pattern Recognition (ICAPR)*, pp. 1–6, 2017.

[39] K. Z. Bousmaha, K. Hamadouche, H. Djouabi, and L. Hadrich-Belguith, "Automatic Algerian Sarcasm Detection from Texts and Images," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 23, no. 108, pp. 1-25, 2024.

[40] M. Marreddy, S. R. Oota, L. S. Vakada, V. C. Chinni, and R. Mamidi, "Am I a resource-poor language? Data sets, embeddings, models and analysis for four different NLP tasks in telugu language," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 22, no. 1, pp. 1–34, 2022.

[41] U. Ahmed, J. C.-W. Lin, and G. Srivastava, "Emotional Intelligence Attention Unsupervised Learning Using Lexicon Analysis for Irony-based Advertising," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 23, no. 1, pp. 1–19, 2024.

[42] X. Tang, M. Zheng, J. Feng, J. Huang, and Y. Gong, "Shortcut Enhanced Syntactic and Semantic Dual-channel Network for Aspect-based Sentiment Analysis," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 22, no. 11, pp. 1–20, 2023.

[43] A. Perti, A. Sinha, and A. Vidyarthi, "Cognitive hybrid deep learning-based multi-modal sentiment analysis for online product reviews," *ACM Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 23, no. 8, pp. 1–14, 2024.

[44] S. Swami, A. Khandelwal, V. Singh, S. S. Akhtar, and M. Shrivastava, "A corpus of english-hindi code-mixed tweets for sarcasm detection," *arXiv preprint*, arXiv1805.11869, pp. 1-9, 2018.

[45] D. Vijay, A. Bohra, V. Singh, S. S. Akhtar, and M. Shrivastava, "A Dataset for Detecting Irony in Hindi-English Code-Mixed Social Media Text.," *EMSASW@ ESWC*, vol. 2111, pp. 38–46, 2018.

[46] D. S. Kulkarni and S. S. Rodd, "Sentiment analysis in Hindi—A survey on the state-of-the-art techniques," *Trans. Asian Low-Resource Lang. Inf. Process.*, vol. 21, no. 1, pp. 1–46, 2021.

[47] R. K. Gupta and Y. Yang, "Crystalnest at semeval-2017 task 4: Using sarcasm detection for enhancing sentiment classification and quantification," in *Proceedings of the Eleventh International Workshop on Semantic Evaluation (SemEval-2017)*, pp. 626–633, 2017.

[48] N. Pawar and S. Bhingarkar, "Machine learning based sarcasm detection on Twitter data," in *Proceedings of the Fifth International Conference on Communication and Electronics Systems (ICCES)*, pp. 957–961, 2020.

Just Accepted

scientific reports

Check for updates

OPEN

# Sustainable material as a column filler in soft clay bed reinforced with encased column: numerical analysis

Srijan[1,2]✉ & A. K. Gupta[1]

**Soil reinforcement is one of the techniques used to enhance the engineer characteristics of the soil. Various techniques can be employed to stabilise problematic soils, such as soft clay. These include the utilisation of portland cement, lime, fly ash, ground freezing, jet grouting, prefabricated vertical drains, and thermal approaches. Similarly, stone columns are one of the most popular methods for enhancing soil and have been adopted across the world to enhance bearing capacity and minimise total and differential settlements of structures built on soft clay. Additionally stone columns serving as vertical drains, thus accelerate consolidation process. But with higher demand of construction material and depletion of natural resources, there is a need of using waste products as a substitute to existing construction materials. In this paper, manufacturing waste like steel slag is used as a sustainable material for column infill and can be affordable and also address the current environmental concern which by removing solid waste... A series of numerical investigation was carried out to study the various behavioural characteristics of virgin soft clay bed, clay bed installed with ordinary steel slag column and also with encased steel slag column. A comparison was made among all for studying the various parameters such as settlement, excess pore water pressure, stress concentration ratio & lateral deformation of columns.**

**Keywords** Numerical Analysis, Soft clay bed, Soil reinforcement, Stone column

Construction of embankments over soft deposits may encounter issues including excesive settlement, significant lateral deformation, and regional or global instability. Compaction piles, displacement type, and replacement type, as well as vacuum pre-consolidation, the utilisation of pre-fabricated vertical drains, and soil reinforcing, are among the several strategies that can be employed to enhance the strength of cohesive soils[1]. The utilisation of granular piles, also known as stone columns, for soil stabilisation in to address various constructions, including embankments and buildings, is seeing an increasing trend[2–12]. Granular columns cause soft soil to consolidate more quickly because they have a shorter drainage channel, higher load bearing capacity, and lower settlement[13–17]. However, because of low lateral strength, building granular columns having undrained strength ($S_u$) below 15 kPa in soft grounds is nearly difficult[18]. To eliminate this limitation, one possible solution is to encase the granular column with a geosynthetic material. This additional layer of confinement enhances the shear resistance of the column, hence preventing excessive bulging[16,19–23].

However, several pressing problems have arisen as a result of the demographic increase and the expansion of urban infrastructure, including the large global constructing materials requirement that is impacted by the depletion of ecological materials, and the enormous stockpiles of waste which are building up in landfills. Because it has become more complex and expensive to get virgin resources, recycling and reusing garbage is a popular choice among governments and other top decision-makers[24]. Consequently, several initiatives have been taken to utilise industrial wastes as substitute construction materials. The majority of earlier research centred on using industrial wastes like fly ash, blast furnace slag, red mud, and others in various civil projects[25–30]. One of the most promising waste products, steel slag, has been the subject of several studies[31–33]. Estimates put the steel slag output at 10–15% of that of basic metal steel (US Geological Survey 2020). Data from the World Steel Association[34]show that only 11.5 million tonnes of the 15.7 million tonnes of steel slag produced as a by-product within Europeare recycled[35]. The use of steel slag again accounts for around 30% of all steel slag in China, the world's largest steel

[1] Department of Civil Engineering, Delhi Technological University, New Delhi 110089, India. [2] Department of Civil Engineering, Maharishi Markandeshwar Engineering College, Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala 133207, India. ✉email: Ssrijanpurvey@gmail.com

producer. In other words, China produces 80 million tonnes of steel slag year annually[36]. About 60% of the complete steel slag output is utilised as a substitute material in the building sector in industrialised nations, compared to 20% in developing countries[37,38]. As a soil enhancement material, Poh et al.[39] combined steel slag fines with two distinct activators, $Na_2SiO_3$ (sodium metasilicate) and CaO (quicklime). The findings indicated that using steel slag fines increased strength and durability while lowering the soil's propensity for expansion. Shen et al.[40] claimed that using steel slag and other solid waste products combination for a base material of road produced specimens with an even better strength for a long time compared to specimens fixed with cement and a higher initial strength than lime-fly ash or lime-soil mixes. Akinwumi et al.[41] researched the effects of putting pulverised steel slag over a lateritic soil's geotechnical characteristics. The findings demonstrated that increasing the amount of steel slag in soil samples from five to ten% reduced the soil's flexibility and swelling potential while improving its permeability and cured strength. Wu et al.[42] developed a superior slag-based composite with better cementation efficacy to alter the expansive soil characteristics. The test findings demonstrated that cementing the slag significantly decreased the possibility of swelling and enhanced the strength properties of the soil. Lang et al.[43] examined the efficacy of adding steel slag to stabilise dredging sludge. The findings suggested that a steel slag percentage of between 5 and 10% was ideal for enhancing the shear strength of dredged sludge stabilised with cement. In their study, Gulsen et al.[44] examined the potential use of steel slags, specifically those from basic oxygen furnace (BOF) and electric arc furnace (EAF), as backfill material in coastal constructions. The study found that the density, porosity, water absorption, and Los Angeles abrasion ratios of steel slags were consistently greater than those of natural aggregates. It is recommended that EAF slags undergo a minimum ageing period of 6 months, whereas BOF slags require at least 24 months of ageing in open air conditions before they may be utilised as backfill material in coastal constructions. Additionally, the maximum particle size should be lowered to 25 mm. Viorica et al.[45] conducted a comprehensive analysis of the chemical, structural, and morphological characteristics of three distinct slag samples that were examined for their potential use in agricultural recycling. The results demonstrate that the combination of landfill slag waste and granulated metallurgical slag can effectively be utilised to restore and preserve damaged areas. Despite all the existing studies, the utilise steel slag as an infill material for column reinforcing the soft soil has not been studied in detail. This paper shows a numerical analysis to understand the working behaviour of ordinary steel slag column (OSSC) reinforcing the soft soil. Additionally, research has been done to understand how encased steel slag column (ESSC) behaves. A comparison was made to determine the response of virgin soil, OSSC and ESSC on various parameters like settlement, excess pore water pressure (Excess PWP), stress concentration ratio (SCR) and lateral deformation of columns. All the said analyses were done for the total construction period of the embankment i.e., 60 days.

## Numerical analysis

### Geometry

In order to simplify the actual geometry of the problem, which typically requires a moderately large number of columns and to solve this problem, a basic geometrical model was used as shown in Fig. 1. Cross-section and plan views of the embankment slice lying on ESSC reinforced soft clay are displayed in Fig. 2, along with numerous monitoring points that were utilised to analyse the simulation findings. The problem is symmetrical, so one half of the embankment cross-section is modelled (in this case the right half is chosen). In the model, a soft clay bed was made on a plane having a breadth 48 m, a thickness of 2.5 m, & a depth of 10 m. A 5 m embankment is created in three layers on top of soil bed having a thickness 2.5 m. The height of the first layer and second layer is 2 m each and height of the third layer is 1 m. The width of the embankment crest is 8 m and width of the whole embankment up to the toe of embankment from origin is 18 m. The slope of embankment has an inclination of 1 V:2 H. Stone columns are provided below the embankment for the entire embankment's depth in clay bed i.e., 10 m deeper. Stone column's diameter is considered as 0.75 m (radius = 0.375 m) and a

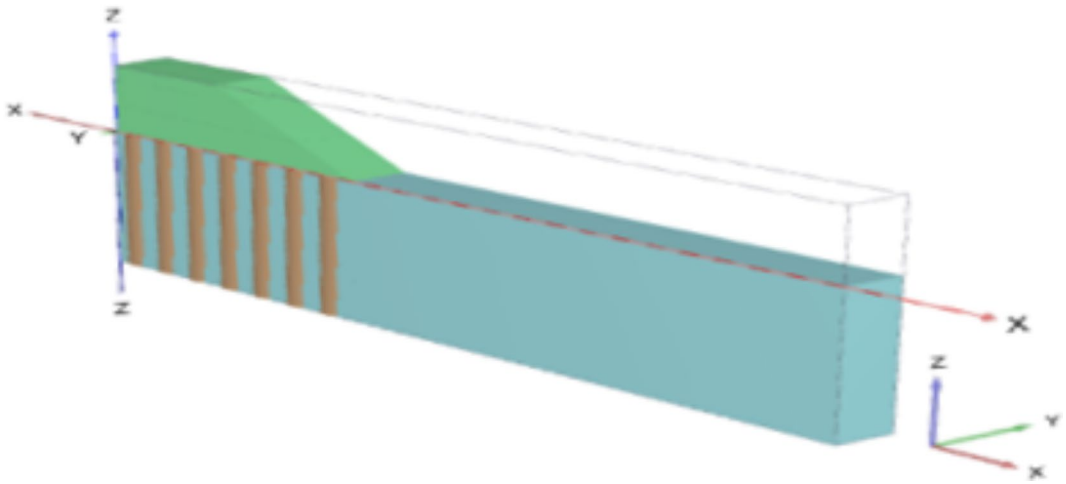



**Fig. 1**. 3D cut representation of the embankment supported by encased stone columns and sitting on soft clay.

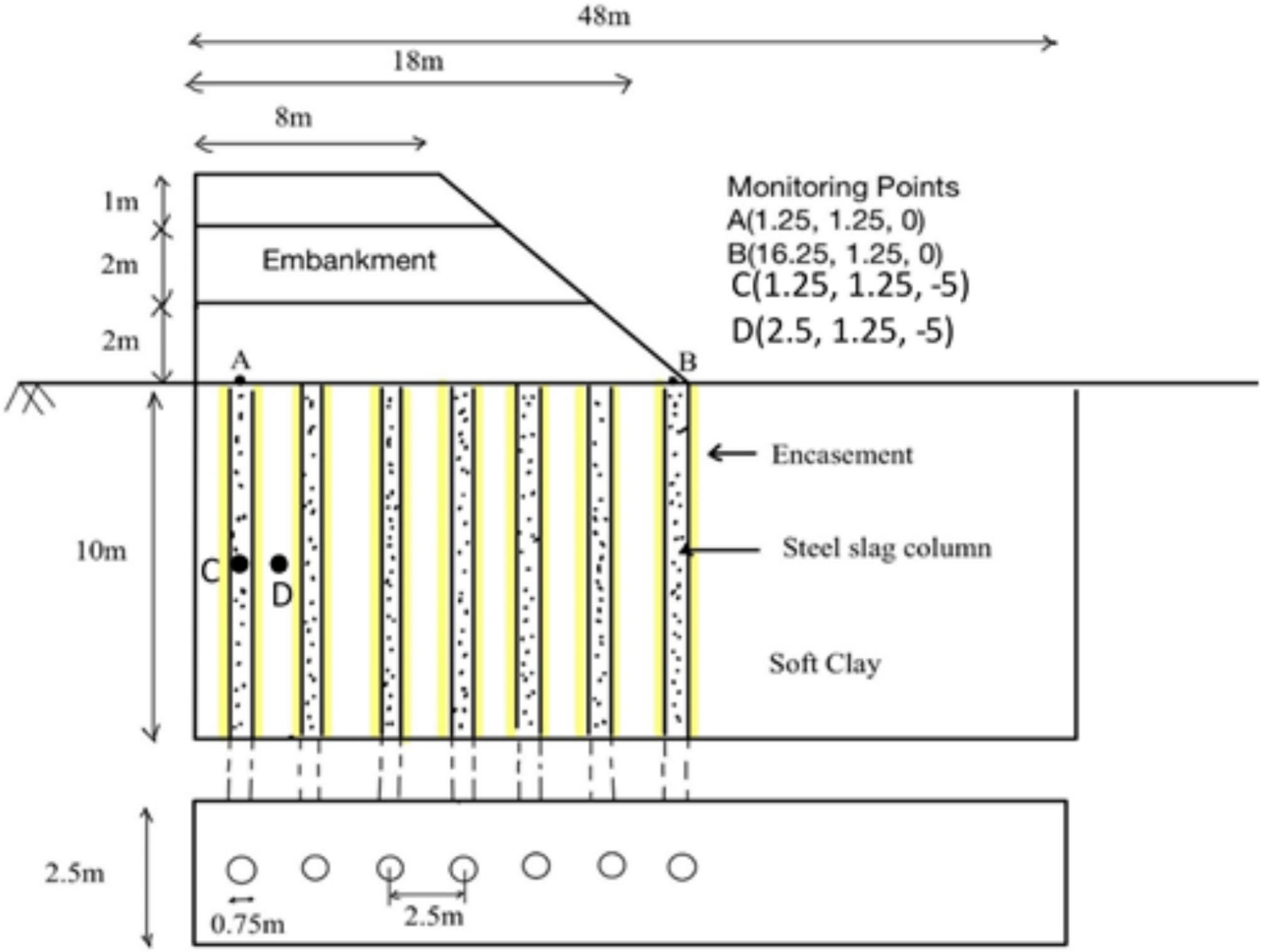



**Fig. 2**. Details displaying cross section & plan of the embankment lying on soft clay improves with ESSC.

2.5 m centre-to-centre distance. The first column is at a distance of 1.25 m from the origin and last column is at distance of 16.25 m from origin.

## Material properties and meshing

The constitutive behaviour of the granular columns & embankment soil is simulated in the current investigation using an elastic perfectly plastic Mohr-Coulomb model. A soft soil model was used to model the foundation clay. Similar analysis was adopted by various previous researchers[8,23]. The water level is considered to be at surface level throughout the investigation, which is typical for very soft clay soils. The material model and related parameters utilised in the investigation are shown in Table 1. Using a linearly elastic model, the geotextile employed for the encasement was modelled as a thin element that could only withstand tensile force. Ten nodded tetrahedral components were used to represent the soft clay, the embankment, & the sand platform, while six nodded triangular surface elements were used to model the geotextile. Twelve nodded interface components of insignificant thickness with a 0.8 strength reduction factor were used to mimic the soil-geosynthetic interaction.

For meshing the whole geometry, a coarse element distribution type was chosen having a global scale factor of 1.20 and a minimum element size factor of $5\times10^{-3}$. For virgin soil total number of generated nodes, stress points and elements were 814, 1400 and 350 respectively. Similarly, for the soft clay bed reinforced with OSSC the total number of nodes, stress points and elements were 23,719, 65,992 and 16,498 respectively. Also, when the same soft clay bed was reinforced with ESSC the total number of nodes, stress points and elements were 32,693, 66,244 and 16,561 respectively. The final meshed diagram for each of the three types i.e., virgin soil, OSSC, and ESSC are represented in Fig. 3.

## Model sequence

Model is constructed for three cases (Virgin Soil, OSSC, ESSC). In all cases firstly a soft clay bed is created then embankment is created on the top of clay layer using a three-stage construction sequence to achieve 5 m of height. First two stage was comprised of two substages, one was used for the construction of the embankment (each 2 m height) and another as consolidation stage. The third stage was constructed for a height of 1 m. Initially, first two layers of embankment is constructed for 10 days each and was allowed to be consolidated for

| Properties | Symbol | Unit | Soft Clay | Embankment | Steel Slag | Geogrid |
|---|---|---|---|---|---|---|
| Material Model | - | - | Soft Soil | Mohr-coulomb | Mohr-coulomb | - |
| Height | H | M | 10 | 5 | 10 | 10 |
| Constrained Modulus | E | kPa | 1800 | 18,000 | 275,000 | - |
| Effective Cohesion | C' | kPa | 5 | 5 | 99 | - |
| Unit Weight | γ | $kN/m^3$ | 17.16 | 20 | 25.90 | - |
| Poisson's Ratio | ν' | - | 0.34 | 0.30 | 0.30 | - |
| Effective Friction Angle | Φ' | ° | 12 | 28 | 56 | - |
| Dilation Angle | Ψ' | ° | - | - | - | - |
| Swelling Index | $C_s$ | - | 0.01 | - | - | - |
| Compressibility Index | $C_c$ | - | 0.12 | - | - | - |
| Permeability in all directions | $K_x$ / $K_y$ / $K_z$ | m/day | $9.87 \times 10^{-3}$ | 0.88 | 12.96 | - |
| Axial Stiffness | EA | kN/m | - | - | - | 2000 |

**Table 1**. Material properties used in the numerical analyses.

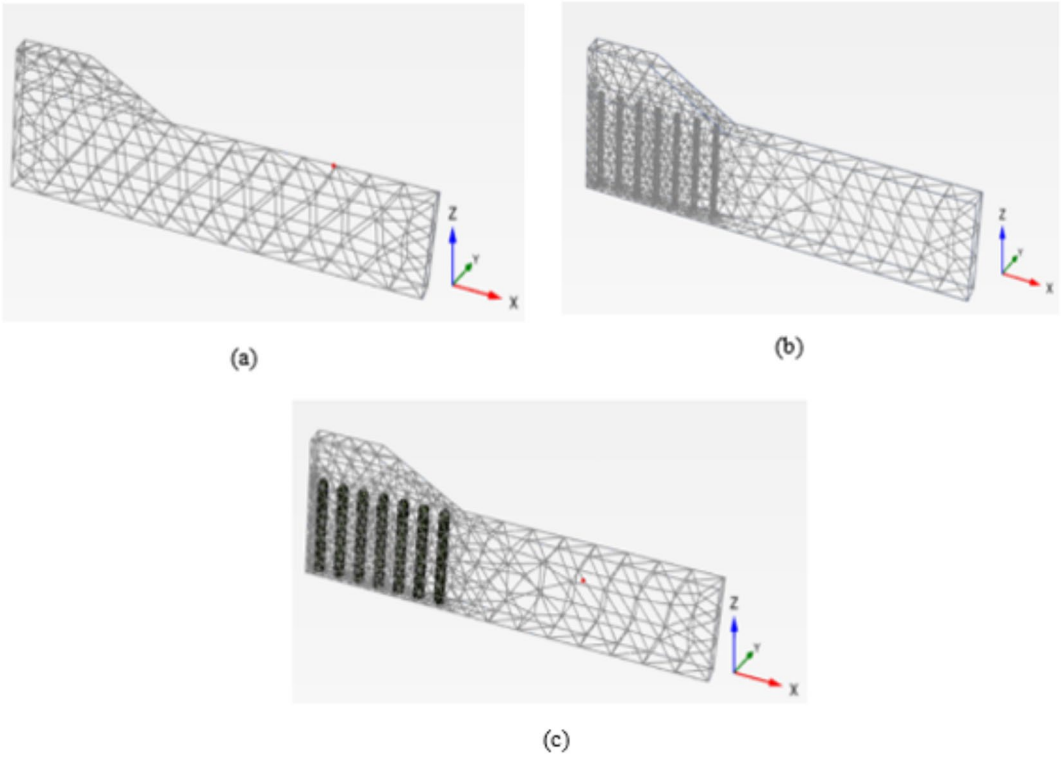



**Fig. 3**. Generated Mesh for (a) Virgin Soil without any reinforcement (b) OSSC (c) ESSC.

15 days each. Finally, third layer of the embankment is constructed and time interval for the construction of third layer is 10 days.

In the virgin soil model after the construction of the embankment, points are selected for the curve for various analyses. In ordinary steel slag stone column (OSSC) model after construction of the embankment, stone columns are provided below the embankment for the whole depth of the embankment in clay bed and then points are selected for the curve. In encased steel slag stone column (ESSC) model, after construction of the embankment, stone columns are provided below the embankment for the whole depth in the clay bed & then geogrid is used to surround stone columns for the complete column's length, afterwards points are selected for the curve.

To reduce the boundary impact, the boundary's distance from the toe of the embankment was determined to be 48 m using sensitivity analysis. The displacement and hydraulic boundaries are chosen in accordance with symmetry. The model is given a full-fixity condition at its base and zero displacements in the direction orthogonal to the symmetric plane. A no-flow boundary state was set up on the plane of symmetry and only vertical flow was allowed through the model's base at the model's core.

## Numerical results and discussions

### Displacement analysis

To ascertain the displacement behaviour in relation to time, the head of the first (A point in Fig. 2) & last column (B point in Fig. 2) were selected for the analysis of all three case of Encased steel slag column (ESSC), Ordinary steel slag column (OSSC), soil with no reinforcement (Virgin Soil) and the corresponding nodal points were chosen. The mean resulting displacement on the column's tops were calculated, permitting the time-displacement graph up to the construction period (60 days) to be drawn accordingly, which is given in Fig. 4.

From Fig. 4(a), it is found that, by using OSSC, at the head of the first column, displacement decreases by 1.48 times than that of virgin soil, and by using ESSC, displacement decreased by 1.23 times than that of OSSC and by 2.43 times than that of virgin soil. From Fig. 4(b), at the head of the last column, it is found by using OSSC, displacement decreases by 1.85 times than that of virgin soil and by using ESSC, displacement decreased by 1.59 times than that of OSSC and by 2.35 times than that of virgin soil.

By reducing the embankment load transmitted to the clay layer and improving the stiffness of the stone column, our findings show how the geosynthetic encasement may significantly minimise the settling of the soft ground. Various other researchers also saw the same enhancement (Hosseinpour et al. 2015; Murugesan and Rajagopal 2010).

### Stress investigation

To ascertain the stress behaviour in relation to time, the centre of the first column at the mid-depth (point C in Fig. 2) was selected and the investigation was done to study the response of excess pore water pressure (Excess PWP). The resultant excess PWP on the said point for virgin soil, OSSC and ESSC due to the embankment loading (in the last stage of embankment construction that is, 60 days) is represented in Fig. 5.

Figure 5 shows the temporal variation of Excess PWP response, and it is found that by including the OSSC, the generated excess PWP was reduced by 1.15 times than unreinforced case and by further encasing the OSSC i.e., ESSC the excess PWP further reduces by 1.08 times as compared to OSSC and by 1.25 times than that of unreinforced case. Additionally, the addition of OSSC and ESSC accelerated the Excess Pore Water Pressure dissipation rate.

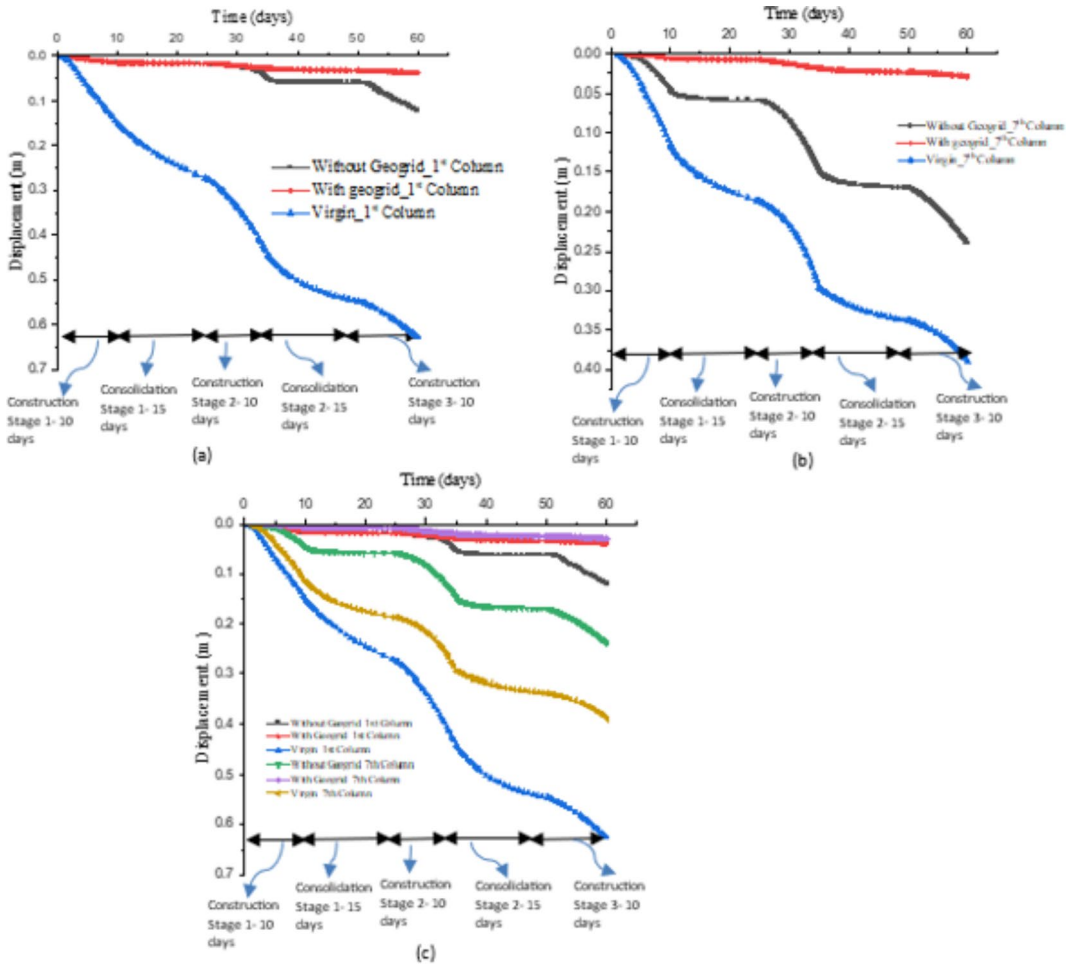



**Fig. 4**. Time-Displacement graph at the head of (a) 1st column (b) 7th column (c)1st and 7th column.

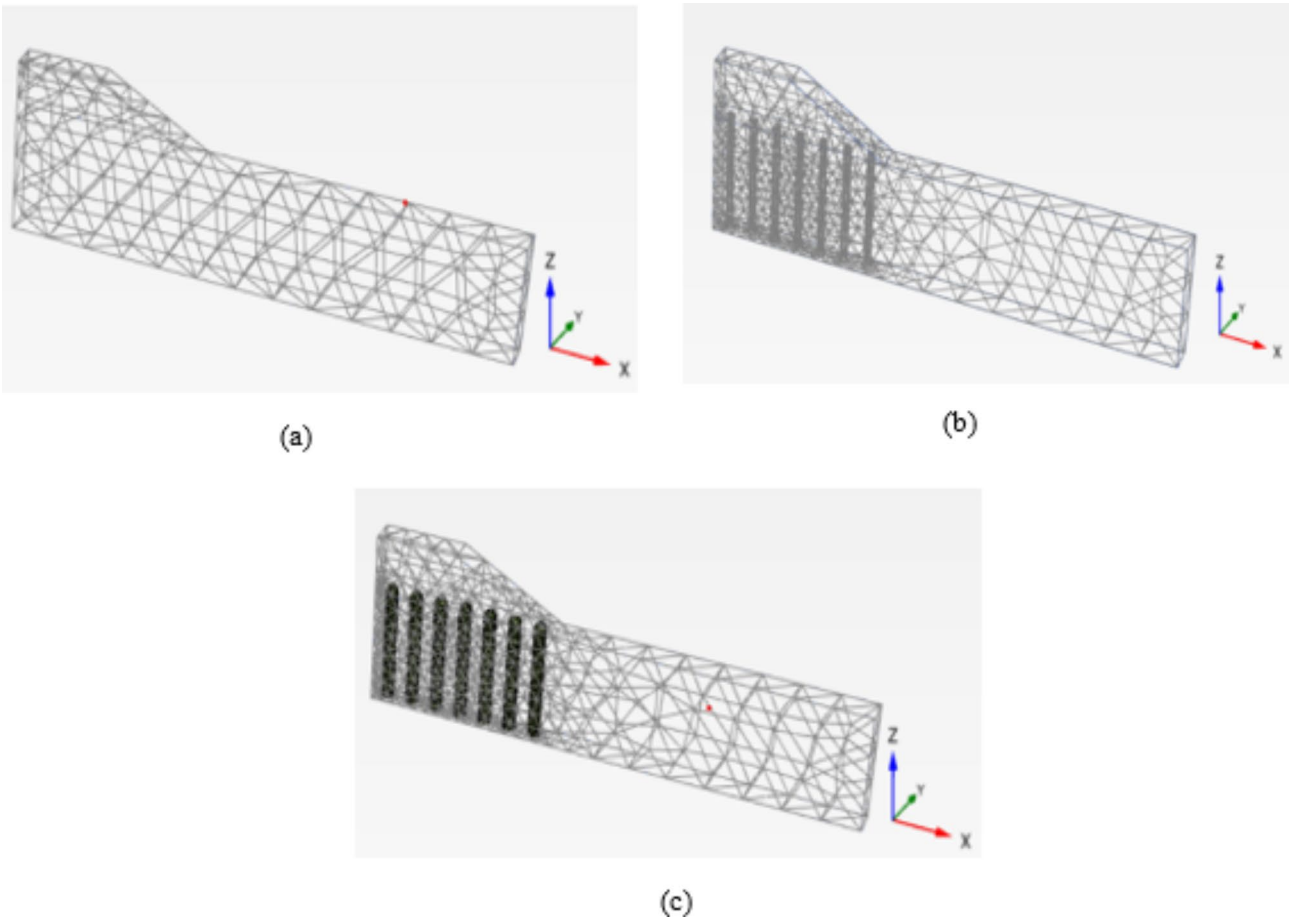



**Fig. 5**. Generation of Excess Pore Water Pressure in final stage of building phase for virgin soil, OSSC and ESSC.

At the conclusion of the embankment building stage, Fig. 6 shows the distribution of developed Excess PWP owing to embankment loading for various situations of unreinforced, OSSC and ESSC reinforced soils. As anticipated, there is non-uniformity in the distribution of produced surplus pore pressure at the completion of the building stage. In each case, the highest pore pressure was found below the embankment crown. Additionally, there is greater pore pressure creation in the area between the first two columns under the crown. Beyond the embankment's toe, the region of clayey soils does not suffer much of an increase in pore pressure production.

## Stress concentration ratio

An index called Stress Concentration Ratio (SCR) is a measure of transmission of embankment loading through stone columns. The SCR is the relation of stress in foundation soil to stress in the stone column at an identical location. The more rigid stone columns bear greater weight compared to the fragile soil in a ground supported by stone columns.

To ascertain the impact of OSSC & ESSC on stress concentration behaviour with respect to time, a point at a mid depth of the first column and another point just adjacent to it in the clay bed at an identical location (point C & point D respectively in Fig. 2) were chosen, time-stress concentration curves were drawn accordingly and is represented in Fig. 7. In Fig. 7, the Stress Concentration Ratio for both ESSC and OSSC vary over time. The SCR is not uniform i.e., changes over time for each of the situations taken into consideration, as shown. For the first 10 days or so after building an embankment, the SCR for these circumstances stays constant (very close to one), but with time, it drops because columns share more stresses than the nearby soft clay. Intriguingly, a rise in OSSC can be seen at the conclusion of the embankment building period, indicating that significantly larger pressures were transmitted to the nearby soil. But when columns are covered in geosynthetic material, SCR drops far less than OSSC. It is found that by using OSSC stress is reduced by 1.15 times than that of virgin Soil and by using ESSC stress is reduced by 1.08 times than that of OSSC and by 1.25 times than that of virgin soil.

The collapse or plastic points for clayey soil strengthened with ESSC of various configurations are shown in Fig. 8 at the conclusion of the building stage. Plastic points can be seen in the embankment structure, at the column junction, and throughout the construction period. Plastic points are more prevalent around the column and in the embankment's body, but they are comparatively less prevalent for columns that are close to the toe.

Figure 9 illustrates how the geogrid encasement affects the principal effective stress produced on the column material and the nearby soft clay beneath the embankment. The findings of the principal effective stress at the

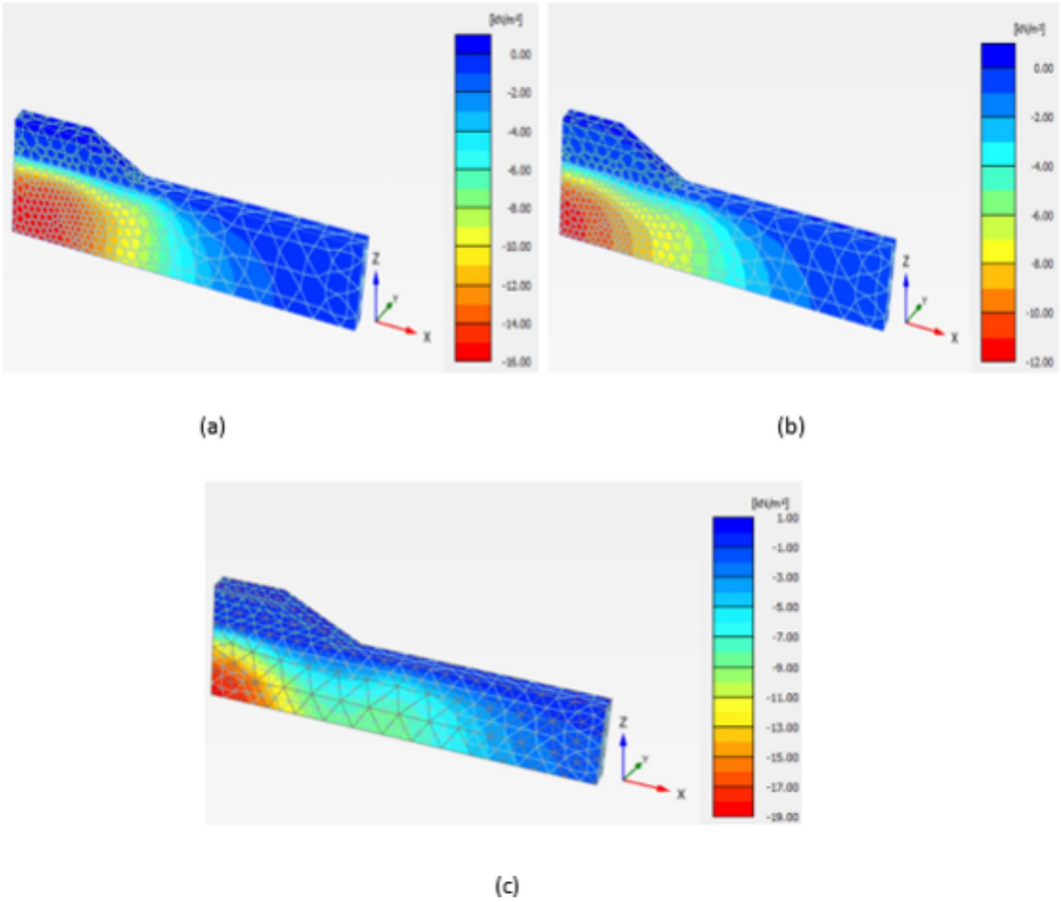



**Fig. 6**. Flow of the developed excess pore pressure at final stage of embankment building (60 days) for (a) virgin soil (b) OSSC (c) ESSC.

head location of each column at the conclusion of the construction stage are displayed. The principal effective stress of unreinforced soft clay increases at the conclusion of the consolidation stage due to an extra rise in geostatic stress caused by embankment construction, however this increase differs based on the geometry of the embankment. It is most intense underneath the embankment's crest and gradually lessens on the sloped face of the embankment. When OSSC are introduced, column material is subjected to greater effective stress than the nearby soft clay. And, on further addition of encasement to the OSSC (i.e., ESSC) results in higher generation of the effective stress. Principal effective stress mobilised in soft clay with the addition of ESSC was substantially lower for equivalent locations than OSSC and for soft clay without any reinforcement. As it is evident from Fig. 9, the first three columns which are below the crest of the embankment have higher principal stresses as compared to other columns.

## Lateral displacement of columns

At the conclusion of the consolidation period, Fig. 10 depicts the patterns of column deformation under embankment stress. Columns are resting on soft clay strengthened by OSSC and ESSC. It is seen that columns beneath the embankment's crest bulge, but throughout the sloping area, column bending appears to predominate. It can be observed from Fig. 10, that the lateral deformation of all the columns keeps on increasing as we move away from the crest portion to the sloped region. This variation was observed in both OSSCs and ESSCs. To comprehend the impact of OSSC and ESSC given the variation in deformation patterns, columns 2 & 7's lateral deformation has been examined which is being shown in Fig. 11. From Fig. 11, it is found that by using ESSC the maximum lateral displacement decreased by 1.09 times than that of OSSC at the second column and by 1.03 times for 7th column. However, by encasing the stone column, the lateral bulging has significantly reduced, mostly because of the increased constrainment the encasing provides, as also suggested by other studies[23,46].

## Comparison of current work with previous research

While it is not feasible to directly compare this study with existing literature due to variations in soil type, model setup, loading criteria, and material parameters, an effort has been made to assess the results to validate this study's effectiveness in relation to previous research. As there are not much study available on the use of steel slag as a granular column under embankment loading where time vs. settlement graphs can be compared, therefore

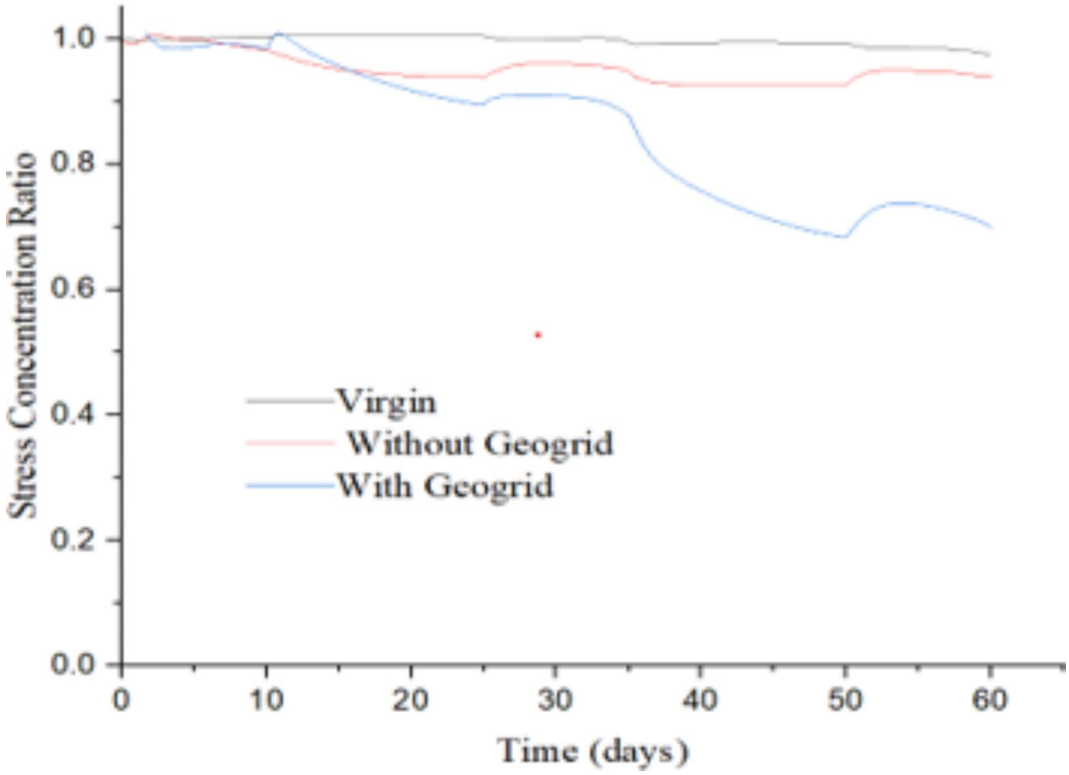



**Fig. 7**. Influence of OSSC and ESSC on stress concentration ratio.

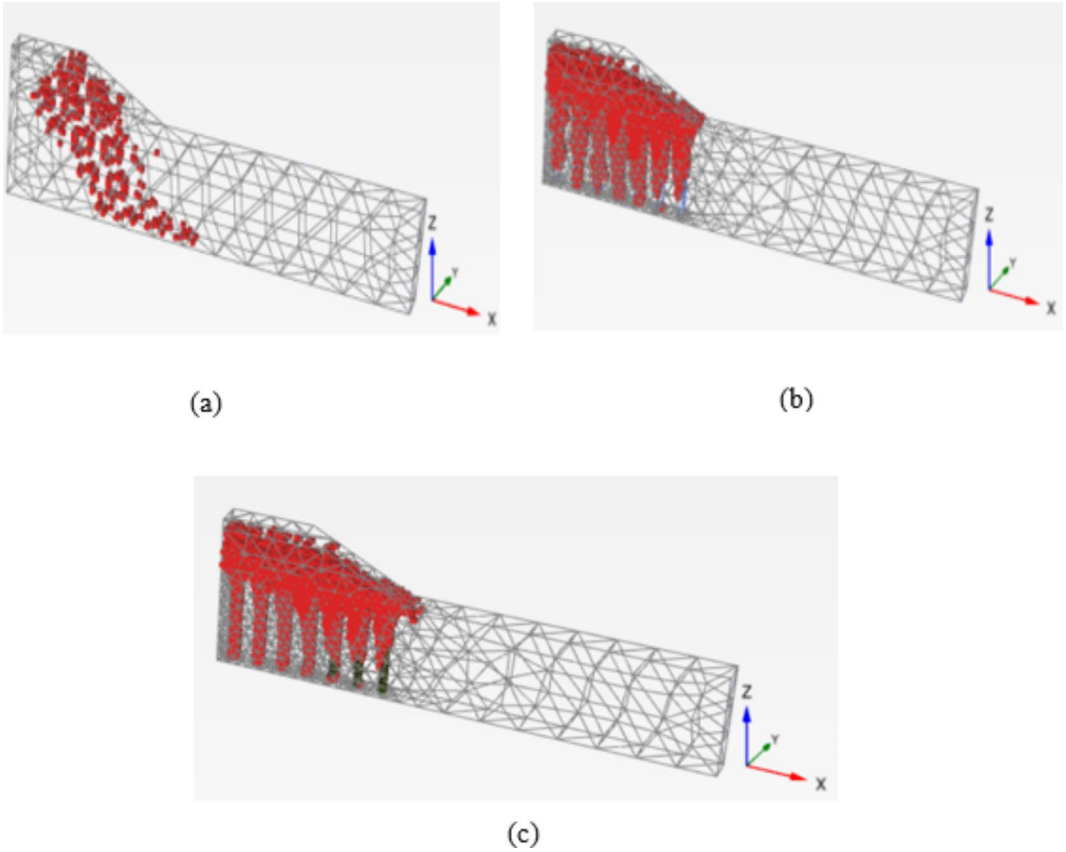



**Fig. 8**. Developed plastic points at final construction stage (a) Unreinforced (b) OSSC (c) ESSC.

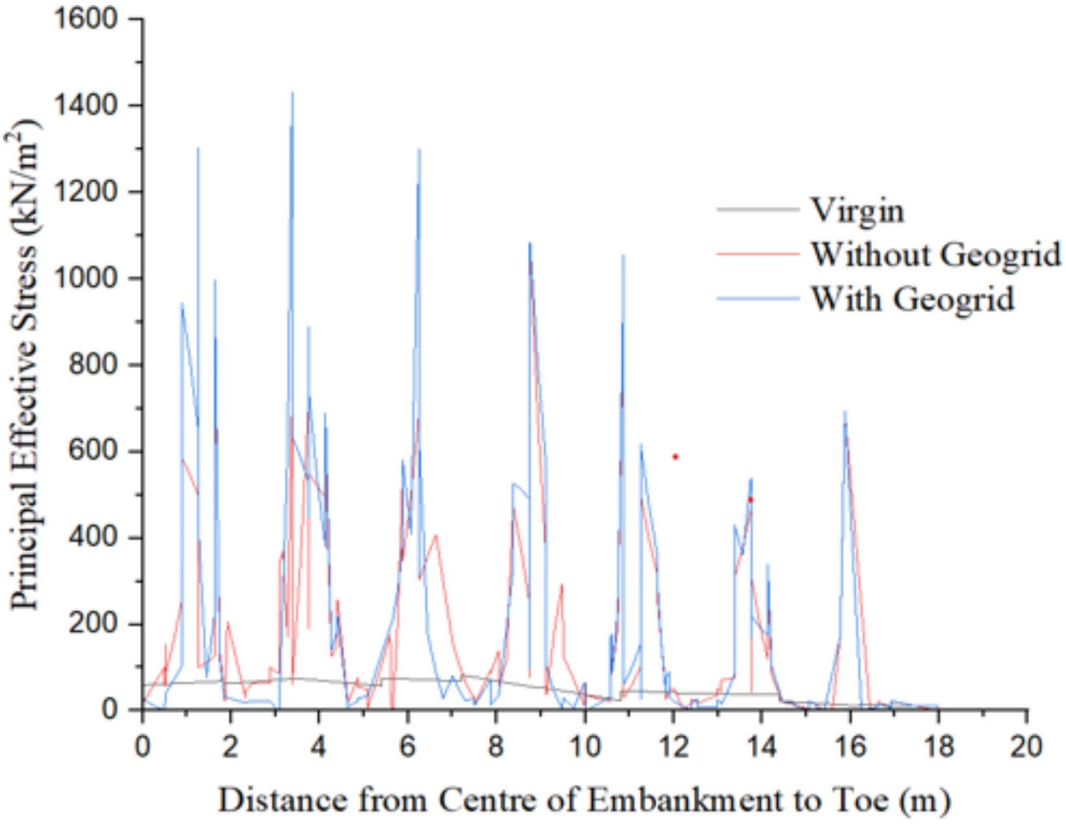



**Fig. 9**. Comparison of variation of principal effective stress below the embankment for the unreinforced case, OSSC and ESSC.

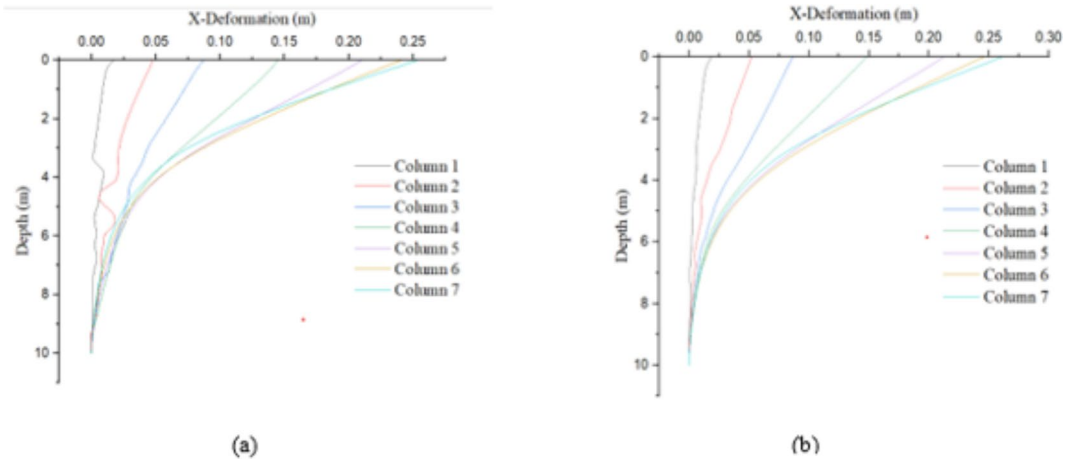



**Fig. 10**. Lateral displacement of all columns for (a) OSSC (b) ESSC.

a comparison has been made from the data of geosynthetic encased granular column to understand the trend. Figure 12 represents the comparison of load-settlement variation of the current study with the existing literature.

Rajesh and Jain[47], reported the findings of a series of numerical investigations conducted to assess the time-dependent performance of ordinary stone columns (OSC) and geosynthetic-encased stone columns (GESC). The tests were conducted with an embankment height of 6 m. The substrata utilised in the study consisted of a layer of soft clay with a thickness of 10 m, which was afterwards followed by a layer of hard rock strata. For both the Overlying Sand Cover (OSC) and the Geosynthetic Encased Sand Column (GESC) methods, a 1-meter-thick sand mat or drainage blanket is positioned on the surface of the soft clay layer. The groundwater table was located at the uppermost surface of the clay layer. The stone column, with a diameter of 400 mm, is installed by driving it from the sand mat and extending it to the complete depth of the soft clay. This results in a length of

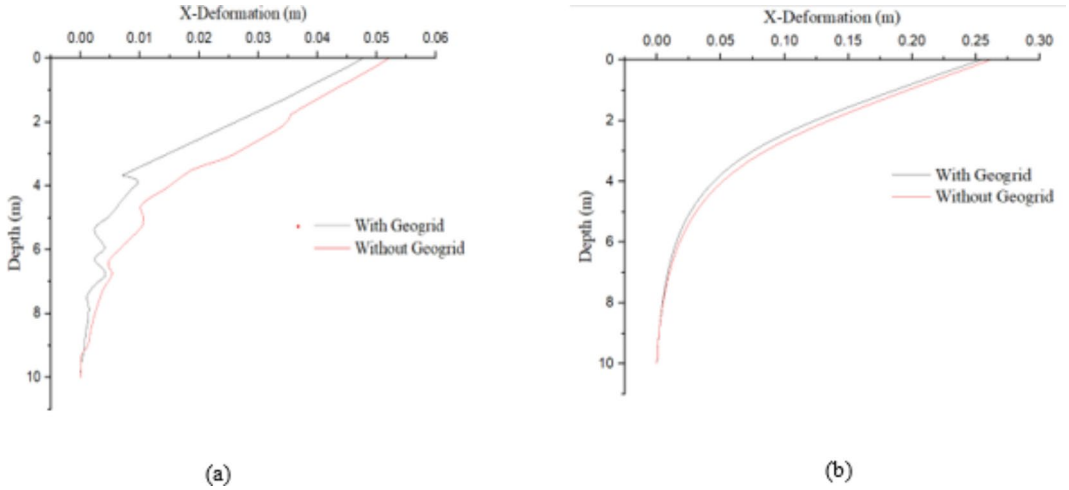



**Fig. 11**. Comparison of lateral deflection of column under OSSC and ESSC for (a) 2nd column (b) 7th column.

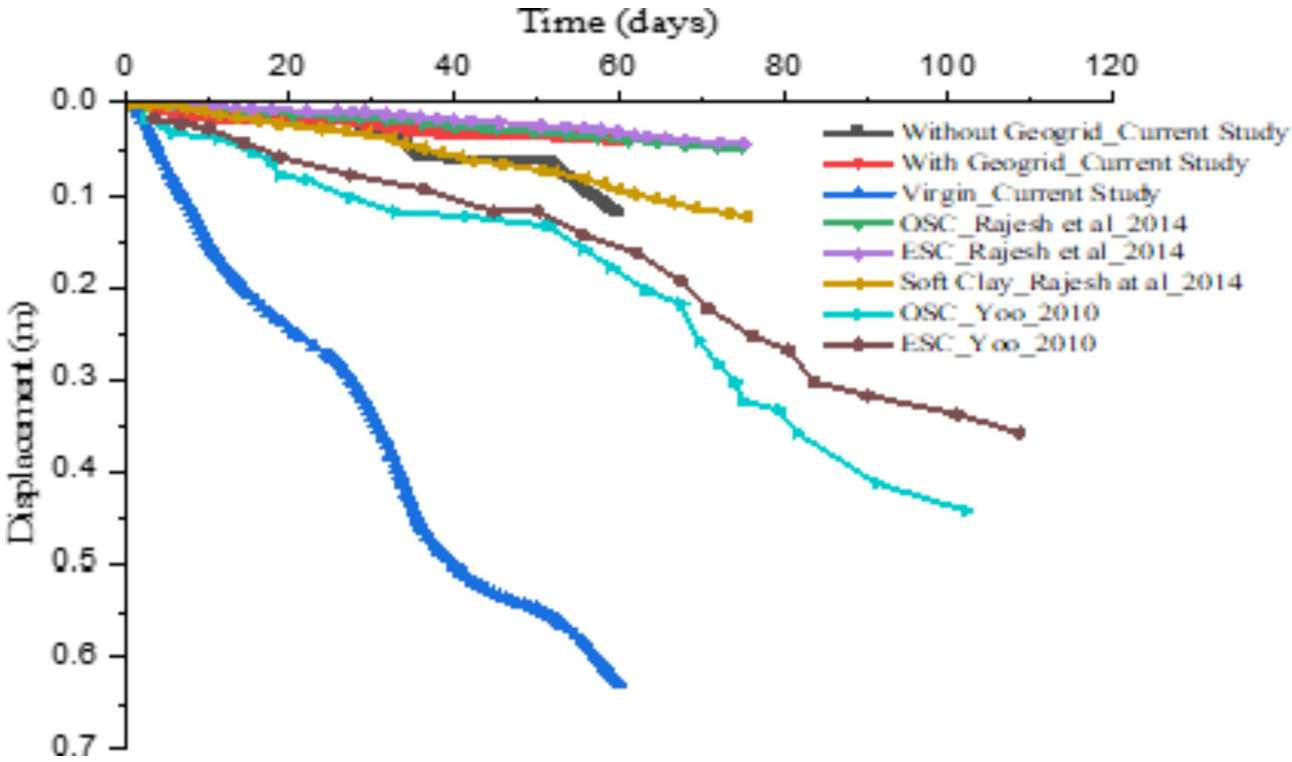



**Fig. 12**. Comparison of load-settlement behaviour of present study with existing literature.

the stone column, denoted as LS, measuring 11 m. In this work, a replacement ratio of 9% is utilised, with the diameter of the axisymmetric unit cell set at 1.338 m. The study determined that the ultimate settlement of OSC and GESC exhibited a reduction of 43% and 61%, respectively, in comparison to soft clay. A significant decrease in the duration required for the consolidation of ground treated with GESC was seen in comparison to ground treated with OSC.

Yoo[48] did the analysis which pertains to a 50-meter-wide embankment featuring side slopes with a 1:1 vertical-to-horizontal ratio. The embankment is constructed using sandy soil and is situated above a homogenous clay layer that is 10 m thick. Below the clay layer, there exists a stiff layer. A square grid-shaped GESCs with a 0.8 m diameter constructed of crushed stone is installed. Before the embankment fill is put, a sand mat with a thickness of 1 m is installed for column drainage. The findings suggest that the inclusion of geosynthetic encasement leads to increased stiffness of the stone column and a decrease in the amount of embankment load transferred to the soft ground. Consequently, this results in a reduction in overall settlement.

The detailed methodology, soil type, loading conditions, material properties and other parameters can be obtained from the referred study. As observed, a higher settlement is observed for untreated soil which reduces when it is being treated with ordinary steel slag column/granular column and further reduces when reinforced with encased steel slag column/granular column. The findings of the current study align with and can be compared to the patterns documented in prior investigations.

## Conclusion

In the current study, it was looked into how an embankment sitting on soft clay and reinforced with steel slag columns would behave over time both with and without encasing. To calculate the influence of the reinforcing method against embankment loading, a numerical model was created. A parametric analysis was performed to comprehend how OSSC and ESSC behave on various outputs such as settlement, excess PWP, SCR and lateral deformation of columns. The following conclusions from the current investigation can be made based on the observed results:

1. The settlement of soft clay at the end of the embankment construction phase has been significantly higher for the chosen configuration and model characteristics, compared to OSSC. Enclosing the column in a suitably stiffer geosynthetic material will result in a further reduction in settlement.
2. Underneath the embankment, where soft clay produces the maximum excess PWP, the amount of excess PWP decreases towards the embankment slopes. Excess PWP created on the untreated soft clay at the conclusion of the embankment building phase has been greater than OSSC for the selected monitoring location. When ESSC was employed, the excess PWP value was further decreased.
3. Stress concentration ratio (SCR) defined as stress in soft soil to the column at equal elevation decreases for OSSC as compared to the unreinforced soil. This decrease is even more pronounced for ESSC, indicating that ESSC is more successful than OSSC as a means of reinforcement.
4. Principal effective stress increases over the column compared to the surrounding soil when ESSC is used as compared to OSSC and unreinforced soil. Maximum stresses were experienced on the columns under the crest of the embankment.
5. The load capacity and rigidity of steel slag columns can be increased by completely encasing them in geosynthetic material. It is found that the lateral bulging of the stone columns is decreased and that the geosynthetic encasement contains them.

## Data availability

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Received: 16 July 2023; Accepted: 27 March 2024
Published online: 11 January 2025

## References

1. Srijan, Gupta, A. K. A Review article on Construction, Parametric Study and Settlement Behaviour of Stone Column. IOP Conference Series: Earth and Environmental Science, vol. 796, 012021. (2021).
2. Murugesan, S. & Rajagopal, K. Studies on the behaviour of single and group of geosynthetic encased stone columns. *J. Geotech. Geoenviron Eng.* **136** (1), 129–139 (2010).
3. Cengiz, C., Kilic, I. E. & Guler, E. Shear failure mode of granular column embedded unit cells subjected to static and cyclic shear loads. *Geotext. Geomembr.* **47**, 193–202 (2019).
4. Araujo, G. L. S., Palmeira, E. M. & Cunha, R. P. Behaviour of geosynthetic-encased granular columns in porous collapsible soil. *Geosynth Int.* **16**, 433–451 (2009).
5. Almeida, M. S. S., Hosseinpour, I., Riccio, M. & Alexiew, D. Behaviour of geotextile-encased granular columns supporting test embankment on soft deposit. *J. Geotech. Geoenviron Eng.* **141**, 04014116 (2015).
6. Black, J., Sivakumar, V., Madhav, M. R. & McCabe, B. An improved experimental test set-up to study the performance of granular columns. *Geotech. Test. J.* **29** (3), 193–199 (2006).
7. Rajagopal, K. & Mohapatra, S. R. Behaviour of geosynthetic encased granular columns under vertical and lateral loading. GA 2016–6th Asian Reg. Geosynth. Infrastruct. Dev. Proc, Conf. Geosynth, pp. KN83–KN99. (2018).
8. Hosseinpour, I., Almeida, M. S. S. & Riccio, M. Full-scale load test and finite-element analysis of soft ground improved by geotextile-encased granular columns. *Geosynth Int.* **22**, 428–438 (2015).
9. Chong, S. H. & Kim, J. Y. Nonlinear vibration analysis of the resonant column test of granular materials. *J. Sound Vib.* **393**, 216–228 (2017).
10. Tan, X., Feng, L. & Hu, Z. The equivalent shear strength properties of the composite soil reinforced by stone columns: an FDM-DEM-coupled numerical evaluation. *Environ. Earth Sci.* **80**, 125 (2021).
11. Srijan & Gupta, A. K. Effectiveness of triangular and square pattern of stone column with varying s/d ratio on consolidation behaviour of soil. Second ASCE India Conference CRISIDE, 1, 1648–1654. (2020).
12. Pandey, B. K. & Rajesh. Chandra, S. Performance of soft clay reinforced with encased stone column: a systematic review. *Int. J. Geosynthetics Ground Eng.* **8**, 40 (2022).
13. Mitchell, J. K. & Huber, T. R. Performance of a stone column foundation. *J. Geotech. Eng.* **111** (2), 205–223 (1985).
14. Xie, K. H., Lu, M. M., Hu, A. F. & Chen, G. H. A general theoretical solution for the consolidation of a composite foundation. *Comput. Geotech.* **36**, 24–30 (2009).
15. Verma, P. & Sahu, A. K. Effect of grouted granular column on the load carrying capacity of the expansive soil. *Int. J. Recent. Technol. Eng.* **8**, 2606–2612 (2019).
16. Hosseinpour, I., Soriano, C. & Almeida, M. S. S. A comparative study for the performance of encased granular columns. *J. Rock. Mech. Geotech. Eng.* **11**, 379–388 (2019).
17. Orekanti, E. R. & Dommaraju, G. V. Load-settlement response of geotextile encased laterally reinforced granular piles in expansive soil under compression. *Int. J. Geosynth Gr Eng.* 5–17. (2019).
18. FHWA (Federal Highway Administration). Highway sub-drainage design manual. Rep. TS-80-224, Washington, DC. (1980).

19. Jayarajan, J. & Karpurapu, R. Settlement analysis of geosynthetic encased granular column treated soft clay deposits. *Int. J. Geotech. Eng.* **14**, 473–489 (2020).
20. Dash, S. K. & Bora, M. C. Influence of geosynthetic encasement on the performance of stone columns floating in soft clay. *Can. Geotech. J.* **50**, 754–765 (2013).
21. Aryal, S. & Kolay, P. K. Long-term durability of ordinary portland cement and polypropylene fibre stabilized kaolin soil using wetting–drying and freezing–thawing test. *Int. J. Geosynth Gr Eng.* **6**, 1–15 (2020).
22. Mohapatra, S. R., Rajagopal, K. & Sharma, J. Direct shear tests on geosynthetic-encased granular columns. *Geotext. Geomembr.* **44**, 396–405 (2016).
23. Srijan, S. & Gupta, A. K. Vertically and horizontally reinforced end-bearing stone column: An experimental and numerical investigation. *Appl. Sc.* **13**, 11016 (2023).
24. Maghool, F., Arulrajah, A., Suksiripattanapong, C., Horpibulsuk, S. & Mohajerani, A. Geotechnical properties of steel slag aggregates: shear strength and stiffness. *Soils Found.* **59** (5), 1591–1601 (2019).
25. Ashango, A. A. & Patra, N. R. Behavior of expansive soil treated with steel slag, rice husk ash, and lime. *J. Mater. Civ. Eng.* **28** (7), 06016008 (2016).
26. Jafer, H. M. Development of a new ternary blended cementitious binder produced from waste materials for use in soft soil stabilisation. *J. Clean. Prod.* **172**, 516–528 (2018).
27. Yang, D., Sasaki, A. & Endo, M. Reclamation of a waste arsenic-bearing gypsum as a soil conditioner via acid treatment and subsequent Fe (II)–As stabilization. *J. Clean. Prod.* **217**, 22–31 (2019).
28. Sharma, A. & Sharma, R. K. Strength and drainage characteristics of poor soils stabilized with construction demolition waste. *Geotech. Geol. Eng.* **38** (5), 4753–4760 (2020).
29. Wani, K. M. N. S. & Mir, B. A. Stabilization of weak dredged soils by employing waste boulder crusher dust: a laboratory study. *Geotech. Geol. Eng.* **38**, 6827–6842 (2020).
30. Zhang, T. Durability of silty soil stabilized with recycled lignin for sustainable engineering materials. *J. Clean. Prod.* **248**, 119293 (2020).
31. Brand, A. S. & Roesler, J. R. Steel furnace slag aggregate expansion and hardened concrete properties. *Cem. Concr Comp.* **60**, 1–9 (2015).
32. Dash, M. K., Patro, S. K. & Rath, A. K. Sustainable use of industrial-waste as partial replacement of fine aggregate for preparation of concrete – a review. *Int. J. Sustain. Build. Environ.* **5**, 484–516 (2016).
33. Ferreira, V. J. et al. Evaluation of the steel slag incorporation as coarse aggregate for road construction: technical requirements and environmental impact assessment. *J. Clean. Prod.* **130**, 175–186 (2016).
34. World Steel Association. Steel Statistical Yearbook 2019—Concise Version, Belgium, Brussels.
35. U.S. Geological Survey, Mineral Commodity Summaries Washington, U.S. (2020).
36. Kang, L., Du, H. L., Zhang, H. & Ma, W. L. Systematic research on the application of steel slag resources under the background of big data. Complexity 1–12. (2018).
37. Hosseinabadi, M. J. R., Bayat, M., Nadi, B. & Rahimi, A. Utilisation of steel slag as a granular column to enhance the lateral load capacity of soil. *Geomech. Geoengin.* **17** (5), 1406–1416 (2021).
38. Kumar, H. & Varma, S. A review on utilization of steel slag in hot mix asphalt. *Int. J. Pavement Res. Technol.* **14**, 232–242 (2021).
39. Poh, H. Y., Ghataora, G. S. & Ghazireh, N. Soil stabilization using basic oxygen steel slag fines. *J. Mater. Civ. Eng.* **18**, 229–240 (2006).
40. Shen, W., Zhou, M., Ma, W., Hu, J. & Cai, Z. Investigation on the application of steel slag-fly ash-phosphogypsum solidified material as road base material. *J. Hazard. Mater.* **164**, 99–104 (2009).
41. Akinwumi, I. I., Adeyeri, J. B. & Ejohwomu, O. A. Effects of steel slag addition on the plasticity, strength, and permeability of lateritic soil. ICSDEC 2012 Dev. Front. Sustain. Des. Eng. Constr. - Proc. 2012 Int, Conf. Sustain. Des. Constr., American Society of Civil Engineers, pp. 457–464. (2013).
42. Wu, J. et al. Expansive soil modified by waste steel slag and its application in subbase layer of highways. *Soils Found.* **59**, 955–965 (2019).
43. Lang, L., Song, C., Xue, L. & Chen, B. Effectiveness of waste steel slag powder on the strength development and associated micro-mechanisms of cement-stabilized dredged sludge. *Constr. Build. Mater.* **240**, 117975 (2020).
44. Tozsin, G., Yonar, F. & Yucel, O. Utilization possibilities of steel slag as backfill material in coastal structures. *Sci. Rep.* **13**, 4318 (2023).
45. Ghisman, V., Muresan, A. C. & Buruiana, D. L. Waste slag benefits for correction of soil acidity. *Sci. Rep.* **12**, 16042 (2022).
46. Srijan, Vatsa, H. & Singh, A. A study on the behaviour of soft clay under embankment loading. *IOP Conf. Ser. Earth Environ. Sci.* **1326**(1), 012120 (2024).
47. Rajesh, S. & Jain, P. Influence of permeability of soft clay on the efficiency of stone columns and geosynthetic-encased stone columns – a numerical study. *Int. J. Geotech. Eng.* **9**(5), 483–493 (2014).
48. Yoo, C. Performance of geosynthetic-encased stone columns in embankment construction: Numerical investigation. *J. Geotech. Geoenviron. Eng.* **136**(8), 1148–1160 (2010).


## Author contributions

Srijan: Conceptualization, Methodology, Writing – original draft, Investigation. A.K. Gupta: Supervision, Writing – review & editing, Resources, Conceptualization.

## Declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to S.

**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

# T&D Circuit Segmentation for Improved Resilience—A Texas Utility Case Study

Sarat Chandra Vegunta[1], Nelson Bacalao[1], Giovani Sasso Scarpati[1], Azadeh Shadman[1], and Ram Nath[1]

[1]Affiliation not available

January 05, 2025

**Abstract**

Several lessons were learned from the post-analysis of the 2021 Winter Storm Uri's devasting impact on Texas' electricity grids. Among the lessons was the need for a greater granularity and flexibility by which Texas electricity utilities can shed load during emergencies as directed by the local Independent System Operator (ISO), while protecting supply to critical loads and minimizing impact on customers. This paper, following a review of principal lessons learned following the storm, summarizes the work, concepts, and findings from a transmission and distribution circuit segmentation study performed on a Texas electricity utility's system.

Posted on 5 Jan 2025 — CC-BY 4.0 — https://doi.org/10.36227/techrxiv.173603691.14780084/v1 — e-Prints posted on TechRxiv are preliminary reports that are not peer reviewed. They should not b...

# T&D Circuit Segmentation for Improved Resilience—A Texas Utility Case Study

Sarat Chandra Vegunta, *Senior Member, IEEE*, Nelson Bacalao, *Member, IEEE*, Giovani Sasso Scarpati, *Member, IEEE*, Azadeh Shadman, *Member, IEEE*, and Ram Nath, *Senior Member, IEEE*

***Abstract*--Several lessons were learned from the post-analysis of the 2021 Winter Storm Uri's devasting impact on Texas' electricity grids. Among the lessons was the need for a greater granularity and flexibility by which Texas electricity utilities can shed load during emergencies as directed by the local Independent System Operator (ISO), while protecting supply to critical loads and minimizing impact on customers. This paper, following a review of principal lessons learned following the storm, summarizes the work, concepts, and findings from a transmission and distribution circuit segmentation study performed on a Texas electricity utility's system.**

***Index Terms*--Climate change, power distribution, power system economics, power system planning, power system reliability, power transmission, power transmission planning, systems operation.**

## I. Introduction

THE 2021 Winter Storm Uri (WSU)—a devastating winter storm during February 13–17, 2021—impacted a large portion of the United States and Northern Mexico. Texas was particularly hard-hit, with millions of residents experiencing prolonged power outages and water supply issues. Oklahoma, Kansas, Missouri, Nebraska, and New Mexico also experienced rolling blackouts and hazardous conditions extended as far north as Canada. WSU, contributed to as many as seven-hundred deaths, and led to approximately US $200B in economic damages, making the storm most expensive natural disaster to ever hit the USA [1, 2]. And there is evidence that weather-related billion-dollar disaster events, impacting electricity grids, are on the rise in the USA and globally [3, 4].

Several lessons were learned from the post-analysis of the impact of the WSU on Texas electricity grids. Among the lessons was the need for an improved electricity Transmission and Distribution (T&D) Circuit Segmentation (CS) to increase granularity and flexibility by which Texas utilities can shed load during emergencies, while protecting supply to critical loads [5].

A Circuit Segmentation Study (CSS) examines as follows: (a) whether and how the utility's electricity T&D circuits can be segmented to rotate outages more evenly across customers and circuits, while maintaining and or improving electricity supply protections offered to critical facilities; (b) feasibility of using sectionalization and or automation to segment circuits that host significant numbers of critical facilities into smaller segments to enable more granular and flexible electricity outage management; (c) identify feeders (fdrs.) with critical facilities that, if equipped with facility-specific backup power, can enhance the utility's outage management flexibility; and (d) estimate—via the CSS—the required investments and improvements with an improved CS.

Although a transfer of critical load, either manually or via an Automatic Transfer Switch (ATS), from an interrupted main supply to a backup is not new, the frequency and scale of such need has been unprecedented, driven by climate change impact on electricity grids. In fact, in the context of the abovementioned aspects, it was only after the WSU, the term circuit segmentation has been extensively used in the literature, and more so since the introduction of the State of Texas' Public utility Regulatory Act (PURA) 38.078 [6].

Additionally, most existing grid-edge Advanced Metering Infrastructure (AMI) based meters have built-in load-interrupting switches. These switches enable electricity utilities to remotely connect or disconnect customer loads, and if used for Load Shed Rotation (LSR) would potentially offer much greater granularity for load shedding without interrupting critical loads during emergencies (this was explored in [7]).

While this is an exciting and promising option for an LSR, the use of existing AMI-based load-interrupting switches for the extent and criticality of the system operations required for LSR applications is not fully understood. This was a risk that both the utility and the paper's authors had to consider when determining the best approach for increased circuit segmentation in the work described below.

This paper, building on the authors' work in [8, 9], following a summary of principal lessons learned following the storm, summarizes the work, developed concepts, and findings from a CSS—this study, detailed here, was completed in 2024 in compliance with the requirements of the State of Texas' PURA 38.078 for a Texas-based electricity T&D utility ("utility").

## II. 2021 Winter Storm Uri and Lessons Learned

The post-WSU analysis of the impact of the storm on the Texas electric power system, and in particular the performance of the power system during the WSU, and the impact of that performance on the end customers interconnected to the system captured several, important lessons for the improvement of the

S. C. Vegunta, N. Bacalao, A. Shadman, and R. Nath are with Siemens Industry, Inc., 400 State St 4th Floor, Schenectady, NY 12305, USA (e-mails: chandra.vegunta@siemens.com, nelson.bacalao@siemens.com, azi.shadman@siemens.com, and ram.nath@siemens.com).

G. S. Scarpati is with Siemens Infrastructure & Industry, Inc., São Paulo, SP 05110-000 Brazil (e-mail: giovani.scarpati@siemens.com).

Texas power system and its resilience. These lessons, broadly, can be grouped into categories as follows: (a) the procedures implemented to conduct the ERCOT's (ISO's) load shed instructions, (b) the availability of electricity generation fleet and performance of that fleet, and (c) the need for a reform of the energy market. Salient lessons, among several lessons (for e.g., detailed in [8]), in these categories are detailed below.

### *A. Loadshedding Procedure Lessons*

Limiting the duration of the electricity outages by more frequent load-shed rotation could have mitigated some of the impacts. There were feeders where critical loads were located that were not shed and all customers in the feeders (critical and not critical) remained with service, while other feeders had extended outages, resulting in important damages to property. Furthermore, the curtailment of gas compressors aggravated the loss of electricity generation; these types of electrical load should have had a high priority and remained uninterrupted during emergencies and reduced electricity generation [10].

### *B. Electricity Generation and Fuel Availability Lessons*

Poor generation weatherization and procedures to maintain the electricity generation availability under extreme weather conditions, as those experienced during the WSU and in general generations reserves, were the central cause of the impacts [2, 10]. Regulatory and markets mechanisms should be implemented to provide signals for enhancing the generation availability and reserves [2, 10].

### *C. Electricity Generation and Fuel Availability Lessons*

Texas utilities—coordinating with other stakeholders, implementing demand response programs, and ensuring adequate fuel reserves for power generation—should develop robust contingency plans for extreme weather events (e.g., WSU). The existing energy market that relies on the scarcity signals failed to incentivize adequate generation reserves and availability signals to avoid the negative impacts of the WSU on the economy and consumers alike. Post-WSU electricity market reforms should address this and "promote the incorporation of dispatchable power generation capacity (e.g., natural gas-fired power plants) located near load centers and increase natural gas storage" [10].

### *D. Texas Legislature Bills*

In response to the WSU's disaster, the Texas Legislature passed the House Bill (HB) 1500 [11]: The bill mandates rulemakings and orders by the PUCT, indicating the bill's wide-ranging impact on the utility regulation and administration. The Senate Bill (SB) 2627 [12], The Powering Texas Forward Act, on the other hand, directed the PUCT to create and administer the Texas Energy Fund to provide grants and loans to finance the construction, maintenance, modernization, and operation of Texas electric facilities. The program (with available funds up to US $7.2B) finances new construction or upgrades to the existing dispatchable electric generating facilities within the ERCOT power region. To be eligible for a loan, the project must add a minimum of 100 MW new generation to the ERCOT grid. The loan could be financed up to 60% of the project's cost. As of the date writing this paper, the PUCT has received a total of 125 notices of interest for 55.9 GW of generation and a request for US $38.9B in financing [13].

The PUCT, in Jan. 2022, approved a blueprint for electricity market design in two Phases [14]: Phase I consist of directly actionable measures and Phase II of the blueprint incorporates long-term market design reforms to promote the supply of dispatchable generation and develop a backstop reliability service. While Phase I measures will improve the reliability of the system, a new reliability product is needed for the long-term reliability of the system and to reduce the likelihood of ERCOT requiring Load Serving Entities (LSEs) to shed load during events like those experienced the WSU repeat.

### *E. WSU Lessons Learned Summary and Recommendations*

In the authors' opinion, the WSU provided a stark example of the dire consequences to people, assets, and economy of not adequately managing and timely recovering from a major event. The shortcomings were multiple and complex, and so were the required solutions.

There is clearly a need for greater initiative, coordinated planning, adaptation, and preparedness for addressing the impacts of climate change on a diverse, complex power system, particularly with an increase in the integration of renewable generation. Many of the WSU issues could have been avoided, and there are mechanisms (enhanced reserve requirements or ancillary systems, better management of load curtailment when required, weatherization, and promotion of new generation entries to mention a few) that are being put in place since the WSU to manage the worst impacts. However, while this is important, the WSU is just one of the many ways that climate change will affect the power system, as the Houstonians were able to appreciate firsthand after the 2024 Hurricane Beryl.

There is a need for an integrated planning, carried out by multidisciplinary teams with advanced knowledge, methods, and tools to proactively anticipate how the system and its customers will be impacted next and prepare for it via a Storm Readiness Working Group (SRWG). The SRWG could be part of various organizations, but the Independent System Operators (ISOs) seem to be its natural home.

## III. The Utility's T&D System for the CSS

The Texas-based utility has about 150 substations and 360 Medium Voltage (MV) distribution feeders. It serves a large group of residential and business customers, connected to Low Voltage (LV) secondary (480 V), MV primary (4.16–25 kV), or directly at transmission high voltage. Overall, the utility has a peak electricity load of about 2.5 GW, approx. 2.7% of ERCOT's peak electricity load demand.

ERCOT manages the electricity supply to over 26 million customers in the State of Texas, USA, ensuring that the electricity supply continuously meets the electricity demand. Following the WSU, ERCOT requires electricity utilities and retail providers in Texas to periodically provide information pertaining to the electricity load shed mechanisms, such as type of customers supplied, procedures for the consideration of critical load, and reducing electricity consumption during

emergencies, when electricity supply falls short of demand and to avoid uncontrolled system blackouts. Upon receiving ERCOT's Energy Emergency Alert (EEA) Level 3, a Transmission Operator (TO) is obligated to shed a percentage of its load—calculated as the TO's share of ERCOT's previous year's peak electricity demand—of ERCOT's requested load shed amount.

Per [15], Critical Loads (CLs) are "Loads for which electric service is considered crucial for the protection or maintenance of public safety; including but not limited to hospitals, police stations, fire stations, critical water and wastewater facilities, and customers with special in-house life-sustaining equipment". And a Critical Natural Gas Facility, per [15], is "A facility designated as a critical customer by the Railroad Commission of Texas under §3.65(b) of this title (relating to Critical Designation of Natural Gas Infrastructure) unless the facility has obtained an exception from its critical status;" this is referred here as the Critical Load Serving Generation (CLSG).

In ERCOT region, except for the Under Frequency Load Shed (UFLS) loads, Under Voltage Load Shed (UVLS) loads, CLSG facilities, and the Water Treatment facilities, the remaining facilities (including CLs) are expected to participate in the LSR.

To avoid disruption to CLs, additional CS granularity to maintain supply to the CLs during the LSR was required.

## IV. Distribution Feeder Clusters Identification

A Representative Distribution Sector (RDS)—a tool for distribution analysis—leverages the fact that the distribution system design generally follows similar and consistent design rules and practices for feeders that supply similar customer load. The RDS methodology cluster feeders with similar customer load and insights drawn analyzing a single Representative Feeder (RFD) in that Feeder Cluster, which can generally be extended to all feeders in that feeder cluster. The authors of this paper tested this expectation to ensure that the tool's applicability to the utility's distribution system (detailed in Section III). Feeder Cluster may be physically connected in the same utility business unit or different business units and or supplied by the same substation or different substations, that are similar among feeders in that cluster. An RFD in a Feeder Cluster represents all feeders in that Feeder Cluster.

The applied methodology to identify the Feeder Clusters and RFDs in those Feeder Clusters was as follows:

- Feeders with majority (≥80%) of a particular load type were identified as feeders of those types: Residential, Commercial or Industrial. Feeders where this condition failed, they were considered as Mixed feeders, generally residential and commercial.
- Residential feeders with per customer consumption intensity 0.01–1 MW/Cust. or >1 MW/Cust. were reconsidered as Commercial or Industrial feeders, respectively. The load per customer was too high for such feeders to be characterized as Residential and, therefore, would be an outlier in that feeder-type cluster. Similarly, Mixed feeders with per customer consumption intensity 0.025–1 MW/Cust. And >1 MW/Cust. were reconsidered as Commercial and Industrial feeders, respectively.
- Feeders with less than two (2) customers—e.g., Industrial feeders with dedicated connections—were not considered in the clustering analysis. Similarly, feeders with <0.5 MW feeder-head MW flow were not considered in the clustering analysis. Feeders with fewer customers or smaller overall load would not lend themselves to the CS analysis, as such feeders can be effectively switched-off at the substation, affecting a few customers, which is the objective of the segmentation.
- Furthermore, in the resultant feeder dataset for each feeder type, statistical z-scores for the feeders' customer consumption intensity and customer density values were calculated (a z-score, per [16], is "the deviation of an individual value from the mean measured in standard deviations"). Feeders with z-scores greater than a value of three (3) on those metrics were considered outliers and removed from clustering analysis input data.
- For the Residential, Commercial, and Mixed feeder types, a set of three (3), three (3), and two (2) Feeder Clusters (eight (8) total), respectively, were identified that best represent the utility's feeders. For each Feeder Cluster, its centroid (supplemented with practical considerations) was used to find the RFDs as it shows the "average" of the cluster (see Figure 1–3).

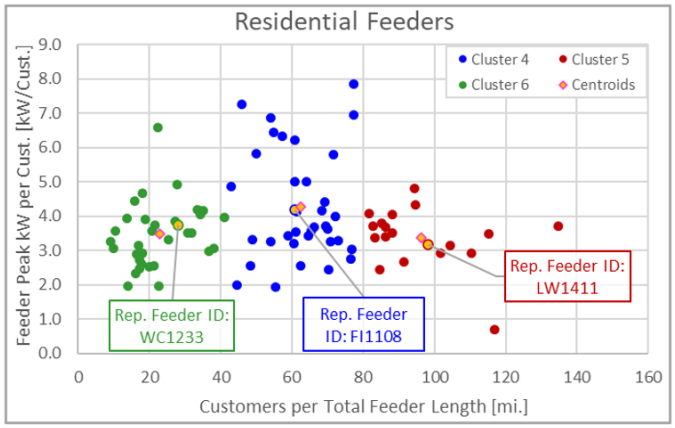



Fig. 1. Identified Residential Feeder Clusters & RFDs in those Clusters

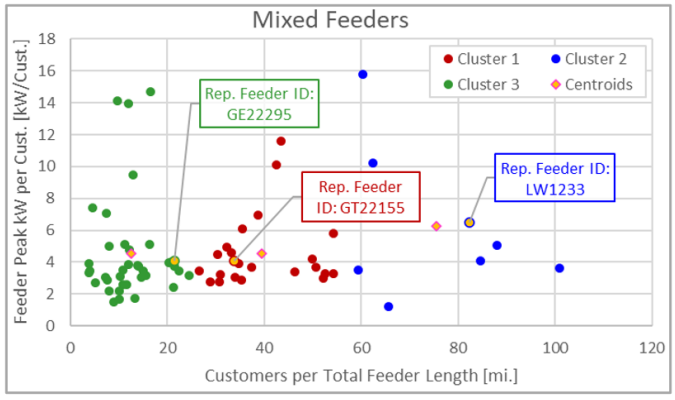



Fig. 2. Identified Mixed Feeder Clusters & RFDs in those Clusters

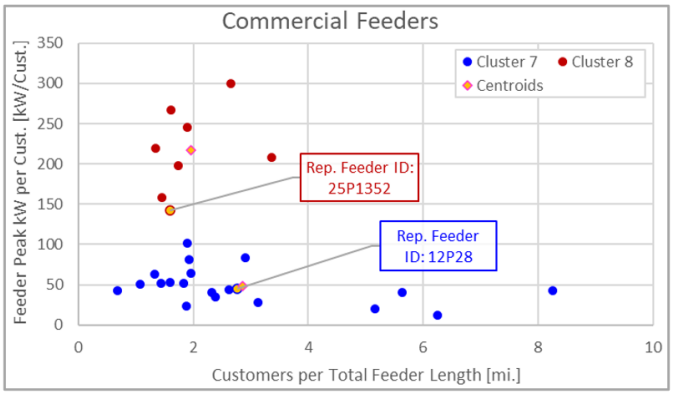



Fig. 3. Identified Commercial Feeder Clusters & RFDs in those Clusters

- Upon further analysis, it was found that feeders in Feeder Cluster 7 and 8 had CLs and CLSGs and were mostly radial, dedicated feeders to those loads. Therefore, the feeders in those clusters cannot participate in the LSR (see details in Section III) and would remain energized during an EEA Level 3 event. As such, no further CS was required for the feeders in Feeder Cluster 7 or 8.

## V. Distribution CSS Methodology

This section outlines the applied CSS methodology, which involves determining the switching points on an RFD within its Feeder Cluster. This process facilitates interruption of a feeder's non-Critical Loads (non-CLs), while ensuring CLs remain energized or supplied via an adjacent (Adj.) feeder or a set of adjacent feeders.

### *A. Feeder Clusters Model Development*

A set of six (6) electrical network area models, one per county with its set of feeders, were prepared. The feeder models included Medium Voltage (MV) network equipment, such as overhead lines and underground cables (including their conductors, impedances, ratings, etc.), capacitor banks, voltage regulators (including their impedances, voltage control range and setpoints, etc.), switches (both along the feeders for segmentation and those interconnecting feeders with normally open points), reclosers and fuses, along the feeders.

MV metered loads along those feeders were modelled as lumped three-phase loads, while the LV metered loads, supplied by MV or LV distribution transformers, were also modelled as three-phase MV connected lumped loads. Each feeder-head's active power in MW was maintained per the peak measurement value and power factor was maintained between unity and 0.95 lagging. The substations were modeled using grid infeed models, maintaining a constant 1.04 pu voltage (a typical voltage control set point for primary substations transformers' tap control) during normal operation and emergencies. The statuses of feeders' switches under normal operation and during typical emergencies were gathered and applied. Where there were minor data or model issues, reasonable assumptions or steps were taken to address them.

### *B. Circuit Segmentation Identification*

In identifying the optimum, minimum-cost CS for the RFDs, the following general considerations were adopted and applied:

- The existing switching infrastructure to formulate feeder network areas (or groups), consisting of non-CLs or CLs, suitable for load shedding or load transfer was used, whenever possible. Where this was not possible, a minimum set of switching devices—supplemented, if necessary, by a minimum set of new circuit sections—were considered to allow for an increased segmentation and or load transfer to adjacent feeder.
- All or most CLs on an RFD were transferred to its adjacent feeder(s), so that those CLs remain energized during an LSR.
- Load flow assessment was performed to confirm that the thermal loadings and voltages on the adjacent feeders to which load was transferred from the RFD were within the feeder's thermal and operational voltage limits.

The CS diagrams, illustrating the layout of the RFDs, have features as follows:

- The RFDs are always shown using green color, while other feeders of interest are shown using other colors.
- Specific to each RFD, sets (or clusters) of CLs are marked with CL-S# (where # in CL-S# is the set number of interest in the RFD).
- Similarly, specific to each RFD, network areas (or groups) of interest are marked with G# (where # in G# is the network group number of interest in the RFD) and are enclosed in dashed envelopes: network areas enclosed in green, dashed envelops are deenergized and those in red, dashed are energized. Approximate load in MW in each network area (or group G#) is also shown in relevant diagrams.
- Locations and the open-close statuses of the existing switches are also marked using Open or Close labels. Where a new switch was required to improve CS, the switch status (Open or Close) was annotated with the text (new).

### *C. Extrapolation of Representative Feeders' Findings*

Upon completion of the CSS on the distribution system feeders, the following general approach (with case specific variations) was applied to extrapolate the results obtained from analyzing the RFDs to all feeders in their respective Feeder Clusters:

- For feeders in each Feeder Cluster that have CLs, the RFD's percentage load shed following circuit segment was used to estimate the likely amount of load shed in each feeder. Furthermore, in each Feeder Cluster, the same minimum set of additional switching devices and or circuit section(s) to allow for an increased CS and or load transfer to adjacent feeders identified for the RFD was assumed for other feeders in that cluster.
- Since the feeders in the Feeder Clusters 7 and 8 were found to have CLSGs, they were not considered to participate in the LSR and would remain energized during an EEA Level 3 event.
- If a feeder in a Feeder Cluster does not have CLs, it was assumed that feeder can fully shed load.

## VI. Distribution CSS Results and Discussion

Presentation of detailed analysis and results for every Feeder Cluster is beyond this paper's scope. Therefore, to demonstrate the application of the CSS methodology described above, detailed CS analysis and results for a set RFDs (FDR#001–003), among the identified eight (8) RFDs, are presented. This general procedure was applied similarly (with case specific variations), for the remainder of the Feeder Clusters (FDR#004–006). Finally, overall results of the CSS applied to all Feeder Clusters and utility's distribution system are presented and discussed.

### *A. RDS-1 Circuit Segmentation Analysis and Results*

In Feeder Cluster 1, before CS, the feeder FDR#001 (highlighted in green in Figure 4) serves as the RFD, drawing power from Substation A to supply mixed loads. Among these are the CLs, identified as Critical Public Safety (CL-S1 to CL-S3, highlighted in orange in Figure 4), that require uninterrupted power supply. The proposed methodology involves transferring these CLs from the RFD FDR#001 to its adjacent feeder FDR#010 (highlighted in blue in Figure 4), ensuring a continuous power supply to these CLs.

The optimum RDS CS was identified to be as follows:

- The RFD was found to have two critical load sets (or clusters) CL-S2 and CL-S3. Although the critical load set CL-S1 is within proximity of the RFD, it can be normally supplied from the adjacent feeder FDR#010. See Figure 4 for details.
- Network area G1 has approx. 0.3 MW load and does not serve critical loads. Therefore, the load in network area G1 can be fully shed with existing switching elements (see the enclosed green, dashed envelop, marked G1, in Figure 4).
- Network area G2 has approx. 0.8 MW load and serves CL-S2. Therefore, the network area G2 was maintained energized by transferring its load to the adjacent feeder FDR#010. To achieve this, installation of a new (normally open) SCADA-controlled switch is recommended, to be closed for the transfer.

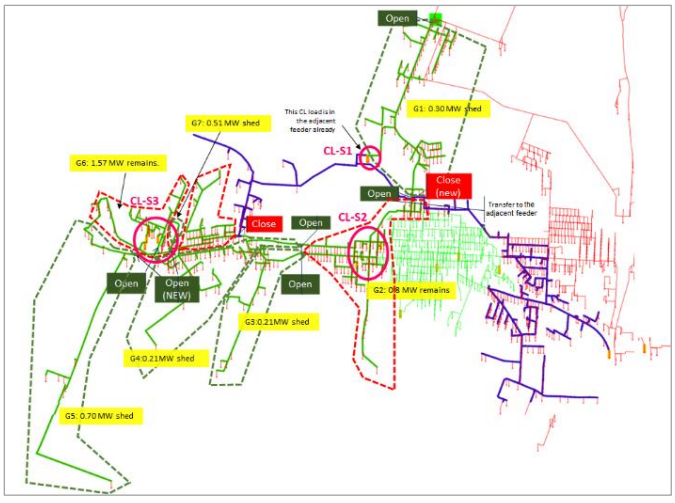



Fig. 4. Layout of the Rep. Fdr. FDR#001 (Green), Adj. Fdr. FDR#010 (Blue), and CLs (Orange) and segmentation of the Rep. Fdr. FDR#001

- The RFD was found to have two critical load sets (or clusters) CL-S2 and CL-S3. Although the critical load set CL-S1 is within proximity of the RFD, it can be normally supplied from the adjacent feeder FDR#010. See Figure 4 for details.
- Network area G1 has approx. 0.3 MW load, and it does not serve critical loads. Therefore, the load in network area G1 can be fully shed with existing switching elements (see the enclosed green, dashed envelop, marked G1, in Figure 4).
- Network area G2 has approx. 0.8 MW load and serves CL-S2. Therefore, the network area G2 was maintained energized by transferring its load to the adjacent feeder FDR#010. To achieve this, installation of a new (normally open) SCADA-controlled switch is recommended, to be closed for the transfer.

The CS methodology detailed in the previous section and above listed considerations for the RFD were simulated in the developed PSS®SINCAL model ("model"), detailed in Figure 4, and analyzed.

Results from the model-based power flow simulations, detailing the loading conditions before and after CS on the RFD (FDR#001) and its adjacent feeder (FDR#010) are summarized in Table I.

Results indicate that about 32% of the load (approx. 1.15 MW) on the RFD FDR#001 can be shed, while the remaining load on the RFD can be transferred to and supplied by the adjacent feeder FDR#010. As detailed in Table I, transfer of the RFD load to the adjacent feeder FDR#010 is not expected to cause any feeder thermal overloads or voltages below 0.9 pu The power flow simulation results (in Table I) indicate that there are no voltage drops or overvoltage violations on either feeder. Additionally, the loading on the feeders did not exceed 80% before or after CS.

TABLE I.
SUMMARY OF RFD FDR#001 LOAD SHED & TRANSFER

| Feeder (Fdr.) Description | Before CS | | | After CS | | | Load Shed | |
|---|---|---|---|---|---|---|---|---|
| | No. of CLs Supplied by Fdr. | Max. Fdr. Load [%] | Min. Fdr. Vol. [pu] | No. of CLs Supplied by Fdr. | Max. Fdr. Load [%] | Min. Fdr.Vol. [pu] | Load Shed [MW] | Load Shed [%] |
| Rep. Fdr. | 7 | 37.7 | 1.02 | - | - | - | 1.15 | 32 |
| Adj. Fdr.#010 | 2 | 22.6 | 1.04 | 9 | 39.3 | 1.03 | - | - |

### *B. RDS-2 Circuit Segmentation Analysis and Results*

In Feeder Cluster 2, before CS, the feeder FDR#002 (highlighted in green in Figure 5) serves as the RFD, drawing power from Substation B to supply mixed loads. Among these are the CLs, identified as Critical Public Safety (CL-S1 to CL-S6, highlighted in orange in Figure 5), that require uninterrupted power supply. The proposed methodology involves transferring these CLs from the RFD FDR#002 to its adjacent feeder FDR#020 (highlighted in blue in Figure 5), ensuring a continuous power supply to these CLs.

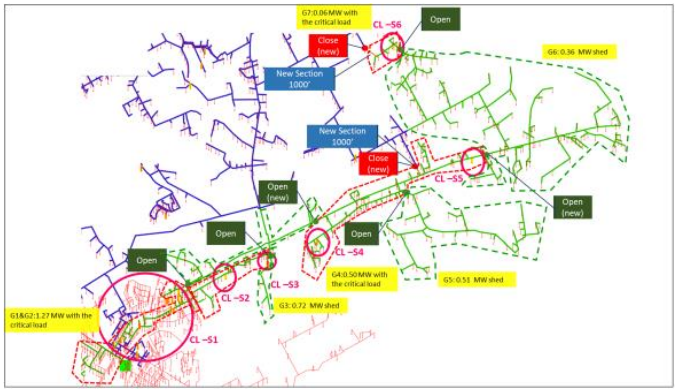



Fig. 5. Layout of the Rep. Fdr. FDR#002 (Green), Adjacent Fdr. FDR#020 (Blue), and CLs (Orange) and segmentation of the Rep. Fdr. FDR#002

The optimum RDS CS was identified to be as follows:

- The RFD was found to have critical load sets (or clusters) CL-S1 to CL-S6. Although the critical load set CL-S1 is within proximity of the RFD, it can be normally supplied from the adjacent feeder FDR#020. See Figure 5 for details.
- Network areas G1 and G2 have approx. 1.27 MW load and include the critical load sets CL-S1 to CL-S3 (see the enclosed red, dashed envelop, marked G1&G2, in Figure 5).
- Network area G3 has approx. 0.72 MW load, and it does not serve critical loads. Therefore, the load in the network area G3 can be fully shed (see the enclosed green, dashed envelop, marked G3, in Figure 5). To achieve this, installation of a new (normally closed) SCADA-controlled switch is recommended.
- Network area G4 has approx. 0.5 MW load and include the critical load sets CL-S4 and CL-S5. Therefore, the network area G4 was maintained energized by transferring its load to the adjacent feeder FDR#020 (see the enclosed red, dashed envelop, marked G4, in Figure 5). To achieve this, installation of a new section of the feeder under 1000 ft., and a new (normally open) SCADA-controlled switch nearby is recommended.
- Network area G5 has approx. 0.51 MW load, and it does not serve critical loads. Therefore, the load in the network area G5 can be fully shed (see the enclosed green, dashed envelop in Figure 5 marked G5). To achieve this, installation of a new (normally closed) SCADA-controlled switch nearby is recommended.
- Network area G6 has approx. 0.36 MW load, and it does not serve critical loads. Therefore, the load in the network area G6 can be fully shed (see the enclosed green, dashed envelop in Figure 5 marked G6).To achieve this, installation of a new (normally closed) SCADA-controlled switch is recommended.
- Network area G7 has approx. 0.06 MW load and include the critical load set CL-S6. Therefore, the network area G7 was maintained energized by transferring its load to the adjacent feeder FDR#020 (see the enclosed red, dashed envelop, marked G7, in Figure 5). To achieve this, installation of a new section of the feeder under 1000 ft., and a normally open SCADA-controlled switch is recommended.

The CS methodology detailed in Section V and the above listed considerations for the RFD were simulated in the developed model, detailed in Figure 5, and analyzed.

Results from the model-based power flow simulations, detailing the loading conditions before and after CS on the RFD (FDR#002) and its adjacent feeder (FDR#020) are summarized in Table II.

Results indicate that about 46% of the load (approx. 1.59 MW) on the RFD FDR#002 can be shed, while the remaining load on the RFD can be transferred to and supplied by the adjacent feeder FDR#020. As detailed in Table II, transfer of the RFD load to the adjacent feeder FDR#020 is not expected to cause any feeder thermal overloads or voltages below 0.9 pu The power flow simulation results (in Table II) indicate that there are no voltage drops or overvoltage violations on either feeder. Additionally, the loading on the feeders did not exceed 80% before or after CS.

TABLE II.
SUMMARY OF RFD FDR#002 LOAD SHED & TRANSFER

| Feeder (Fdr.) Description | Before CS | | | After CS | | | Load Shed | |
|---|---|---|---|---|---|---|---|---|
| | No. of CLs Supplied by Fdr. | Max. Fdr. Load [%] | Min. Fdr. Vol. [pu] | No. of CLs Supplied by Fdr. | Max. Fdr. Load [%] | Min. Fdr. Vol. [pu] | Load Shed [MW] | Load Shed [%] |
| Rep. Fdr. | 16 | 28.8 | 1.02 | 14 | 10.61 | 1.04 | 1.59 | 46 |
| Adj. Fdr.#020 | 27 | 35.7 | 1.01 | 29 | 42.23 | 1.01 | - | - |

### *C. RDS-3 Circuit Segmentation Analysis and Results*

In Feeder Cluster 3, before CS, the feeder FDR#003 (highlighted in green in Figure 6) serves as the RFD, drawing power from Substation C to supply mixed loads. Among these are the CLs, identified as Critical Public Safety (CL-S1 to CL-S6, highlighted in orange in Figure 6), that require uninterrupted power supply. The proposed methodology involves transferring these CLs from the RFD FDR#003 to its adjacent feeders FDR#030, and FDR#031 (highlighted in purple and red, in Figure 6), ensuring a continuous power supply to these CLs.

The optimum RDS CS was identified to be as follows:

- The RFD was found to have critical load sets (or clusters) CL-S1 and CL-S2. Although the critical load sets CL-S3 and CL-S4 are within proximity of the RFD, it can be normally supplied from the adjacent feeders FDR#031 and FDR#032. See Figure 6 for details.
- Network area G1 has approx. 2.14 MW load and include the critical load sets CL-S1, and CL-S2 (see the enclosed red, dashed envelop, marked G1, in Figure 6).
- Network area G2 has approx. 1.4 MW load, and it does not serve critical loads. Therefore, the load in the network area G2 can be fully shed load (see the enclosed green, dashed envelop, marked G2, in Figure 6). To achieve this, installation of a new (normally closed) SCADA-controlled switch is recommended.
- Network area G3 has approx. 1.1 MW load and include

the critical load sets CL-S3. Therefore, the network area G3 was maintained energized by transferring its load to the adjacent feeder FDR#031 (see the enclosed red, dashed envelop, marked G3, in Figure 6). To achieve this, installation of a new (normally open) SCADA-controlled switch is recommended.

- Network area G4 has approx. 0.75 MW load and include the critical load set CL-S4. Therefore, the network area G4 was maintained energized by transferring its load to the adjacent feeder FDR#030 (see the enclosed red, dashed envelop, marked G4, in Figure 6). To achieve this, installation of a new (normally open) SCADA-controlled switch nearby is recommended.
- Network area G5 has approx. 1.3 MW load, and it does not serve critical loads. Therefore, the load in the network area G5 can be fully shed (see the enclosed green, dashed envelop in Figure 6 marked G5).To achieve this, installation of a new (normally closed) SCADA-controlled switch nearby is recommended.

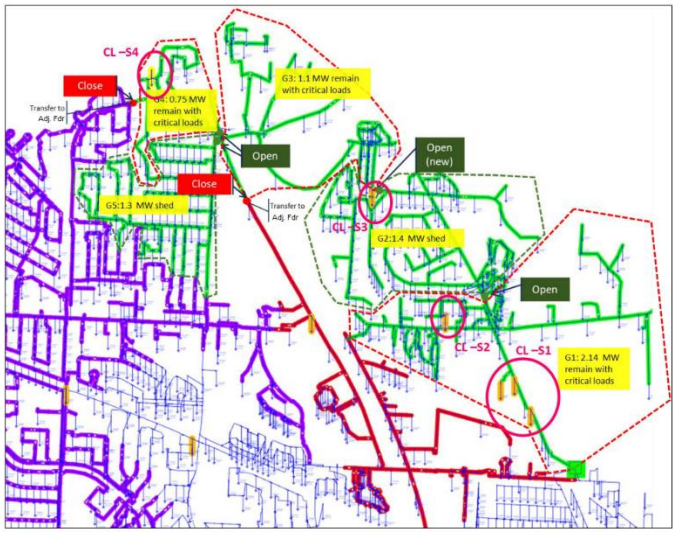



Fig. 6. Layout of the Rep. Fdr. FDR#003 (Green), Adj. Fdr. FDR#030 (Purple), Adj. Fdr. FDR#031 (Red), and CLs (Orange) and segmentation of the Rep. Fdr. FDR#003

The CS methodology detailed in Section V and above listed considerations for the RFD were applied in the developed model, detailed in Figure 6, and analyzed.

Results from the model-based power flow simulations, detailing the loading conditions before and after circuit segmentation on the RFD (FDR#003) and its adjacent feeders (FDR#030 and FDR#031) are summarized in Table III.

Results indicate that about 32% of the load (approx. 3.12 MW) on the RFD FDR#002 can be shed, while the remaining load on the RFD can be transferred to and supplied by the adjacent feeders FDR#30 and FDR#31. As detailed in Table III, transfer of the RFD load to the adjacent feeder FDR#30 and FDR#31 is not expected to cause any feeder thermal overloads or voltages below 0.9 pu The power flow simulation results (in Table III) indicate that there are no voltage drops or overvoltage violations on either feeder. Additionally, the loading on the feeders did not exceed 80% before or after CS.

TABLE III.
SUMMARY OF RFD FDR#003 LOAD SHED & TRANSFER

| Feeder (Fdr.) Description | Before CS | | | After CS | | | Load Shed | |
|---|---|---|---|---|---|---|---|---|
| | No. of CLs Supplied by Fdr. | Max. Fdr. Load [%] | Min. Fdr. Vol. [pu] | No. of CLs Supplied by Fdr. | Max. Fdr. Load [%] | Min. Fdr. Vol. [pu] | Load Shed [MW] | Load Shed [%] |
| Rep. Fdr. | 6 | 42.6 | 1.02 | 4 | 17.9 | 0.95 | 3.12 | 32 |
| Adj. Fdr.#30 | 4 | 32.3 | 1.03 | 5 | 39.9 | 1.03 | - | - |
| Adj. Fdr.#31 | 3 | 34.8 | 1.02 | 4 | 40.2 | 1.03 | - | - |

### *D. Utility Distribution Wide Analysis Results and Summary*

Table IV, for the Feeder Clusters, shows the extrapolated results of their RFDs applied to all other feeders in those clusters, calculating the total Feeder Cluster load shed before and after circuit segmentation. Results show that the potential load shed for RFDs ranges from a low 6.4% on Cluster 7 to a high of 77.4% on Cluster 2.

Industrial loads and Feeder Cluster identification outliers (among Mixed, Residential, and Commercial feeder types), depending on critical loads they have, are expected to be partially shed (see Table IV for details). This was identified by the analysis of the outliers.

TABLE IV.
SUMMARY OF THE CS ANALYSIS TECHNICAL RESULTS

| RFD (Fdr.) Cluster ID | Feeder Type | No. of CLs Supplied by Rep. Fdr. Cluster | Total Rep. Fdr. Cluster Load Shed Before CS [MW] | Total Rep. Fdr. Cluster Potential Load Shed After CS | |
|---|---|---|---|---|---|
| | | | | [MW] | [%]** |
| 1 | Mixed | 158 | 78.6 | 26.7 | 34.0 |
| 2 | Mixed | 31 | 35.9 | 27.8 | 77.4 |
| 3 | Mixed | 348 | 74.3 | 46.0 | 62.0 |
| 4 | Resident. | 255 | 202.0 | 103.7 | 51.4 |
| 5 | Resident. | 112 | 101.7 | 45.4 | 44.6 |
| 6 | Resident. | 304 | 110.8 | 36.5 | 32.9 |
| 7 | Comm. | 123 | 67.1 | 4.3 | 6.4 |
| 8 | Comm. | 54 | 47.1 | 12.4 | 26.3 |
| Non-Clustered Fdr. Groups (Outliers* or Indust. Loads**) | Mixed | 51 | 20.1 | 1.2 | 6.1 |
| | Resident. | 18 | 15.1 | 0.02 | 0.2 |
| | Comm. | 13 | 9.1 | 4.5 | 49.0 |
| | Indust. | 4 | 184.0 | 161.2 | 87.6 |
| Totals Across utility's Distribution System | | | 945.7 | 461.7 | 49.7 |

*Outliers feeders were not covered by the cluster identification procedure, hence only feeders with no critical loads were considered eligible for load shed. **Industrial feeders have very few customers and can be opened from the substation (no segmentation was possible), thus only feeders with no-critical loads were shed, as they have no critical loads, are expected to completely shed load. ***The percentage value is calculated as 100 times the ratio of the Total Rep. Fdr. Cluster Load Shed Before CS [MW] to the Total Rep. Fdr. Cluster Load Shed Before CS [MW].

Combining the above two findings, the total potential load shed after circuit segmentation during emergencies—while ensuring the critical loads remain energized or supplied via an adjacent feeder or adjacent feeders—is approx. 49.7% of the utility peak load (see Table IV for details).

To evaluate the investments required to implement the proposed circuit segmentation (results summarized in Table V), two options were considered: Base and Enhanced. The Base option consists of lower-cost SCADA-controlled motorized (load break) switches, which could be opened or closed as required by the utility outage management system's LSR program or system operators; and (b) an Enhanced option that instead uses reclosers, which could also be part of a feeder automation scheme like the Fault Location Isolation and Service Restoration scheme implemented by various utilities.

The unit costs for the required additional switches were derived using the RSMeans cost estimating database [15], adapted to the Texas prevailing costs and regional factors. The RSMeans estimating database provides costs at EPC (Engineering, Procurement, and Commissioning) level, and it includes costs for engineering, materials, tools, labor, contractor overhead and profits, local taxes, and insurance. To these costs, the authors assumed an added contingency of 25%, as there is significant uncertainty regarding actual field conditions and a 20% owner's cost that captures all additional costs that the utility will incur.

TABLE V.
REQUIRED INVESTMENTS (CAPEX, IN 2024 PRICES) WITH BASIC SOLUTION FOR ADDITIONAL CIRCUIT SEGMENTATION

| Rep. Fdr. Cluster ID | No. of New Switches per Rep. Fdr. | Cost per Solution-Set [US $K**/Solution-Set] | | Total of New Motorized Switches per Rep. Fdr. Cluster | Required Investment per Rep. Fdr. Cluster [US $M] | |
|---|---|---|---|---|---|---|
| | | Basic | Enhanced | | Basic | Enhanced |
| 1 | 7 | 62.7 | 113.3 | 161 | 10.08 | 18.23 |
| 2 | 2 | 62.7 | 113.3 | 14 | 0.87 | 1.58 |
| 3 | 8 | 108.1§ | 158.7§ | 304 | 22.49 | 37.88 |
| 4 | 2 | 62.7 | 113.3 | 66 | 4.13 | 7.47 |
| 5 | 6 | 62.7 | 113.3 | 114 | 7.14 | 12.91 |
| 6 | 2 | 62.7 | 113.3 | 68 | 4.26 | 7.70 |
| 7 | 0* | - | - | - | - | - |
| 8 | 0* | - | - | - | - | - |
| Totals across utility's distribution system | | | | 727 | 48.9 | 85.8 |

*Feeders with no new equipment required for the proposed circuit segmentation. **Symbols K, M denote thousands and millions of US dollars, respectively. ‡These do not include substation breakers. §For RFD ID 3, two (2) new OHL sections are required.

For the Base option, the unit price, which includes the cost for a remote Overhead SCADA-controlled three-phase switch, plus contingency and owner's cost, is estimated to be approximately US $62.7K. The same for the Enhanced option but with a SCADA-controlled vacuum recloser instead is estimated to be approximately US $113.3K. Additionally, a 1,000 ft overhead line with four (4) 40 ft. wood pole, one (1) crossarm with insulators and hardware each, and 3xAL 1/0 kcmil is expected to cost approximately US $45.4K, including the contingency and owner's cost.

For either option, SCADA upgrades, specific to the utility's system, are not expected to result in additional costs.

Based on the number of new switches required per RFD and extrapolating those numbers for their corresponding Feeder Clusters, the total required potential investment across the entire utility's distribution system is approx. US $49M for the Base Option to approx. US $86M (in 2024 prices).

## VII. TRANSMISSION SYSTEM ANALYSIS, RESULTS, AND DISCUSSION

A review of the utility's data by the authors identified that some transmission-connected industrial loads in the utility's system were part of automatic UFLS and UVLS schemes, and, therefore, as expected, these industrial loads are not part of the utility's LSR program. The remaining transmission-connected loads were categorized into groups based on their characteristics, such as Oil Field, Pipeline, Cryogenic, Refinery, Data Center, and Water Treatment Center. To gauge their potential for load shedding during emergencies, the response of industrial facilities during the WSU was analyzed.

Authors' review of the utility's WSU telemetry confirmed that the transmission-connected loads (e.g., industrial) voluntarily participated in the LSR, or were encouraged to reduce consumption given the extremely high energy prices. It was identified that approx. 110–214 MW of industrial load (i.e., approx. 5–10% of the MW rating of those loads) could be incorporated into the utility's existing LSR.

An industrial load shed survey conducted by the utility during this study had limited responses (only two), and it did not support the availability of backup power systems at industrial facilities. It is, however, important to note that the limited survey response samples may not be representative of most industrial loads and of their backup power systems used in LSR during emergencies.

A key insight drawn from the study was the need for discussions between the utility and its industrial load customers to gather and verify as follows: whether they (customers) can reduce non-essential loads during extreme emergencies, determining the time required for load reduction, the amount of load each customer can offer to reduce, the duration of load reduction, and the time interval needed between sequential load reductions for each industrial customer. These details should be worked out by utility collaborating with its industrial customers when modifying utility existing LSR program.

Although the collaboration between the utility and its industrial customer and the participation from industrial customers is crucial to achieving success in incorporating non-critical transmission-connected industrial loads into the utility's LSR, during the study (detailed here), a lack of interest among

industrial customers to provide survey responses was observed. Without the collaboration of such customers, it becomes a challenge at a transmission system level to achieve the desired load shed reduction goals.

## VIII. References


[1] "The Failure of Fossil Fuels: Learning From Winter Storm Uri," Sierra Club Legislative, Washington, DC, USA, [Online] Feb. 2022. Accessed: June 1, 2024. [Online]. Available: https://www.sierraclub.org/sites/default/files/The%20Failure%20of%20Fossil%20Fuels%20-%20Winter%20Storm%20Uri%20Report%20-%20Feb%202022.pdf

[2] "The Timeline and Events of the February 2021 Texas Electric Grid Blackouts," The University of Texas Austin, Austin, TX, USA, [Online] July 2021. [Online]. Available: https://energy.utexas.edu/sites/default/files/UTAustin%20%282021%29%20EventsFebruary2021TexasBlackout%2020210714.pdf

[3] "NOAA National Centers for Environmental Information (NCEI) U.S. Billion-Dollar Weather and Climate Disasters (2024)," National Centers for Environmental Information (NCEI), Asheville, NC, USA, [Online] 2024. Accessed: Oct. 20, 2024. [Online]. Available: https://www.ncei.noaa.gov/access/billions/

[4] "Electricity Grids and Secure Energy Transitions," International Energy Agency (IEA), Paris, France, [Online] Oct. 2023. Accessed: Oct. 20, 2024. [Online]. Available: https://iea.blob.core.windows.net/assets/ea2ff609-8180-4312-8de9-494bcf21696d/ElectricityGridsandSecureEnergyTransitions.pdf

[5] C. M. Wyman, "Unpacking House Bill 1500," GridMonitor, Austin, TX, USA, [Online] July 11 2023. Accessed: Oct. 20, 2024. [Online]. Available: https://dash2.gridmonitor.com/news/articles/?id=8501

[6] Texas Admin. Code, Public Utility Regulatory Act (PURA), part 2, chap. 25(C), rule 38.078, Enacted 2023. [Online]. Available: https://statutes.capitol.texas.gov/Docs/UT/htm/UT.38.htm

[7] C. Rocha, J. Green, V. Holsomback, and B. Deaver, "Strategic Load Shedding for Reduced Impact of Extreme Weather Events on Distribution Systems," in 2024 CIRED Workshop, Chicago, IL, USA, Nov. 7–8, 2024.

[8] S. C. Vegunta, N. Bacalao, G. S. Scarpati, and A. Shadman, "Uri Winter Storm Impact on the Texas Electricity Grid & Lessons Learned—A Review," in 2024 CIRED Workshop, Chicago, IL, USA, Nov. 7–8, 2024.

[9] S. C. Vegunta, N. Bacalao, G. S. Scarpati, A. Shadman, and R. Nath, "T&D Circuit Segmentation for Improved Resilience—A Case Study," in 2024 CIRED Workshop, Chicago, IL, USA, Nov. 7–8, 2024.

[10] J. Wolff, "Lessons from 2021 Winter Storm Uri for the August 2023 Heat Wave in Texas," [Online] Feb. 2022. Accessed: June 1, 2024. [Online]. Available: https://stillwaterassociates.com/lessons-from-2021-winter-storm-uri-for-the-august-2023-heat-wave-in-texas/

[11] Texas Legislature, 88th Texas Legislature, 88(R) Session Session. (2023, June 9). *Relating to the continuation and functions of the Public Utility Commission of Texas and the Office of Public Utility Counsel, and the functions of the independent organization certified for the ERCOT power region; increasing an administrative penalty*.

[12] Texas Legislature, 88th Texas Legislature, 88(R) Session Session. (2023, June 9). *Relating to funding mechanisms to support the construction, maintenance, modernization, and operation of electric generating facilities*.

[13] "Public Utility Commission of Texas Sees Strong Industry Response to Texas Energy Fund Loan Program," Austin, TX, USA, [Online] June 1 2024. Accessed: June 1, 2024. [Online]. Available: https://ftp.puc.texas.gov/public/puct-info/agency/resources/pubs/news/2024/PUCT_Sees_Strong_Industry_Response_to_Texas_Energy_Fund_Loan_Program.pdf

[14] "Approval of Blueprint for Wholesale Electric Market Design and Directives to ERCOT," Public Utility Commission of Texas (PUCT), Austin, TX, USA, [Online] 52373-336-1180125, Jan. 13 2022. [Online]. Available: https://interchange.puc.texas.gov/Documents/52373_336_1180125.PDF

[15] Texas Admin. Code, Reliability and Continuity of Service, part 2, chap. 2(C), rule 25.52, Enacted 2021. [Online]. Available: https://www.puc.texas.gov/agency/rulesnlaws/subrules/electric/25.52/25.52.pdf

[16] "z-Score," in *Encyclopedia of Public Health*, W. Kirch Ed. Dordrecht: Springer Netherlands, 2008, pp. 1484-1484.


## IX. Biographies



**Sarat Chandra Vegunta** (M'09, SM'16) received his BE in Power Electronics (2002), Nagpur University, India, MSc in Electrical Power Engineering (2004), UMIST, UK, and PhD in Electrical Energy & Power Systems (2008), The University of Manchester, UK.

His interests are in the fields of power systems and power quality. Over the years, he performed power system studies, analyses, engineering design tasks, and power quality-related troubleshooting investigations of electricity Transmission and Distribution (T&D), Distributed Energy Resource (DER), microgrid, and industrial power systems, leading to a successful integration and or continuous operation of those systems.

He has been with Siemens since 2022; as a Consulting Director, he leads a team that provides consultancy services related to T&D, industrial, DERs, and microgrid power systems' planning, protection, power quality, and control.



**Nelson Bacalao** (M'80) received his Engineering Degree from the Simon Bolívar University in Venezuela (1979), Master of Engineering from Rensselaer Polytechnic Institute (1980), and PhD in Electrical Engineering from the University of British Columbia (1987). His interests include integrated system planning, optimal capacity expansion and Transmission and Distribution (T&D) planning under uncertainty, including using probabilistic methods. He has been with Siemens since 2006, and he is a Principal Consultant.



**Giovani Sasso Scarpati** (M'13) received his Bachelor of Electrical Engineering from Estácio University, Brazil (2012) and a Specialization (master's equivalent) in Power System Protection from Paulista University, Brazil (2022).

His interests are in the fields of power system protection and dynamics. He performed numerous protection coordination, protective relay settings calculation, and arc-flash studies, power system dynamic studies (including tuning of power system stabilizers and automatic voltage regulators for conventional generators), and onsite tasks commissioning protective devices and their settings.

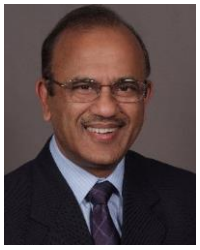
He joined Siemens in 2023 as a Senior Power System Consultant. Here, he focuses on and performs power flow, short-circuit, black-start, protection coordination, arc-flash, and stability studies for transmission, distribution, and industrial power systems and conventional and renewable generation systems.



**Azadeh Shadman** (M'13) received her BE in Electrical & Telecommunications Engineering (2003), Islamic Azad University, Iran, MSc in Electrical Power Engineering (2015), University of Wisconsin-Milwaukee, USA.

Her interests are in the fields of power systems, protection coordination, arc-flash, and Electromagnetic Transient (EMT) analyses and integration of Inverter-Based Resources (IBRs) into utility transmission and distribution power systems. She performed numerous power system studies in these fields for renewable-energy-based generation, industrial, and utility power systems in the USA.

She has been with Siemens since 2023 as a Power System Consultant. Here, with a focus on distribution systems, she has been performing studies ranging from power flow, short circuit, voltage flicker, voltage stability, EMT, harmonics, protection, energy losses to AC/DC arc-flash phenomena.



**Ram Nath** (M'98, SM'99) received his B.Sc. Degree in Mechanical Engineering (1974), Delhi College of Engineering, Delhi University, India, PhD in Power Systems (1985), Indian Institute of Technology, Delhi, India.

His interests are in the field of power systems and electromagnetic transient studies. His interests are in the transmission integration of large-scale renewable generation projects, especially offshore wind and solar PV, sub-synchronous resonance, control interaction, ferroresonance, and analysis of other EMT phenomenon. He has been with Siemens since 2008, and he is a Principal Consultant.

# THYMES: A Framework for Detecting Suicidal Ideation from Social Media Posts Using Hyperbolic Learning

Surendrabikram Thapa
*Virginia Tech, USA*

Mohammad Salman
*University of Stavanger, Norway*

Siddhant Bikram Shah
*Northeastern University, USA*

Shuvam Shiwakoti
*Delhi Technological University, India*

Qi Zhang
*Tongji University, China*

Liang Hu
*Tongji University, China*

Imran Razzak
*UNSW, Australia*

Usman Naseem
*Macquarie University, Australia*

2024 IEEE International Conference on Big Data (BigData) | 979-8-3503-6248-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/BigData62323.2024.10825881

***Abstract*—Mental health concerns are a critical issue in today's digital age, posing a threat to both individual and societal well-being and making the identification of at-risk individuals crucial. Analyzing an individual's social media post history can offer insights into their mental health state and help identify the presence of suicidal ideation. However, the complexity of linguistic and temporal data, along with sparsity and time irregularities, poses a formidable challenge in machine learning. Previous methods in this domain either rely on Euclidean space for processing which does not adequately model the power-law properties of social media posts, or lose information due to the discretization of the time axis. To address these challenges, we propose a novel framework, THYMES, which leverages pre-trained encoders and a rich representation learning paradigm with hyperbolic learning to model power-law features for enhanced sequence modeling. We perform experiments on two datasets and demonstrate that THYMES outperforms previously proposed methods while maintaining classification fairness under heavy data imbalances. Additionally, we qualitatively analyze commonly misclassified samples to reveal the shortcomings of models in this domain.**

***Index Terms*—Mental Health, Social Media, NLP**

## I. Introduction

The critical importance of addressing mental health concerns cannot be overstated, as they profoundly impact individual well-being and societal health. According to the World Health Organization, around one in eight people globally experience mental health disorders [24]. Depression and anxiety disorders are among the most common, often resulting in serious and harmful effects. While seeking professional help is the most effective way to address mental health issues, approximately 80% of individuals suffering from mental health problems do not receive the necessary mental health assistance [20]. This highlights an urgent need for alternative methods to surveil public mental health.

The proliferation of social media platforms has created unprecedented opportunities for understanding human behavior and mental health [22]. While a significant number of at-risk individuals do not disclose their condition or seek help [25], many individuals turn to social media to voice their grievances [4, 38]. User-generated data on social media, such as tweets, posts, and comments, offer a rich source of information that can be leveraged to detect signs of mental health issues and potentially help in the prevention of more severe outcomes like as self-harm and suicide. Suicide is one of the leading causes of death globally [25], underscoring it as a critical challenge to public health. Therefore, the early detection of suicidal intentions through social media content can significantly enhance timely interventions and potentially save lives.

With the increasing amount of social media content generated, computational methods for sentiment analysis and emotion recognition have seen significant growth [8, 26, 35]. Further, natural language processing (NLP) techniques have proven to be valuable tools in assessing users' mental health [18, 30, 31]. While generalized models such as the Bidirectional Encoder Representations from Transformers (BERT) family of transformers [9] have performed well on tasks such as sentiment analysis and emotion recognition [12, 16, 17, 39] upon fine-tuning, factors such as user diversity and the need for contextual and linguistic cues make reliable detection of mental health issues a challenging task. To overcome these challenges, STATENet [28] incorporates emotional and temporal contextual cues for suicide risk assessment. The time-aware transformer-based model focuses specifically on detecting suicidal intent in English tweets by integrating historical context into linguistic models. However, the model's capabilities fall short when it comes to effectively capturing and modeling temporal patterns or irregularities in social media data.

The detection of suicidal ideation on social media platforms poses a complex challenge in big data analytics due to its alignment with the **5Vs** of big data. The **Volume** of user-generated content is massive, making manual analysis impractical. The **Velocity** of social media posts necessitates real-time analysis due to the topic's sensitivity. The **Variety** of textual data, including acronyms, slang, and sarcasm,



Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:05:21 UTC from IEEE Xplore. Restrictions apply.

necessitates robust NLP methods. The **Value** lies in the opportunity to save lives by identifying at-risk individuals. **Veracity** is a challenge, as social media posts can be ambiguous or misleading, requiring advanced methods with near-human precision for making critical decisions. Addressing these 5Vs underscores the need for big data solutions capable of handling large datasets, processing information rapidly, analyzing diverse content, extracting valuable information, and ensuring data accuracy.

Building upon these motivations, we propose THYMES (**T**emporal **HY**perbolic **M**od**E**ling for **S**uicide ideation detection), a framework to identify users with suicidal ideation based on their post history on social media. Specifically, we utilize the pre-trained BERT [9] and Plutchik EmoNet [28] models to capture the linguistic features of each post and encapsulate them into a compact intermediate representation. We utilize Continuous Value Embeddings (CVE) [34] to prevent the loss of fine-grained temporal information when the time-axis is discretized. We represent the classification task in the hyperbolic space by integrating HYPHEN [1] with T-LSTM2 [41] to enhance the sequence modeling capabilities of THYMES. We perform experiments on two real-world suicide intent detection datasets and find that THYMES outperforms previously proposed methods in this domain. We qualitatively analyze commonly misclassified samples to help interpret the decision-making process of THYMES. We further discuss the limitations of the proposed framework in practical scenarios. Through THYMES, we endeavor to ensure that high-risk individuals receive prompt attention and support while minimizing errors based on the irregularities prevalent in real-world data.

## II. Related Work

### A. Identifying Mental Health Problems on Social Media

Several studies have tried to assess users' mental health through social media data. Ji et al. [14] introduce MentalBERT and MentalRoBERTa, two pre-trained language models specifically designed for mental healthcare applications, trained on 13 million sentences from posts on Reddit focused on mental health topics. Zogan et al. [42] propose DepressionNet, a framework to addresses the inefficiencies in existing models by using a summarization strategy to select relevant content from user tweets and combining it with user behavior features in a CNN-GRU deep learning model. Similarly, advancements have also been made in detecting suicide ideation through social data. Tadesse et al. [33] propose a CNN-LSTM model for suicide ideation detection, showcasing how deep learning frameworks outperform traditional machine learning methods. Sawhney et al. [29] used a deep graph neural framework to incorporate personal emotional history and social interactions for suicide ideation detection. They use Hyperbolic Graph Convolution Networks (HGCN) combined with the Hawkes process to model users' historical emotional spectrum, which addresses limitations related to individual subjective differences. Similarly, Cao et al. [5] introduced a suicide-oriented knowledge graph integrated with deep neural networks. The personal knowledge graph is built using psychological insights to model the key risk factors for suicidal ideation. They further used an LSTM coupled with a two-layered attention mechanism and the knowledge graph for detection. Sawhney et al. [28] present the STATENet framework, a time-aware transformer-based model designed to assess suicidal risk by incorporating historical emotional context from users' tweets. The model combines linguistic features with temporal and emotional cues to improve detection accuracy.

### B. Enhancing Sequence Modeling Capabilities

Sequence modeling in NLP has seen substantial advancements over the past decade, largely driven by the development of deep learning architectures. While early architectures such as Recurrent Neural Networks (RNNs) [13] and Gated Recurrent Units (GRUs) [6] maintained contextual information as memory, they struggled with modeling long-term dependencies due to vanishing gradients. The introduction of the attention mechanism in Transformer models [36] marked a significant breakthrough in NLP; specifically, pre-trained transformers showed impressive performance on downstream tasks [37]. However, these models may fail to capture the intricacies of complex data due to the loss of information when discretizing the time-axis [34]. Temporal language sequences in domains such as mental health exhibit properties similar to a power-law distribution [27], and Hyperbolic learning outperforms traditional Euclidean learning in modeling such distributions. The HYPHEN architecture [1] was proposed to jointly learn time irregularities and power-law properties of temporal sequences. However, HYPHEN utilizes Long Short-Term Memory (LSTM) modules to learn hyperbolic spaces, which cannot make long-term predictions due to its singular time gate. To alleviate these problems, our proposed framework THYMES combines hyperbolic learning with the T-LSTM2 architecture to better understand time irregularities by predicting both the current and future values of the sequence. To overcome the problem of loss of information in sparse and irregular time-series data, we create CVE representations to ensure the preservation of fine-grained multivariate information.

## III. Methodology

In this section, we define the problem statement and describe the methodology of our proposed framework, THYMES. It first extracts contextual linguistic information as representations by using the BERT model. Then, the emotional context of a post's history is extracted by using the Plutchik EmoNet transformer. This emotional context is encoded into CVE representations, preserving fine-grained temporal information. We implement a hyperbolic layer to better capture the power-law distribution exhibited by longitudinal mental health data, and concatenate the representations obtained from BERT and the hyperbolic layer before classification. A schematic diagram of THYMES is presented in Table 1.

### A. Problem Definition

We aim to identify the presence of suicidal intent within a social media user based on their post history. For each


Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:05:21 UTC from IEEE Xplore. Restrictions apply.

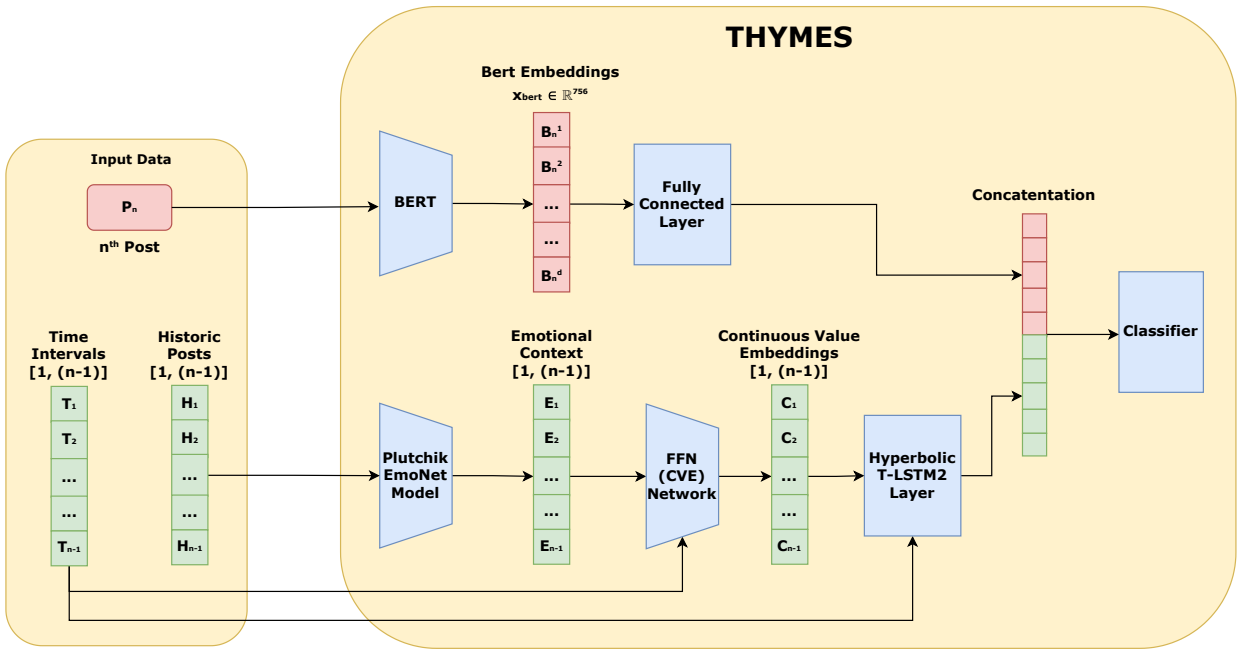



Fig. 1. The overall workflow of our proposed framework, THYMES. We use pre-trained BERT and Plutchik EmoNet transformers to convert the input post and historic posts into their respective representations. We then use an FFN model to convert the emotional context model into CVE embeddings. These embeddings are passed to a hyperbolic T-LSTM2 layer and concatenated with the projected BERT embeddings for classification.

user $u$, the posts authored by the user can be represented as $p_i \in P = \{p_1, p_2, ...p_n\}$ posted at time $t_i \in T = \{t_1, t_2, ...t_n\}$, where $t_{n-1} > t_{n-2} > ... > t_1$. For post $p_i$, all its preceding posts from $p_1$ to $p_{n-1}$ form its post history $h_i = \{(p_1, t_1), (p_2, t_2), ...(p_{n-1}, t_{n-1})\}$. The classification problem is to predict whether the user $u$ exhibits suicidal tendencies.

### B. Pre-trained Embedding Layers

Language transformers such as BERT [9] and Longformer [3] demonstrate superior performance as representation learners compared to static models such as Word2Vec [21] as they can encapsulate the contextual information of each text into a set of rich linguistic features. Accordingly, we implement the pre-trained BERT transformer as a text encoder for THYMES. For each text sample $t_i$, we extract the last hidden state of the transformer to obtain BERT embeddings $x_{bert} \in \mathbb{R}^{768}$.

$$x_{bert} = BERT(t_i) \tag{1}$$

We further add a dense layer to project the embedding to $x_{bert} \in \mathbb{R}^d$ with dimension d.

The emotional context of a user's post history can be instrumental in diagnosing depression and suicidal tendencies [23]. Parallel to the BERT embedding, we model the emotional context of each historic post $h_i$ of a user by using the Pluchtik EmoNet Transformer to obtain an emotion vector $x_{emo} \in \mathbb{R}^{768}$.

$$x_{emo} = EmoNet(h_i) \tag{2}$$

### C. Preventing Loss of Fine-Grained Information

Multivariate time-series data often suffer from sparsity and irregular time intervals, which may cause the loss of fine-grained information when the time-axis is discretized. To address this challenge, we incorporate Continuous Value Embedding (CVE) into THYMES to create parameterized embeddings for continuous time intervals and measured values, ensuring the preservation of fine-grained multivariate information. We treat an input sequence as a series of observation triplets consisting of time, feature, and value. Therefore, the sequence can be written as $T = \{(t_i, f_i, v_i)\}_{i=1}^{n}$. For each triplet $i$, we obtain its corresponding feature embedding $e_i^f$, time embedding $e_i^t$, and value embedding $e_i^v$ by passing the triplet through a one-to-many feed-forward network (FFN).

$$e_i^f, e_i^t, e_i^v = FFN(t_i, f_i, v_i) \tag{3}$$

We obtain the final initial embedding $x_{cve}$ for each triplet $i$ by summing its embeddings $e_i^f$, $e_i^t$, and $e_i^v$.

$$x_{cve} = e_i^f + e_i^t + e_i^v \tag{4}$$

By incorporating CVE into THYMES, we aim to enhance the expressiveness of the representations with respect to the time-axis.

### D. Hyperbolic Learning for Time-Sequences

Longitudinal analyses in sensitive domains such as mental health often exhibit the characteristics of a power-law distribution; a small number of instances can disproportionately affect the overall prediction. Hyperbolic learning has proven


Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:05:21 UTC from IEEE Xplore. Restrictions apply.

to be more effective in modeling power-law dynamics than Euclidean learning across various domains [32]. Accordingly, we integrate HYPHEN into THYMES as a hyperbolic layer for its ability to count time irregularities in scale-free sequences and use multi-hyperbolic space to encode varying levels of hyperbolic dynamics in sequences. As the hyperbolic space is a non-Euclidean space with a constant negative curvature $c$, we aim to learn $c$, which controls the degree of hyperbolic properties represented by the space.

The Time-aware LSTM (T-LSTM) [2] in HYPHEN exhibits excellent performance in longitudinal tasks by incorporating the time differences between each post for representation learning. However, it falls short of modeling long-term states due to its singular time gate. To alleviate this problem, we modify the HYPHEN architecture and integrate it with T-LSTM2, allowing the model to understand time irregularities between posts better. T-LSTM2 consists of two time-gates: $t1_k$, which exploits historical time information for current prediction, and $t2_k$, which stores time interval information to model long-term outcomes. The operations corresponding to this architecture can be defined as follows:

$$
\begin{aligned}
i_k &= \sigma \log_o \left(W^i \otimes h_{k-1} \oplus U^i \otimes m_k \oplus b^i\right) \\
f_k &= \sigma \log_o \left(W^f \otimes h_{k-1} \oplus U^f \otimes m_k \oplus b^f\right) \\
o_k &= \sigma \log_o \left(W^o \otimes h_{k-1} \oplus U^o \otimes m_k \oplus b^o\right)
\end{aligned}
$$

Here, $f_k$, $i_k$, and $o_k$ represent the input, forget, and output gates respectively. For timestep k, $m_k$ represents the input and $U^i$ represents the weight matrices for gate $i$. For timestep $k-1$, $h_{k-1}$ represents the hidden state and $W^i$ represents the weight matrices for gate $i$. $b^i$ represents the bias for gate $i$. In this context, the symbols $\otimes$, $\oplus$, and $\odot$ represent the Mobius Addition, Mobius Matrix Multiplication, and Mobius Pointwise Multiplication operations respectively. Formally, $t1_k$ and $t2_k$ are defined as:

$$
\begin{aligned}
t1_k &= \sigma \log_o \left(W_{t1} \otimes m_k + \sigma \left(u^t \otimes g\left(\Delta t_k\right)\right) \oplus b_{t1}\right) \\
t2_k &= \sigma \log_o \left(W_{t2} \otimes m_k + \sigma \left(u^t \otimes g\left(\Delta t_k\right)\right) \oplus b_{t2}\right)
\end{aligned}
$$

where $u^t$ is the weight vector corresponding to the time decay effect. The time-lapses in the hyperbolic layer can be defined as:

$$
\begin{aligned}
c^s_{k-1} &= \exp_o \left(\tanh \left(\log_o \left(W^d \otimes c_{k-1} \oplus b^d\right)\right)\right) \\
\tilde{c}_k &= \sigma \log_o \left(W^c \otimes h_{k-1} \oplus U^c \otimes m_k \oplus b^c\right) \\
\bar{c}_k &= f_k \odot c^s_{k-1} \oplus i_k \odot t1_k \odot \tilde{c}_k \\
c_k &= f_k \odot c^s_{k-1} \oplus i_k \odot t2_k \odot \tilde{c}_k
\end{aligned}
$$

where $c^s_{k-1}$, $\tilde{c}_k$, $\bar{c}_k$, and $c_k$ represent the memory cell's state at timestep $k-1$, candidate state at timestep $k$, temporary state at timestep $k$, and the final state at timestep $k$ respectively. The function $g(\cdot)$ denotes a heuristic decaying function. In our case, we set $g(\Delta k) = 1/\Delta k$. Finally, integrating T-LSTM2 with the hyperbolic layer, we get:

$$
\begin{aligned}
h_k &= o_k \odot \exp_o \left(\tanh \left(\bar{c}_k\right)\right) \\
h_j &= HTTN\left(m_j, \Delta_j, h_{j-1}\right), \quad j \in [1, T]
\end{aligned}
$$

where $h_k$ represents the hidden state at timestep $k$ and HTTN represents the Hyperbolic Time Aware Temporal Network. We obtain the embedding $x_{HYP}$ from the hyperbolic layer $H_{T-LSTM2}$.

$$
x_{hyp} = H_{T-LSTM2}(x_{cve}) \tag{5}
$$

### E. Training Objective

We concatenate $x_{BERT}$ and $x_{HYP}$ to obtain a final embedding $x_{final}$ used for classification. We add a linear layer at the end of the THYMES as a classifier to map the representations to the number of classes (two in our case). To make THYMES more robust to the data imbalances inherent in real-world scenarios, we employ the Class-Balanced loss function [7]. This loss function assigns a weight to each class that is inversely proportional to the number of samples in the class.

## IV. Experimental Setup

### A. Datasets

We conducted experiments on two benchmark publicly available datasets commonly used in previous similar studies [1, 28, 32]. Specifically, we used the Suicide and the eRisk dataset. Below, we briefly describle each dataset.

**Suicide.** The Suicide dataset [10] is composed of Reddit posts from 500 users exhibiting various mental health states, screened from an initial pool of 270,000 users across multiple subreddits focused on mental health. The posts covered topics such as depression, anxiety, alcoholism, and suicide, and were annotated into five risk levels by using the Columbia Suicide Severity Risk Scale (C-SSRS).

**eRisk.** The eRisk dataset [19] is a part of the 'Early Risk Prediction on The Internet' annual challenge organized by the Conference and Labs of the Evaluation Forum (CLEF). It consists of a user's posts on social media in chronological order. The posts covered topics such as depression and self-harm.

### B. Baselines

**STATENet.** STATENet [28] is a model proposed to identify the suicidal intent of a social media user by processing their post history. The model uses a time-aware LSTM (T-LSTM) to facilitate the early screening of suicidal intent on social media. We use STATENet model as the backbone for each of the other architectures except the RISE model.

**Time2Vec.** Time2Vec [15] is used as a trainable vector representation learner to encode time features instead of utilizing traditional time features.

**T-LSTM2.** The T-LSTM in STATENet is replaced with T-LSTM2 [41] for enhanced sequence modeling capabilities owing to its two time-gates.

**T-LSTM3.** The T-LSTM3 model [41] is incorporated into STATENet. The architecture uses coupled input and forget gates to reduce the number of parameters.

**RxLSTM.** The Rx-refill LSTM (RxLSTM) model [40] is used in place of STATENet's LSTM layer. The architecture uses a graph neural network to encode heterogenous sequences.


Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:05:21 UTC from IEEE Xplore. Restrictions apply.

TABLE I

CLASSIFICATION PERFORMANCE OF THE MODELS ON THE ERISK DATASET. STATENET IS USED AS THE BACKBONE MODEL FOR ALL BASELINES EXCEPT RISE. PERFORMANCE IS REPORTED ACROSS FOUR EVALUATION METRICS: ACCURACY, MACRO PRECISION, MACRO RECALL, AND MACRO F1-SCORE. THE MEAN PERFORMANCE OF THE MODELS ACROSS FIVE-FOLD VALIDATION IS REPORTED. THE BEST PERFORMANCE IS HIGHLIGHTED IN **BOLD**.

| Method | Accuracy | Precision | Recall | F1 |
|---|---|---|---|---|
| STATENet | 82.17 | 58.93 | 56.61 | 57.37 |
| Time2Vec | 79.57 | 56.67 | 56.43 | 56.54 |
| T-LSTM2 | 82.56 | 55.81 | 55.92 | 55.82 |
| T-LSTM3 | 83.45 | 60.18 | 56.02 | 56.98 |
| RxLSTM | 82.56 | 59.34 | 56.54 | 57.38 |
| BiLSTM | 79.17 | 53.92 | 55.83 | 53.90 |
| CVE | 82.27 | 55.68 | 58.15 | 56.07 |
| HYPHEN | **86.51** | 43.26 | 50.00 | 46.38 |
| RISE | 81.33 | 56.14 | 58.77 | 57.18 |
| Ours | 77.29 | **61.00** | **66.10** | **62.25** |

**BiLSTM.** The Bi-directional LSTM model [11] processes sequences in both forward and backward directions, capturing information from both past and future contexts.

**CVE.** The CVE paradigm [34] along with a one-to-many FFN is integrated with STATENet to create embeddings for continuous times and measured values, ensuring the preservation of fine-grained information.

**HYPHEN.** The HYPHEN architecture [1] is used to replace the T-LSTM in STATENet to formulate the classification problem in the hyperbolic space instead of the Euclidean space.

**RISE.** Robust Early-exiting Internal Classifiers (RISE) model [32] is a recently proposed hyperbolic framework that consists of multiple early-exiting classifiers.

### C. Implementation Details

We conducted our experiments with the HYPHEN models on an NVIDIA GeForce RTX 3060 GPU with 12 GB of dedicated memory. All other experiments were conducted on the NVIDIA A100 GPU with 80 GB of dedicated memory by using Google Colab[1]. We perform 5-fold validation and report the mean performance across all runs for each metric. We empirically find the best learning rate, batch size, and number of epochs for each model.

## V. RESULTS AND ANALYSIS

### A. Experimental Results

Table I presents the classification performance of the methods on the eRisk dataset. The default STATENet method provides a strong baseline for comparison. Adding different LSTM methods does not alter performance in a major way. THYMES generally performs competitively against contemporary state-of-the-art methods. While THYMES is surpassed in terms of accuracy by various methods, including STATENet, the eRisk dataset harbors a class imbalance ratio of over 85% across just two classes; therefore, accuracy does not necessarily signify a better method as these methods may be overfitting on the majority class. This can be observed with the results for the HYPHEN model, where the accuracy is the highest amongst all other methods. In contrast, all the other metrics suggest that the model is overfitting. On the other hand, the precision, recall, and F1 scores of THYMES are the highest among all the methods, which suggests that it has fewer false positives and negatives than the other methods. The comparative results between HYPHEN, RISE, and THYMES further elucidate the prominence of our enhanced hyperbolic learning layer over other hyperbolic frameworks.

[1]https://colab.research.google.com/

TABLE II

CLASSIFICATION PERFORMANCE OF THE MODELS ON THE SUICIDE DATASET. STATENET IS USED AS THE BACKBONE MODEL FOR ALL BASELINES EXCEPT RISE. PERFORMANCE IS REPORTED ACROSS THREE EVALUATION METRICS: ACCURACY, MACRO F1-SCORE, AND WEIGHTED F1-SCORE. THE MEAN PERFORMANCE OF THE MODELS ACROSS FIVE-FOLD VALIDATION IS REPORTED. THE BEST PERFORMANCE IS HIGHLIGHTED IN **BOLD**.

| Method | Accuracy | Macro F1 | Weighted F1 |
|---|---|---|---|
| STATENet | 87.64 | 71.93 | 88.11 |
| T-LSTM2 | 70.31 | 57.65 | 77.19 |
| CVE | 85.29 | 62.09 | 86.16 |
| HYPHEN | **92.96** | 48.17 | 89.58 |
| RISE | 88.29 | 64.97 | 85.06 |
| Ours | 90.63 | **77.34** | **90.43** |

The classification performance of THYMES and the baseline methods on the Suicide dataset are presented in Table II. The Suicide dataset carries a class imbalance ratio of over 90%, exceeding the eRisk dataset's imbalance. The HYPHEN model achieves the highest accuracy again but suffers from a low Macro F1-Score, suggesting that the model does not have a fair decision boundary for classification across the two classes. Adding the T-LSTM2 layer or generating CVE representations causes a marked decrease in performance, suggesting that altering these modules deteriorates the ability of the STATENet backbone to process time and language features, respectively. Unlike the other methods, the RISE framework performs similarly to THYMES and avoids overfitting. However, THYMES performs the best overall in both Macro F1-Score and Weighted F1-Score, suggesting that the combination of hyperbolic learning and T-LSTM2 allows THYMES to better understand temporal and linguistic data.

### B. Ablation Study

**Component wise Analysis.** We ablate each component of THYMES to understand the contribution of each module towards its performance. The results of the ablation study are presented in Table III. We start with the baseline pre-trained encoders and gradually integrate each module of THYMES. The pre-trained BERT and Plutchik EmoNet models act as strong representation learners and provide a good foundation for THYMES. Adding the hyperbolic layer initially causes the performance to degrade due to overfitting; however, it paves



Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:05:21 UTC from IEEE Xplore. Restrictions apply.

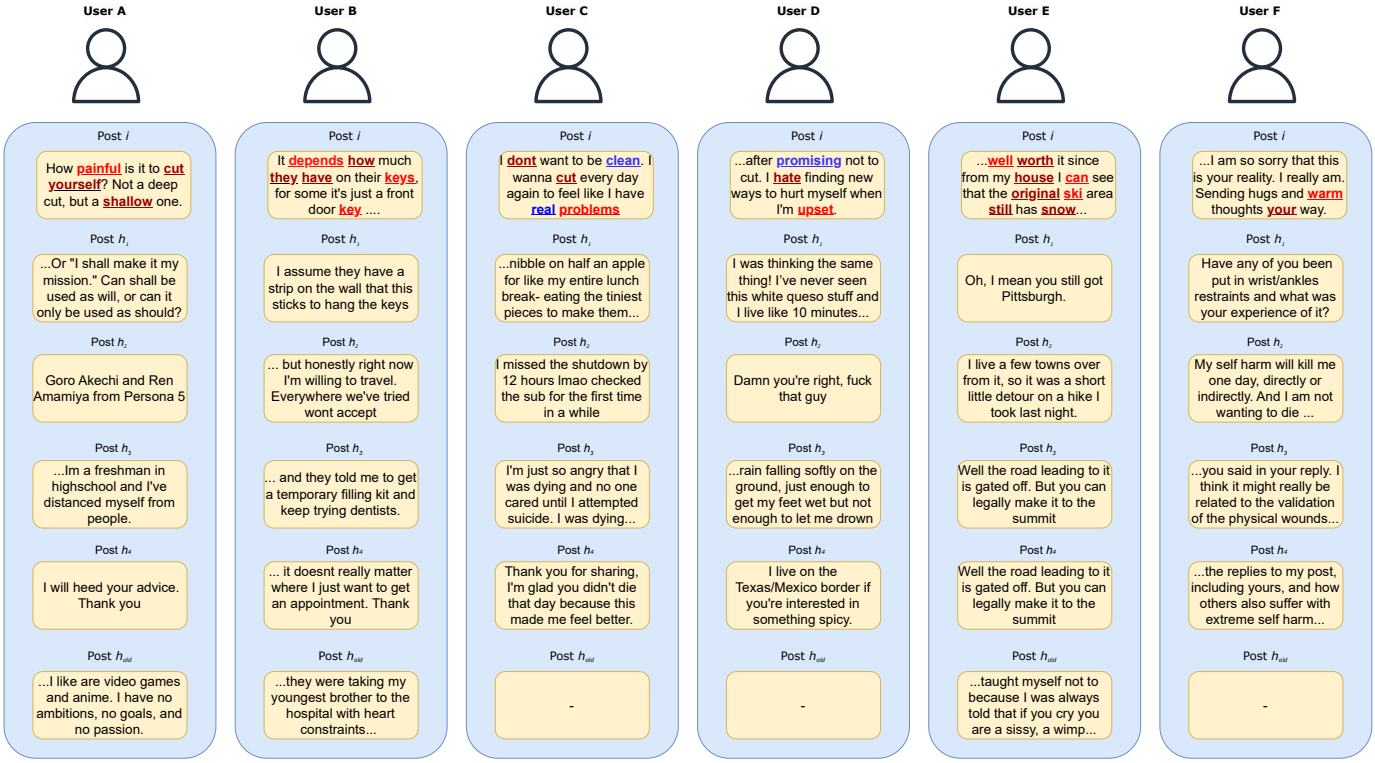



Fig. 2. Samples from 6 social media users taken from the eRisk dataset. We analyze the topmost post $i$ and use all previous posts as the user's post history. If the user has no more posts before post $h_4$, $h_{old}$ is shown as blank.

the way for better representations that model the classification task. Converting the embeddings to CVE helps the model encapsulate more information in its representations, causing an increase in performance. T-LSTM2 completes THYMES and causes a further increase in performance due to enhanced sequence modeling capabilities over the default T-LSTM.

**Comparison Between Model Variants.** We compare Hyperbolic and Euclidean variants of THYMES to derive the advantages of Hyperbolic learning over Euclidean learning for suicidal intent detection. Further, we measure the impact of using Hawkes attention over the default (Vanilla) attention on THYMES. The results of this comparison are presented in Table IV. We find that the Hyperbolic models outperform the Euclidean models across all metrics in both datasets, strengthening our argument that the linguistic features in the mental health domain are more effectively represented in a Hyperbolic space over an Euclidean space. We observe that Hawkes's attention surpasses Vanilla's attention on both Euclidean and Hyperbolic models across most metrics except accuracy for the eRisk dataset; thus, the Vanilla attention model shows signs of overfitting. This is further elucidated by the results where the addition of Vanilla attention to THYMES degrades the model in terms of Macro F1-Score.

TABLE III
CLASSIFICATION PERFORMANCE OF EACH COMPONENT OF THYMES ON THE ERISK DATASET. PE SIGNIFIES THE PRE-TRAINED ENCODERS AND HYP SIGNIFIES THE HYPERBOLIC LAYER RESPECTIVELY. THE LAST ROW PRESENTS THE COMPLETE THYMES MODEL. PERFORMANCE IS REPORTED ACROSS FOUR EVALUATION METRICS: ACCURACY, MACRO PRECISION, MACRO RECALL, AND MACRO F1-SCORE. THE MEAN PERFORMANCE OF THE MODELS ACROSS FIVE-FOLD VALIDATION IS REPORTED. THE BEST PERFORMANCE IS HIGHLIGHTED IN **BOLD**.

| PE | HYP | CVE | T-LSTM2 | Acc. | Prec. | Rec. | F1 |
|---|---|---|---|---|---|---|---|
| ✓ | | | | 82.17 | 58.93 | 56.61 | 57.37 |
| ✓ | ✓ | | | **86.51** | 43.26 | 50.00 | 46.38 |
| ✓ | ✓ | ✓ | | 80.78 | 60.68 | 61.72 | 61.14 |
| ✓ | ✓ | ✓ | ✓ | 77.34 | **61.00** | **66.10** | **62.25** |

**Qualitative Analysis.** We conduct a qualitative analysis of misclassified samples from the eRisk dataset to help interpret how the model makes decisions based on the content of a user's post. We choose 6 random users and analyze the last 5 of their posts. These samples are presented in Figure 2. The model predictions are presented in Table V.

In User A's post, which exhibited suicidal intention and was classified as sadness, none of the models predicted correctly due to the absence of historical suicidal tweets. The most negative post for this user is shown in $h^A_{old}$. User B's tweet, a regular post predominantly conveying joy, stood out among other emotional categories. The last 5 tweets indicated no signs of depression, although $h^B_{old}$ revealed profound sadness in earlier posts. User C's post displayed suicidal intent, and their historical tweets suggested the same. This user was only classified correctly by


Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:05:21 UTC from IEEE Xplore. Restrictions apply.

TABLE IV
COMPARISON OF THE CLASSIFICATION PERFORMANCE OF EUCLIDEAN AND HYPERBOLIC THYMES VARIANTS ON THE ERISK AND SUICIDE DATASETS WITH DIFFERENT TYPES OF ATTENTION. PERFORMANCE IS REPORTED ACROSS THREE EVALUATION METRICS: ACCURACY, MACRO F1-SCORE, AND WEIGHTED F1-SCORE. THE MEAN PERFORMANCE OF THE MODELS ACROSS FIVE-FOLD VALIDATION IS REPORTED. THE BEST PERFORMANCE IS HIGHLIGHTED IN **BOLD**.

| Model Type | Attention | eRisk Dataset | | | Suicide Dataset | | |
|---|---|---|---|---|---|---|---|
| | | Accuracy | Macro F1 | Weighted F1 | Accuracy | Macro F1 | Weighted F1 |
| Euclidean | None | 81.46 | 57.62 | 78.33 | 70.31 | 57.65 | 77.19 |
| | Vanilla | 76.88 | 57.41 | 78.07 | 70.31 | 57.65 | 77.19 |
| | Hawkes | 76.41 | 61.97 | 78.96 | 73.44 | 58.97 | 79.53 |
| Hyperbolic | None | 75.94 | 60.11 | 78.73 | 78.91 | 63.60 | 83.52 |
| | Vanilla | **79.06** | 57.85 | **79.40** | 78.91 | 60.79 | 83.28 |
| | Hawkes | 77.29 | **62.25** | **79.40** | **90.63** | **77.43** | **90.43** |

TABLE V
RESULTS BY DIFFERENT METHODS ON 6 SELECTED SAMPLES FROM THE ERISK DATASET FOR THE QUALITATIVE ANALYSIS.

| User | STATENet | RxLSTM | Time2Vec | BiLSTM | T-LSTM3 | THYMES |
|---|---|---|---|---|---|---|
| *A* | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| *B* | ✓ | ✗ | ✗ | ✗ | ✓ | ✓ |
| *C* | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ |
| *D* | ✗ | ✗ | ✗ | ✗ | ✓ | ✗ |
| *E* | ✓ | ✗ | ✗ | ✓ | ✗ | ✗ |
| *F* | ✗ | ✗ | ✓ | ✗ | ✓ | ✓ |

our model due to it leveraging fine-grained timing irregularities and hyperbolic attention. User D's post, categorized as anger with indications of self-injury, surprisingly had no clear negative sentiment in its history. T-LSTM3 was the only model that could classify this user correctly. User E's post radiated joy, and the last 5 historical posts appeared normal, although earlier posts in $h^E_{old}$ contained negative emotions. Three models correctly predicted this user. User F's tweet expressed joy despite nearly all of the 5 most recent tweets hinting at severe suicidal intent within a short timeframe. Three different models correctly classified this sample.

### C. Discussion and Future Work

THYMES works in the hyperbolic space, a non-Euclidean space with distinct geometrical properties. To leverage these unique characteristics and the structure of hyperbolic geometry, we used the Möbius transformation. While the Möbius transformations and hyperbolic isometries have representations similar to matrices, they operate in different mathematical settings. Möbius transformations are defined on the complex plane and primarily focus on conformal mappings, while hyperbolic isometries operate on hyperbolic space and preserve the hyperbolic metric. In the future, alternative approaches such as Fractional linear transformations and Riemannian geometry can be studied to enhance further our understanding of neural modeling capabilities within the hyperbolic space. Further, a major issue in this domain is the data imbalances that exist in the datasets that are used to train models such as THYMES. We do not employ data-centric techniques to attenuate class imbalances for a fair comparison with previous methods. However, in the future, the performance of suicide ideation detection frameworks can greatly be increased with techniques such as data augmentation with GANs.

## VI. CONCLUSION

In this work, we propose THYMES, a framework leveraging BERT and Plutchik EmoNet models for extracting linguistic and affective features, respectively, Continuous Value Embeddings for preserving temporal information, and hyperbolic learning with T-LSTM2 for enhanced sequence modeling. Our experiments on two real-world datasets demonstrate THYMES's superior performance in suicide intent detection, highlighting its potential to provide timely support to high-risk individuals. While THYMES improves upon the previously proposed frameworks in this domain, we acknowledge that the diverse, intricate, and ever-changing nature of social media may present challenges to the accurate identification of suicide ideation. Further, due to the extremely sensitive nature of datasets in this domain, they are limited by their size and remain inadequate for training large deep-learning models. We present THYMES as a step towards addressing common real-world problems in social media, helping to better diagnose at-risk individuals in the digital landscape.

## ETHICAL CONSIDERATIONS

**User Privacy:** As the datasets used to train THYMES contained user data, we preprocessed the data to ensure that all user identifiers, such as usernames and tags, were absent.

**Broader Impact:** We are aware of the potential impact of our work on individuals and communities. Our primary goal is to provide a tool that can aid in early detection and intervention, ultimately contributing to suicide prevention efforts.

**Biases:** We acknowledge that due to the subjective nature of the topic and the relatively small sizes of the datasets used to train THYMES, the model may exhibit biased decision-making. All such biases are unintentional, and we ask readers to be aware of them.

## LIMITATIONS

Due to the sensitive nature of suicide ideation detection, we acknowledge that model limitations can lead to serious consequences. While we believe that THYMES is a promising framework, it has limitations that can be addressed in future work. Scaling hyperbolic learning methods to larger datasets


Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:05:21 UTC from IEEE Xplore. Restrictions apply.

or higher dimensional spaces can be challenging. As the dimensionality increases, the computational cost and memory requirements can grow significantly, potentially limiting the applicability of THYMES to large-scale users with many posts. Further, while the model may help in the early identification of at-risk users, it may be prone to false predictions, and a human-in-the-loop system may be more practical.

## REFERENCES


[1] Shivam Agarwal, Ramit Sawhney, Sanchit Ahuja, Ritesh Soun, and Sudheer Chava. Hyphen: Hyperbolic hawkes attention for text streams. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 620–627, 2022.

[2] Inci M Baytas, Cao Xiao, Xi Zhang, Fei Wang, Anil K Jain, and Jiayu Zhou. Patient subtyping via time-aware lstm networks. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 65–74, 2017.

[3] Iz Beltagy, Matthew E Peters, and Arman Cohan. Longformer: The long-document transformer. *arXiv preprint arXiv:2004.05150*, 2020.

[4] Luca Braghieri, Ro'ee Levy, and Alexey Makarin. Social media and mental health. *American Economic Review*, 112(11):3660–3693, 2022.

[5] Lei Cao, Huijun Zhang, and Ling Feng. Building and using personal knowledge graph to improve suicidal ideation detection on social media. *IEEE Transactions on Multimedia*, 24:87–102, 2020.

[6] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, 2014.

[7] Yin Cui, Menglin Jia, Tsung-Yi Lin, Yang Song, and Serge Belongie. Class-balanced loss based on effective number of samples. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9268–9277, 2019.

[8] Xiang Deng, Vasilisa Bashlovkina, Feng Han, Simon Baumgartner, and Michael Bendersky. Llms to the moon? reddit market sentiment analysis with large language models. In *Companion Proceedings of the ACM Web Conference 2023*, pages 1014–1019, 2023.

[9] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

[10] Manas Gaur, Amanuel Alambo, Joy Prakash Sain, Ugur Kursuncu, Krishnaprasad Thirunarayan, Ramakanth Kavuluru, Amit Sheth, Randy Welton, and Jyotishman Pathak. Knowledge-aware assessment of severity of suicide risk for early intervention. In *The world wide web conference*, pages 514–525, 2019.

[11] Alex Graves and Jürgen Schmidhuber. Framewise phoneme classification with bidirectional lstm and other neural network architectures. *Neural networks*, 18(5-6):602–610, 2005.

[12] Mickel Hoang, Oskar Alija Bihorac, and Jacobo Rouces. Aspect-based sentiment analysis using bert. In *Proceedings of the 22nd nordic conference on computational linguistics*, pages 187–196, 2019.

[13] John J Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, 79(8):2554–2558, 1982.

[14] Shaoxiong Ji, Tianlin Zhang, Luna Ansari, Jie Fu, Prayag Tiwari, and Erik Cambria. Mentalbert: Publicly available pretrained language models for mental healthcare. *arXiv preprint arXiv:2110.15621*, 2021.

[15] Seyed Mehran Kazemi, Rishab Goel, Sepehr Eghbali, Janahan Ramanan, Jaspreet Sahota, Sanjay Thakur, Stella Wu, Cathal Smyth, Pascal Poupart, and Marcus Brubaker. Time2vec: Learning a vector representation of time. *arXiv preprint arXiv:1907.05321*, 2019.

[16] Puneet Kumar and Balasubramanian Raman. A bert-based dual-channel explainable text emotion recognition system. *Neural Networks*, 150:392–407, 2022.

[17] Xin Li, Lidong Bing, Wenxuan Zhang, and Wai Lam. Exploiting bert for end-to-end aspect-based sentiment analysis. *arXiv preprint arXiv:1910.00883*, 2019.

[18] Chenhao Lin, Pengwei Hu, Hui Su, Shaochun Li, Jing Mei, Jie Zhou, and Henry Leung. Sensemood: depression detection on social media. In *Proceedings of the 2020 international conference on multimedia retrieval*, pages 407–411, 2020.

[19] David E Losada, Fabio Crestani, and Javier Parapar. Overview of erisk at clef 2019: Early risk prediction on the internet (extended overview). *CLEF (Working Notes)*, 2019.

[20] Catherine M McHugh, Amy Corderoy, Christopher James Ryan, Ian B Hickie, and Matthew Michael Large. Association between suicidal ideation and suicide: meta-analyses of odds ratios, sensitivity, specificity and positive predictive value. *BJPsych open*, 5(2):e18, 2019.

[21] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.

[22] Usman Naseem, Surendrabikram Thapa, Qi Zhang, Liang Hu, Junaid Rashid, and Mehwish Nasim. Incorporating historical information by disentangling hidden representations for mental health surveillance on social media. *Social Network Analysis and Mining*, 14(1):9, 2023.

[23] Andrada D Neacsiu, Caitlin M Fang, Marcus Rodriguez, and M Zachary Rosenthal. Suicidal behavior and problems with emotion regulation. *Suicide and Life-Threatening Behavior*, 48(1):52–74, 2018.

[24] World Health Organization. Mental disorders, 2024.

[25] World Health Organization. Suicide, 2024.

[26] Sara Rosenthal, Noura Farra, and Preslav Nakov. Semeval-




Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:05:21 UTC from IEEE Xplore. Restrictions apply.

2017 task 4: Sentiment analysis in twitter. *arXiv preprint arXiv:1912.00741*, 2019.

[27] Ramit Sawhney, Shivam Agarwal, Megh Thakkar, Arnav Wadhwa, and Rajiv Ratn Shah. Hyperbolic online time stream modeling. In *Proceedings of the 44th International ACM SIGIR conference on research and development in information retrieval*, pages 1682–1686, 2021.

[28] Ramit Sawhney, Harshit Joshi, Saumya Gandhi, and Rajiv Shah. A time-aware transformer based model for suicide ideation detection on social media. In *Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP)*, pages 7685–7697, 2020.

[29] Ramit Sawhney, Harshit Joshi, Rajiv Shah, and Lucie Flek. Suicide ideation detection via social and temporal user representations using hyperbolic learning. In *Proceedings of the 2021 conference of the North American Chapter of the Association for Computational Linguistics: human language technologies*, pages 2176–2190, 2021.

[30] Siddhant Bikram Shah, Shubham Garg, and Aikaterini Bourazeri. Emotion recognition in speech by multimodal analysis of audio and text. In *2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pages 257–263. IEEE, 2023.

[31] Guangyao Shen, Jia Jia, Liqiang Nie, Fuli Feng, Cunjun Zhang, Tianrui Hu, Tat-Seng Chua, Wenwu Zhu, et al. Depression detection via harvesting social media: A multimodal dictionary learning solution. In *IJCAI*, pages 3838–3844, 2017.

[32] Ritesh Singh Soun, Atula Tejaswi Neerkaje, Ramit Sawhney, Nikolaos Aletras, and Preslav Nakov. Rise: Robust early-exiting internal classifiers for suicide risk evaluation. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 14134–14145, 2024.

[33] Michael Mesfin Tadesse, Hongfei Lin, Bo Xu, and Liang Yang. Detection of suicide ideation in social media forums using deep learning. *Algorithms*, 13(1):7, 2019.

[34] Sindhu Tipirneni and Chandan K Reddy. Self-supervised transformer for sparse and irregularly sampled multivariate clinical time-series. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 16(6):1–17, 2022.

[35] Krishna B Vamshi, Ajeet Kumar Pandey, and Kumar AP Siva. Topic model based opinion mining and sentiment analysis. In *2018 International Conference on Computer Communication and Informatics (ICCCI)*, pages 1–4. IEEE, 2018.

[36] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

[37] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations*, pages 38–45, 2020.

[38] Akkapon Wongkoblap, Miguel A Vadillo, and Vasa Curcin. Researching mental health disorders in the era of social media: systematic review. *Journal of medical Internet research*, 19(6):e228, 2017.

[39] Hu Xu, Bing Liu, Lei Shu, and Philip S Yu. Bert post-training for review reading comprehension and aspect-based sentiment analysis. *arXiv preprint arXiv:1904.02232*, 2019.

[40] Jianfei Zhang, Ai-Te Kuo, Jianan Zhao, Qianlong Wen, Erin Winstanley, Chuxu Zhang, and Yanfang Ye. Rxnet: Rx-refill graph neural network for overprescribing detection. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, pages 2537–2546, 2021.

[41] Yu Zhu, Hao Li, Yikang Liao, Beidou Wang, Ziyu Guan, Haifeng Liu, and Deng Cai. What to do next: Modeling user behaviors by time-lstm. In *IJCAI*, volume 17, pages 3602–3608, 2017.

[42] Hamad Zogan, Imran Razzak, Shoaib Jameel, and Guandong Xu. Depressionnet: A novel summarization boosted deep framework for depression detection on social media. *arXiv preprint arXiv:2105.10878*, 2021.



Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:05:21 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for publication in IEEE Transactions on Cognitive and Developmental Systems. 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/TCDS.2025.3529669



# TRANSIT-EEG — A Framework for Cross-Subject Classification with Subject Specific Adaptation

Chirag Ahuja and Divyashikha Sethia

***Abstract*—Electroencephalography (EEG) is pivotal in monitoring and analysing cerebral activity across diverse domains, including medical diagnostics, cognitive neuroscience, and brain-computer interfaces. However, the inherent intricacy of EEG signals and their subject-specific characteristics pose formidable challenges in devising robust and generalisable classification models. Traditional EEG signal classification paradigms rely on extensive subject-specific datasets. Also, the domain adaption for new subjects often leads to "catastrophic forgetting," thereby diminishing the performance of model trained on prior subjects. This paper proposes a novel framework, Transfer and Robust Adaptation of New Subjects in EEG Technology (TRANSIT-EEG), designed to adapt adeptly to new subjects. TRANSIT-EEG demonstrates resilience to subject-specific artefacts by integrating synthetic data generation using the proposed subject-specific augmentation model - Individualised Diffusion Probabilistic Model (IDPM). Also, it employs a robust Self Organising Graph Attention Transformer (SOGAT) that dynamically constructs a graph for each subject, fostering a more accurate classification. Moreover, TRANSIT-EEG introduces adapter-based finetuning using Low-Rank Adaptation (LoRA) for new subjects, enriching the adaptation process. The TRANSIT-EEG framework presents a promising avenue for advancing the realm of EEG signal classification. Evaluation of widely studied datasets, specifically focusing on two significant tasks, SEED for Emotion Recognition and PhyAat for Auditory Activity Recognition, substantiates the efficacy and versatility of TRANSIT-EEG. This validation indicates a substantial stride toward achieving more generalisable and accurate EEG signal classification.**

***Index Terms*—Transfer Learning, Few-Shot Learning, Diffusion Model, LoRA, Graph Neural Network, Cross-Subject Transfer.**

## I. Introduction

Electroencephalography (EEG) plays an indispensable role in real-time monitoring and analysis of neural activities, finding applications across various domains such as medical diagnostics and Brain-Computer Interfaces (BCIs) [1; 2]. However, the path from raw EEG data to actionable insights presents formidable challenges. The intricate and individual-specific nature of neural signals poses a significant hurdle, making generalization a complex endeavor. Supervised models, both Machine Learning and Deep Learning, require large, subject-specific datasets, which often introduce challenges due to constraints in time, cost, and ethical considerations.

In response to these challenges, researchers explore cross-subject learning as a cornerstone in neuroscience, medical diagnostics, and BCIs. The primary objective is to construct robust models capable of generalizing across diverse subjects, giving rise to cross-subject modeling [3; 4]. Recent advancements in this field include the Self-Organised Graph Neural Network (SOGNN) [3], which attains high cross-subject accuracy on benchmark datasets, outperforming predecessors such as Regularized Graph Neural Networks (RGNN) [5] and the Bi-Hemispheric Discrepancy Model (BiHDM) [6]. These studies indicate that for multivariate signals like EEG, graphical models often outperform standard Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks.

While SOGNN lays a solid foundation for cross-subject modeling using Graph Neural Networks (GNNs), it still lacks in areas such as inter-channel relationship learning, which could potentially improve using techniques like Graph Attention Transformers (GAT) [7]. Moreover, existing cross-subject models focus primarily on robustness to new subjects but often neglect the potential benefits of adapting to individual subjects for customized performance. Some approaches, such as the cross-subject transfer in Deep Domain Adaptation [8], explore using Deep Neural Network finetuning strategies to adapt models for unseen subjects, but there remains significant room for improvement.

Deep learning models often overfit during training and fine-tuning, particularly when working with limited data. Researchers turn to data augmentation as a potential solution [9; 10; 11]. Recent advancements in EEG data augmentation leverage generative models to address the challenge of limited data in deep learning applications. Notable examples include GANSER [10], which utilizes Generative Adversarial Networks (GAN), and approaches employing Variational Autoencoders (VAE) for EEG data augmentation [9]. While these methods show promise in enhancing model performance, they often fall short in capturing subject-specific characteristics crucial for accurate EEG classification. The lack of subject-specific augmentation techniques remains a significant gap in the field, limiting the generalization capabilities of EEG classification models across different individuals.

The recent success of Denoising Diffusion Probabilistic Models (DDPMs) in image generation, surpassing the performance of GANs, indicates their potential superiority in EEG data generation [12]. Torma et.al [13] further corroborate this potential, demonstrating DDPMs' efficacy in synthesizing high-quality Event-Related Potentials (ERPs) and Motor Imagery (MI) signals. Their research reveals that DDPMs can generate EEG signals of sufficient quality to augment training data and enhance deep learning classification models. Building

C. Ahuja is with the Department of Computer Science and Engineering, Delhi Technology University, Delhi, India e-mail: (chiragahuja_2k20phdco13@dtu.ac.in).

D. Sethia is with Department of Software Engineering, Delhi Technology University, Delhi, India e-mail: (divyashikha@dtu.ac.in)

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:26:00 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Cognitive and Developmental Systems. 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/TCDS.2025.3529669



on these advancements, this paper introduces Individualized Diffusion Probabilistic Models (IDPM) for augmentation. IDPM enhances the DDPM approach by distinctly segmenting noise and subject-specific informative signals during the denoising process, thereby aiming to produce more accurate and individualized EEG data augmentation.

Pushing the boundaries of cross-subject EEG analysis, this paper introduces a novel framework: Transfer and Robust Adaptation for New Subjects in EEG Technology (TRANSIT-EEG). It aims to address the challenges in cross-subject learning models for EEG signal processing during each stage of modelling by integrating subject-specific data augmentation, subject-specific model, and finally adapting the model to new subject.

The contributions of this paper are as follows:

1) Introduction of the novel TRANSIT-EEG framework, the first of its kind for effective cross-subject transfer, incorporating subject-specific components such as the model, augmentation technique, and finetuning methodology.
   a) Extension of Self Organising Graph Neural Network (SOGNN) [3] by replacing Graph Convolutional Networks (GCN) [14] with Graph Attention Transformers (GAT) [7] to introduce Self Organising Graph Attention Transformer (SOGAT) model for cross-subject learning.
   b) Introduction of the Individualised Diffusion Probabilistic Model (IDPM) augmentation, which uses diffusion models to separate noise from subject-specific signals and generates synthetic samples by augmenting only the subject-specific part for training and finetuning the model.
   c) Evaluation of different finetuning techniques such as full finetuning, classification head finetuning and adapter finetuning of graph layers using full-rank and low-rank adapters (LoRA) for few-shot domain adaptation with new samples.
2) Evaluation of the end-to-end TRANSIT-EEG framework on the SEED [15] and PhyAat [16] datasets, demonstrating improved classification accuracy for new subjects.
3) Interpretability of different EEG topographic maps and adjacency matrix of channels for different emotional states of SEED [15].

The experimental results demonstrate that using the SOGAT model, graph-based learning effectively identifies the relevant labels on both SEED [15] and PhyAat [16] datasets. Combined with a subject-specific data augmentation technique - IDPM, it significantly improves the quality of cross-subject transfer. The TRANSIT-EEG framework rapidly adapts to a new subject in a few-shot learning setup using Low-Rank Adaptation (LoRA) while achieving an F1-score of 91.53% and 87.78% on SEED [15] and PhyAat [16], respectively.

The remainder of this paper is structured as follows: Section II provides a more detailed review of existing work in subject-specific modeling and adaptation. Section III formulates the problem statement and introduces the proposed TRANSIT-EEG framework. Section IV presents the results of our experiments, and Section V concludes the work with a discussion of future directions.

## II. Related work and background

The problem of cross-subject EEG signal classification has garnered significant attention due to the increasing demand for scalable and robust EEG analysis frameworks in various domains. This section reviews the related work in three aspects: cross-subject learning using GNN, data augmentation using generative models, and domain adaptation for cross-subject transfer.

### A. Cross-Subject learning using GNN

The application of GNNs to EEG signal processing has seen significant advancements in recent years. Foundational work on GCN [14] and GAT [7] marked a substantial leap in structured data analysis. Building upon these foundations, research in EEG signal analysis has produced models such as EEG-GNN [17], EEG-GAT [18], RGNN [5], BiHDM [6] and SOGNN [3], demonstrating the efficacy of graph-based models for analyzing the multivariate nature of EEG signals over non-spatial models as shown for the first time by EEG-GNN [17]

Early approaches like EEG-GNN, EEG-GAT, RGNN, BiHDM utilized fixed graph structures, with nodes representing electrodes and edges defined by either pre-determined or attention-based adjacency matrices. However, these fixed graph approaches have inherent limitations. They struggle to adapt to individual variations in EEG signals and fail to capture the dynamic patterns inherent in EEG data.

SOGNN addressed the limitations of fixed graph structures by introducing a dynamic graph construction strategy based on input signals. This method enhanced EEG signal classification accuracy across subjects by incorporating a subject-specific dimension into graph formation. By allowing the graph structure to evolve based on input data, SOGNN can better capture the unique characteristics of each subject's EEG patterns, leading to improved cross-subject generalization compared to its predecessors, such as RGNN and BiHDM. Notably, all previous work, including SOGNN, relied on GCN. Although EEG-GAT has demonstrated that GAT can outperform GCN, recent studies have yet to incorporate GAT into their architectural designs.

Table I compares EEG-GNN [17], EEG-GAT [18], and SOGNN [3] in terms of graph structure construction, neuroscientific interpretability, application focus, learning mechanisms, channel selection, and computational efficiency. While EEG-GNN [17] and EEG-GAT [18] are designed for general EEG signal classification, SOGNN[3] focuses on cross-subject EEG emotion recognition. The inherent limitations in fixed graph structures prompted the development of SOGAT, which leverages subject-specific graph construction and advanced GNN architecture.

### B. Data Augmentation using generative models

Recent advancements in EEG data augmentation have significantly leveraged the capabilities of Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs).

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:26:00 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Cognitive and Developmental Systems. 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/TCDS.2025.3529669



TABLE I: Comparative Analysis of EEG-GNN, EEG-GAT, and SOGNN

| Factors | EEG-GNN [17] | EEG-GAT [18] | SOGNN [3] |
|---|---|---|---|
| Graph structure construction | Flexible yet pre-defined based on neural connectivity or heuristic distances | Adaptive through multi-head attention mechanism to learn graph topology | Dynamically constructed by self-organised module for each signal |
| Neuroscientific Interpretability | Enhanced through visualisation of connectivity between salient nodes | NA | NA |
| Application focus | Generalised EEG signal classification | Generalised EEG signal classification | Specialised in cross-subject EEG emotion recognition |
| Learning mechanism | Graph Convolution Network (GCN) | Multi-head attention mechanism to parameterise the Graph shift pperator | Self-organised Graph Neural Networks followed by Convolutional Neural Network (CNN) |
| Channel selection and computational efficiency | Emphasized for reducing computational costs and designing portable EEG headsets | NA | NA |

These techniques address the chronic issue of data scarcity in EEG research, particularly in the context of deep learning applications.

GANSER [10] employs GAN based self-supervised framework specifically designed for EEG signal generation. GANSER introduces several key innovations that set it apart in the field of EEG augmentation. Primarily, it can generate synthetic multi-channel EEG data, preserving the spatial correlations crucial for accurate EEG analysis. The model also incorporates a temporal constraint, ensuring that the generated EEG signals maintain temporal coherence, a critical aspect of EEG data. The authors demonstrate GANSER's effectiveness in augmenting data for EEG-based emotion recognition tasks, showing significant improvements in classification accuracy. The architecture of GANSER consists of a generator that produces synthetic EEG data and a discriminator that distinguishes between real and synthetic samples. The self-supervised aspect involves pre-training the discriminator on an auxiliary task of predicting rotated EEG signals, which helps in learning meaningful EEG features. This innovative approach allows GANSER to capture complex EEG characteristics more effectively than traditional GAN architectures.

Variational Autoencoders have also shown promise in EEG data augmentation, offering a different approach compared to GANs. Aznan et al. [19] propose a Conditional Variational Autoencoder (CVAE) for generating synthetic Steady-State Visual Evoked Potential (SSVEP) EEG signals. Their model conditions the generation process on specific SSVEP frequencies, allowing for targeted augmentation of under-represented classes. This approach is particularly valuable in scenarios where certain EEG patterns are rare or difficult to collect. Addressing the temporal aspects of EEG data, Louizos et al. [20] introduce a recurrent VAE architecture that captures the temporal dynamics of EEG signals. This approach is particularly effective for generating time-series EEG data, preserving both the spectral and temporal characteristics of the signals. The recurrent nature of this VAE allows it to model the complex temporal dependencies present in EEG data, a crucial factor in many EEG-based applications. To overcome the tendency of traditional VAEs to produce overly smooth outputs, Luo et al. [21] combine VAE with adversarial training, creating an Adversarial Variational Autoencoder (AVAE) for EEG generation. This hybrid approach aims to produce more realistic and detailed synthetic EEG signals by leveraging the strengths of both VAE and GAN architectures. The adversarial component helps in generating sharper, more realistic EEG patterns, while the VAE framework ensures a stable training process and enables meaningful latent space manipulations. VAE-based methods offer several advantages for EEG augmentation. They provide a probabilistic framework for modeling the underlying distribution of EEG data, allowing for more nuanced control over the generated signals. The latent space of VAEs can be manipulated to generate diverse yet realistic EEG samples, which is particularly useful for data augmentation tasks. Moreover, VAEs often exhibit more stable training compared to GANs, which can be beneficial when working with complex EEG data.

Despite these advancements, both GANSER and VAE-based approaches face challenges in capturing subject-specific characteristics of EEG data. While they excel at generating plausible EEG signals, they often struggle to preserve the unique individual variations that are crucial for personalized EEG analysis and cross-subject generalization. This limitation underscores the need for more advanced techniques that can incorporate subject-specific information in the augmentation process, a gap that our proposed Individualised Diffusion Probabilistic Model (IDPM) aims to address.

### *C. Domain adaptation for cross-subject transfer*

Domain adaptation involves applying learned knowledge from one domain (source domain) to another (target domain) with differing data distributions. In EEG signal processing, cross-subject transfer is a domain adaptation problem due to individual physiological and neurological differences. Hang et al. [4] tackled this with a Deep Domain Adaptation Network (DDAN), employing a Convolutional Neural Network (CNN) for feature extraction from raw EEG data and Maximum Mean Discrepancy (MMD) to minimise distribution discrepancy between source and target subjects. Centre-based Discriminative Feature Learning (CDFL) enhanced recognition performance for target domain EEG data.

Two prevalent domain adaptation techniques are finetuning the last layers of a pre-trained model and finetuning deeper

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:26:00 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Cognitive and Developmental Systems. 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/TCDS.2025.3529669



embedding layers via Adapters [22]. Adapter-based models, including Low-Rank Adapters (LoRA) [23], introduce small, task-specific modules into pre-trained models, reducing parameter and computational overhead. LoRA has demonstrated competitive performance in NLP and CV [23; 24; 25] domains and holds potential for EEG domain adaptation, suggesting a promising avenue for more efficient and scalable solutions in cross-subject EEG signal recognition.

## III. Methodology

This section delineates the methodology of the proposed TRANSIT-EEG framework, a meticulously engineered solution to address the challenges inherent in a few-shot subject transfer within EEG signal classification.

### A. Overall Framework

Building on the foundations laid for Domain Adaptation, such as those employing CS-GAN [26] and Domain Adversarial Sparse Learning model (DASL), which integrates Domain Adversarial Neural Networks (DANN) and GAN [27]. The proposed TRANSIT-EEG framework amalgamates cutting-edge techniques such as data augmentation and domain adaptation to improve cross-subject classification accuracy and ensure a cohesive subject-transfer process. Fig. 1 illustrates the three cardinal phases of the TRANSIT-EEG framework: Pretraining, Finetuning, and Inferencing. Each box denotes an individual phase and contains steps executed in each of them sequentially. All three phases contribute to efficient subject transfer with a few examples of the new subject.

**Phase 1 - Pretraining**: This phase lays the foundation for subject transfer by pretraining two critical components: the IDPM synthetic signal generator and the SOGAT classifier. It trains the IDPM for all subjects and then uses the trained generator to sample new signals, augmenting the training data. Subsequently, it trains the SOGAT classifier on a combination of original and synthetic data to capture the subject-specific characteristics inherent in EEG signals.

**Phase 2 - Finetuning**: The model obtained after Phase 1 lacks information about new subjects. Therefore, the proposed TRANSIT-EEG finetunes the pretrained model using adapter finetuning [28]. The process involves generating synthetic signals for the new subject by sampling from the IDPM trained in Phase 1 and finetuning the base model. There are three standard methodologies for finetuning:

1) *Full finetuning*: Updates the network layers' weights.
2) *Classification Head finetuning*: Updates the weights only of the last layer of the network, which, in the case of a classification model, is the classification head.
3) *Adapter finetuning*:
   a) Adds a new layer on top of the existing layers and fine-tunes only this new layer, referred to as Adapter.
   b) Low-Rank Adapter (LoRA) is a type of adapter finetuning that fine-tunes two small dimensional matrices to save computational resources.

**Phase 3 - Inferencing**: The final phase, inferencing, leverages the fine-tuned model obtained in Phase 2 to perform inferencing on new signals from the adopted subject. The novel design of the preceding phases ensures that the model is adeptly tuned to the specific characteristics of the new subject, thereby promising enhanced classification performance for the new subject and sessions.

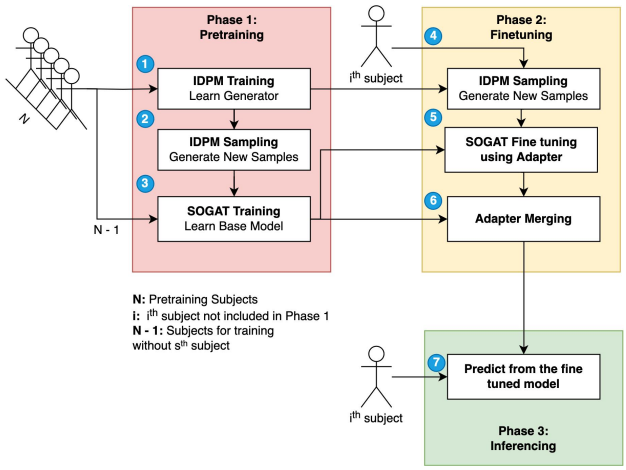



Fig. 1: Schematic representation of the proposed TRANSIT-EEG framework.

The proposed TRANSIT-EEG framework provides a robust solution to the challenge of few-shot subject transfer in EEG signal classification. By combining subject-specific synthetic data generation, self-organising graph attention model, and adapter-based finetuning, it advances subject-specific modelling and transfer methods in EEG signal processing.

TRANSIT-EEG comprises of three main components, each addressing a crucial aspect of the framework. The first component is the data augmentation technique, which utilizes the IDPM. This technique, detailed in subsection III-B, enables the generation of subject-specific synthetic data, enhancing the model's ability to generalize across subjects. The second component is the classification model, which employs SOGAT. This model, described in subsection III-C, leverages graph-based attention mechanisms to capture complex inter-channel relationships in EEG signals. The third component focuses on domain adaptation using LoRA, a technique explained in subsection III-D. This approach facilitates efficient finetuning of the model for new subjects, enabling rapid adaptation with minimal computational overhead.

Fig. 1 illustrates these components and their interactions. The following subsections explain each component in detail and describe its role in the respective phase of the framework.

### B. Subject-specific data augmentation

The proposed IDPM is a generative model that creates signals by decomposing the original signal into subject-specific and clean components. It removes the need for preprocessing before augmentation and artifact removal. The IDPM sampler can generate many signal augmentations beyond standard geometric transformations. The core of IDPM is its sampler, which uses accurate EEG signals to produce synthetic EEG data. It separates subject-specific noise during the denoising

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:26:00 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Cognitive and Developmental Systems. 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/TCDS.2025.3529669



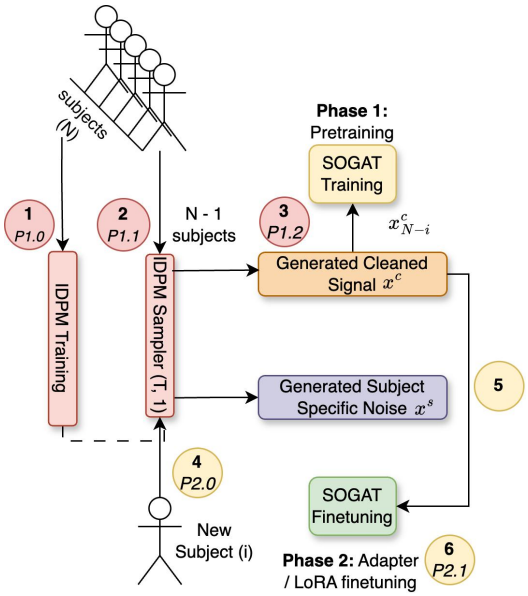



Fig. 2: Process of augmenting data for model training and finetuning using proposed Individualised Diffusion Process Model (IDPM)

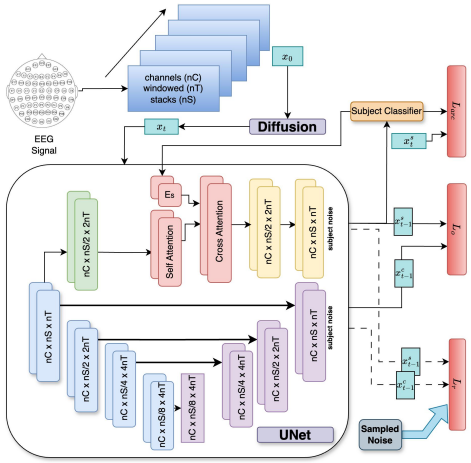



Fig. 3: Proposed Individualised Diffusion Process Model (IDPM). $x$ = input signal, $x^C$ = cleaned signal, $x^S$ = subject-specific artifacts

process unlike the vanilla DDPM [29] which denoises the input signal only without having seperate branches for artifacts and the informative part of the signal.

Fig.2 illustrates how IDPM decomposes the signal into cleaned and subject-specific artifacts, which are used further for training and finetuning purposes. Fig.3 illustrates how IDPM continuously adds noise into the original signal $x_0$ through a Markov process, sampling variables $\{x_0, x_1, \ldots, x_{t-1}, x_t, \ldots, x_T\}$ until $x_T$ becomes a normal noise distribution $p(x_T) \sim \mathcal{N}(x_T; 0, I)$, which is called the *diffusion process*.

To achieve signal decomposition, IDPM refines the standard denoising process into two outputs: cleaned subject-specific and noise signals. Unlike the standard denoising diffusion processes, which do not typically target subject-specific noise, IDPM is explicitly designed to handle subject artifacts in EEG data. It models human artifact removal as a generative denoising process, learning subject-specific noise alongside invariant brain signals. The denoising process is decomposed at each step into subject-specific noise and invariant signal.

It first captures the unique attributes of EEG signals from individual subjects through a conditioning process. Then, it employs a series of forward diffusion and reverse diffusion steps to generate new samples. This process is structured around a probabilistic framework where the forward process gradually introduces noise into clean EEG signals, and the reverse process aims to reconstruct the original signal by removing this noise, thereby generating new, synthetic EEG samples.

1) **Forward diffusion process:** The Equation 1 defines the forward process:

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t I), \quad (1)$$

where $x_0$ is the original clean EEG signal, $\beta_t$ is a variance schedule across diffusion steps, and $I$ is the identity matrix. This process incrementally adds Gaussian noise, transforming the EEG data into a latent space where the noise level is maximised.

2) **Reverse diffusion process:** The Equations 2 describes the reverse process that aims to reconstruct the clean signal from the noisy latent representation by iteratively reducing the noise.

$$p(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}; \mu(x_t, t), \sigma^2(x_t, t)), \quad (2)$$

where $\mu$ and $\sigma^2$ are functions learned by the model to effectively denoise the signal.

3) **Loss functions:** Three primary loss functions Reverse Loss, Orthogonal Loss, and Arc-Margin Loss guides the effectiveness of the Reverse Diffusion Process. These functions are essential for ensuring the quality of the generated EEG samples and enhancing the model's ability to generalise across subjects.

   a) *Reverse loss ($L_r$)*: This ensures that the model accurately reconstructs the clean EEG signals from the noisy data, as defined by Equation 3:

$$L_r = \mathbb{E}_{x_0,\epsilon_t}\left[\|x_0 - \hat{x}_0(x_t, t)\|^2\right], \quad (3)$$

   where $\hat{x}_0(x_t, t)$ is the estimate of the original signal $x_0$ from the noisy data $x_t$ at diffusion time $t$.

   b) *Orthogonal loss ($L_o$)*: This loss function maintains the orthogonality between the components of subject-specific noise and the intrinsic EEG signals, ensuring that the augmentation does not introduce bias, as defined by Equation 4:

$$L_o = \|(\mu_\theta(x_t, t, s) \cdot \mu_\phi(x_t, t))\|_F^2, \quad (4)$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:26:00 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Cognitive and Developmental Systems. 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/TCDS.2025.3529669



where $\mu_\theta$ and $\mu_\phi$ denote mappings for noise and clean signal components, respectively, and $\|\cdot\|_F$ represents the Frobenius norm.

c) *Arc-Margin loss ($L_{arc}$)*: The Arc-Margin Loss enhances the separability of EEG samples from different subjects, which is crucial for tasks such as subject-specific analysis and classification. Equation 5 defines this loss:

$$L_{arc} = -\log \frac{e^{s(\cos(\theta_{y_i}+m))}}{e^{s(\cos(\theta_{y_i}+m))} + \sum_{j\neq y_i} e^{s\cos\theta_j}}, \quad (5)$$

Where, $\theta_{y_i}$ represents the angle between the feature vector of the $i$-th subject and the corresponding class weight, $m$ denotes the margin penalty, and $s$ serves as a scaling factor.

The structural design of IDPM includes the standard UNet [30] architecture, as illustrated in Fig. 3. It receives diffusion results $x_t$ and simultaneously generates subject-specific noise $x_s^{(t-1)}$ and the clean data $x_c^{(t-1)}$ through two distinct streams. Both $x_s^{(t-1)}$ and $x_c^{(t-1)}$ are supervised by the three loss functions: Reverse Loss ($L_r$), Orthogonal Loss ($L_o$) and Arc-Margin Loss ($L_{arc}$).

The model parameters are optimised by randomly selecting a subject $s$ and the corresponding signal $x_0$. It further pretrains a subject classifier, namely EEGNet $\omega$, using pairs of signals and their associated subject IDs. Subsequently, the algorithm selects a random time step $t$ and introduces noise $e$. Using these parameters, the model calculates various components such as $e_\theta(x_t, s)$, $e_\phi(x_t, t)$, and $\omega(e_\theta(x_t, t, s))$ based on the model's parameters $\theta$ and $\phi$.

The primary objective is to minimise the combination loss $L = \lambda_r L_r + \lambda_o L_o + \lambda_{arc} L_{arc}$ through Adam optimiser with default parameters. This iterative process continues until convergence. The finetuning of the model's parameters ensures the production of denoised EEG signals that closely match real data and are tailored to individual subjects. This refinement is critical for enhancing the model's subject-specific analysis and classification capabilities. Fig. 3 outlines the IDPM training architecture which is further detailed in Algorithm 1.

During the inferencing phase of the IDPM, the model employs a progressive denoising strategy to refine EEG signals, as depicted in Fig. 2. Initially, the model starts with a noisy EEG signal. Through a series of generative steps, it incrementally constructs a cleaner version of the signal. At each time step, IDPM utilises two distinct generative pathways: one is tailored for mitigating subject-specific noise, and the other enhances the clean signal content. These pathways operate concurrently and are conditioned on the current noisy signal, subject-specific embeddings sourced from EEGNet, and the time step itself.

The outputs from these pathways are combined at each step, progressively reducing noise and subject-specific artefacts. This iterative refinement process culminates in the generation of the final cleaned signal, denoted as $x_0$ in the Algorithm (2). The resulting signal is optimally free from noise and artefacts, rendering it highly suitable for subsequent classification tasks using the SOGAT model.

**Algorithm 1** IDPM training algorithm for generating synthetic denoised EEG Signals

**repeat** ▷ Begin the iterative training process
Sample $(x_0, s) \sim q(x_0, s)$ ▷ Randomly sample initial clean EEG signal $x_0$ and subject $s$
Pretrain classifier $\omega$ using pairs $\{(x_0, s)\}$ ▷ Enhance classifier to predict subject based on EEG
Sample $t \sim \text{Uniform}(\{1, \ldots, T\})$ ▷ Randomly select a time step for simulation
Sample $e \sim \mathcal{N}(0, 1)$ ▷ Sample noise $e$ for the diffusion process
Compute $e_\theta(x_t, s)$, $e_\phi(x_t, t)$, $\omega(e_\theta(x_t, t, s))$ ▷ Using model parameters $\theta$ and $\phi$
Update $\theta$ and $\phi$ via gradient descent to minimize $L$ ▷ Adjust model parameters to minimize loss
Combination Loss: $L = \lambda_r L_r + \lambda_o L_o + \lambda_{arc} L_{arc}$ ▷ Compute total loss from different components
**until** convergence ▷ Repeat until the model parameters stabilize

**Algorithm 2** IDPM sampling algorithm for generating denoised EEG Signals from Noise

$X_T \sim \mathcal{N}(0, 1)$ ▷ Initialize with a noise distribution
**for** $t = T, T-1, \ldots, 1$ **do**
  **if** $t > 1$ **then**
    $z \sim \mathcal{N}(0, 1)$ ▷ Sample noise for the reverse diffusion process
  **else**
    $z = 0$
  **end if**
  $X_{t-1}^s = \frac{1}{\sqrt{\alpha_t}} \left( X_t - \frac{1-\alpha_t}{\sqrt{1-\alpha_t}} e_\theta(X_t, t, s) + \sigma_t z \right)$
  $X_{t-1}^c = \frac{1}{\sqrt{\alpha_t}} \left( X_t - \frac{1-\alpha_t}{\sqrt{1-\alpha_t}} e_\phi(X_t, t) + \sigma_t z \right)$
**end for**
**return** $x_0^s, x_0^c$

IDPM augments data for all subjects during the pretraining phase to enhance the model's generalisation. In contrast, during the finetuning phase, samples are generated tailored explicitly to an individual subject, ensuring more precise adaptation and improved performance in subject-specific signal analysis.

Algorithm 1 begins by initialising the input signal $x_0$ and the noise vector $\epsilon_0$ with standard Gaussian distributions. Subsequently, it iteratively updates the denoised signal $x_t$ and the noise vector $\epsilon_t$ at each time step $t$ using the diffusion process. This process involves applying transformations learned during training to the input signal and the noise vector. The algorithm terminates after a fixed number of time steps, resulting in the final denoised signal. The post-processing step involves subtracting the mean and dividing by the standard deviation to obtain a normalised denoised signal.

Algorithm 2 outlines sampling domain variance based on subject $s$ and generating unconditional EEG signals from pure

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:26:00 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Cognitive and Developmental Systems. 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/TCDS.2025.3529669



noise. To enhance a given signal, subject-specific domain variance $x_s^0$ is sampled and added directly to $x_0$. However, this method overlooks inherent signal noise. In real-world scenarios, denoising is more practical, not solely based on subject $s$ but also on the signal itself. The denoising process $x_T, \ldots, x_t, x_{t-1}, \ldots, x_0$, noise is systematically eliminated at each iteration. Denoising recorded EEG signals, assumed to be in an intermediate state $x^{1/2}$, allows direct separation of domain variance and clean content from the raw signal $x_{\text{raw}}$. This assumption replaces the iterative sampling process with a single inference.

### C. Classification model: SOGAT for signal classification

The core of the pretraining phase lies in the SOGAT model, designed to master the intricacies of EEG signal classification. It is an amalgam of dynamic graph construction and GAT, each contributing to capturing the subtle and complex relationships among EEG channels. Following are the various components that are core to SOGAT model.

1) **Feature engineering using Differential Entropy (DE)**: A cornerstone of the signal classification pipeline is the judicious selection of features that encapsulate the underlying characteristics of EEG signals. DE emerges as the feature of choice, buoyed by its demonstrated efficacy as reported by [3]. It is a barometer of uncertainty associated with continuous random variables, embodying the complexity and irregularity of EEG signal distributions over time. Equation 6 provides the mathematical expression for differential entropy.

$$h(X) = -\int_{-\infty}^{\infty} f(x) \log f(x)\, dx \tag{6}$$

where, $h(X)$ represents the differential entropy of a continuous random variable $X$ with probability density function $f(x)$. The employment of DE enables the SOGAT model to delve into the intricate patterns embedded in the EEG data, thereby fostering more accurate and robust classification.

2) **Dynamic graph construction**: EEG signals can be structured as graphs, where nodes represent EEG electrodes and edges represent functional relationships, encapsulated in an adjacency matrix [31; 32; 33; 34]. Traditional brain graphs use pre-defined distance functions; however, these fixed structures may not adequately model dynamic brain signals across different subjects. SOGNN [3] introduces a self-organised graph construction module that determines the graph based on input brain signals. Equation 7 represents the creation of the graph adjacency matrix across different channels of the signal.

$$a_{ij} = f(v_i, v_j) = \frac{\exp(\theta(v_i W)\theta(v_j W)^T)}{\sum_{i=1}^{N} \exp(\theta(v_i W)\theta(v_j W)^T)} \tag{7}$$

Here, $v$ is the feature vector of a node (EEG electrode) within $\mathcal{V}$ (a matrix in $\mathbb{R}^{N \times F}$), $N$ is the number of nodes, $W$ is the weight matrix, and $\theta$ is the $\tanh$ activation

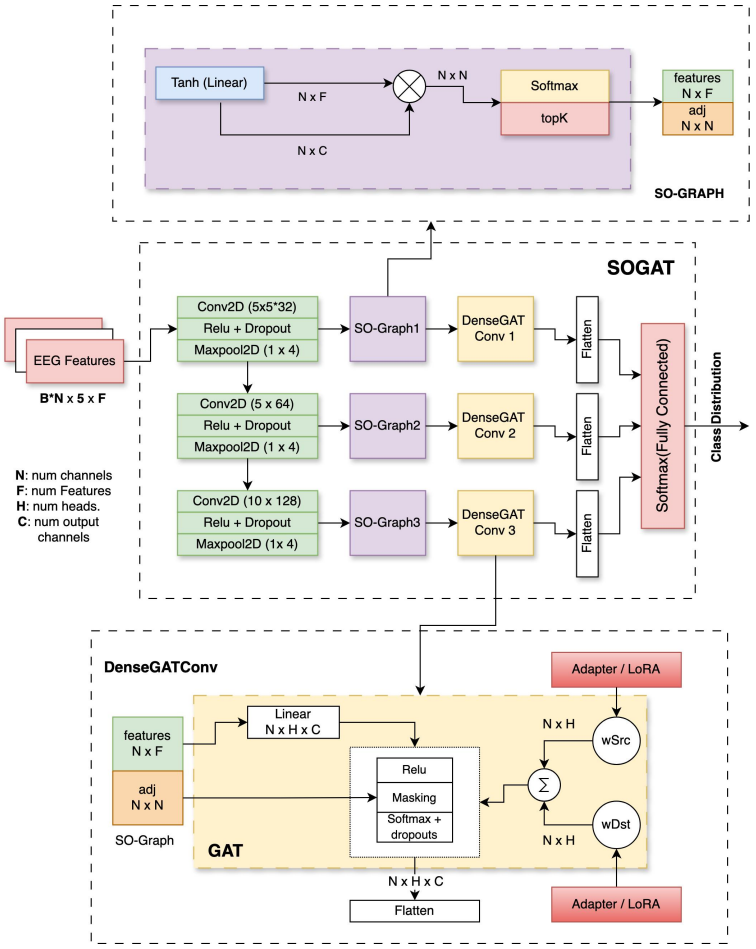



Fig. 4: Architecture of proposed Self-organised Graph Attention Transformer (SOGAT) model for EEG signal classification.

function of a linear layer. The exponential function in the softmax activation normalises the adjacency weights.

$$G = \tanh(\mathcal{V}W) \tag{8}$$

$$A = \text{Softmax}(GG^T) \tag{9}$$

Equation 8 computes the adjacency matrix $G$ by applying $\tanh$ to $\mathcal{V}W$. Equation 9 applies softmax to $GG^T$ to yield a normalised adjacency matrix $A$.

To optimise computational efficiency, a sparse graph retains only the $k$ largest weights from the adjacency matrix, setting smaller weights to zero as shown in Equation 10:

$$\begin{aligned} A' &= A \\ A'_{i,\text{argtop}_k(A[i,:])} &= 0 \quad \forall\, i \in \{1, 2, \ldots, N\} \end{aligned} \tag{10}$$

where, the function $\text{argtop}_k(\cdot)$ identifies the indices of the top-$k$ values in each vector $A[i,:]$. This top-$k$ strategy is

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:26:00 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Cognitive and Developmental Systems. 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/TCDS.2025.3529669



similar to a modified max-pooling layer and facilitates network parameter updates via backpropagation. The self-organised graph construction module dynamically adapts the graph structure to input EEG features. Subsequently, graph convolutional layers process these graphs to extract local and global features for emotion recognition. Similar to SOGNN [3], it consists of three conv-pool blocks, three self-organised graph layers, three graph convolution layers, a fully connected layer, and an output layer designed to handle EEG features organised by channels, bands, and timestamps.

3) **Signal classification model using GAT**: The core strength of SOGAT lies in its use of a GAT for signal classification. The TRANSIT-EEG chooses GAT because it effectively understands complex relationships within data, which is crucial for analysing interactions between brain regions and EEG channels for accurate classification. It uses the graph construction method from [3], which dynamically creates a graph for each subject, as discussed in the following subsection.
   Fig. 4 shows a schematic diagram of the SOGAT Network. The diagram displays the transmission of DE features across five bands to a CNN block, which extracts features at various scales (32, 64, 128). It then uses these extracted features to compute a sparse graph while retaining top-$k$ equals 10 connections as found best by original SOGNN [3] which is shown in purple as an SO-Graph1, SO-Graph2 and SO-Graph3. It inputs the resulting graph into the GAT layer, which learns the weights between the extracted features. Finally, it flattens, concatenates, and forwards the outputs from all three GAT layers to a fully connected layer to produce the output class distribution. The SOGAT architecture comprises 386,991 trainable parameters, therefore based on the size of the dataset on modern data GPUs (Nvidia T4), this model is faster to train with around 300 - 500 seconds per epoch.

### *D. Domain Adaptation: Adaptation to new subjects using LoRA*

The endeavour towards efficacious model adaptation for new subjects begins with the conventional practice of finetuning. Nevertheless, traditional finetuning often contends with "catastrophic forgetting," a difficulty where the model, while acclimating to new data, tends to obfuscate the learned features from the original data.

The narrative pivots with the advent of adapter-based finetuning, which introduces small, trainable modules called adapters instead of retraining the entire model. These adapters learn to tailor the model to new subjects without altering the pre-trained parameters, thereby alleviating catastrophic forgetting and representing a computationally efficient alternative to conventional finetuning.

The saga of model adaptation introduces Low-Rank Adaptation (LoRA), which fine-tunes the model by adapting only a low-rank subspace of the weight update matrices. Depicted in Equation 11, matrices $A$ and $B$ represent the low-rank decomposition of the weight update matrix $\delta_W$, contrasting with conventional finetuning.

$$\delta_W = A * B \tag{11}$$

By decomposing $\delta_W$ into two lower-dimensional matrices, $A$ and $B$, LoRA reduces the number of parameters that need to be fine-tuned. Matrix $A$ maps the high-dimensional weight space to a lower-dimensional subspace, while matrix $B$ maps it back to the original space. This low-rank approximation significantly reduces computational complexity and memory requirements, making finetuning more efficient.

The integration of LoRA into the SOGAT model, which is optimised for dense graph structures, highlights its utility in practical applications. The 'DenseGATConv' layer is particularly crucial in this context. This layer is designed to handle dense adjacency matrices, facilitating effective feature aggregation across all nodes:

1) **Low-Rank decomposition:** In the 'DenseGATConv' layer, the source (*att_src*) and destination (*att_dst*) matrices each undergo decomposition into two smaller matrices, $A$ and $B$. This approach reduces the full matrix $W$ to $W \approx AB$, simplifying the model's complexity and reducing the number of trainable parameters.
2) **Integration into DenseGATConv:**
   a) *Modify parameter initialisation :* Instead of initialising large full-rank matrices, smaller, low-rank matrices $A$ and $B$ are initialised for both *att_src* and *att_dst*.
   b) *Adjust the forward pass:* The forward method of 'DenseGATConv' replaces the computations of *att_src* and *att_dst* with the product of their corresponding low-rank matrices, calculating *att_src* and *att_dst* as $A_{\text{src}} \times B_{\text{src}}$ and $A_{\text{dst}} \times B_{\text{dst}}$, respectively.
3) **Training:** During the training phase, the low-rank matrices $A$ and $B$ for both source and destination receive updates exclusively. This targeted optimisation facilitates effective adaptation to new subjects while maintaining the stability and robustness of the learned features.

Post-finetuning, the adapters are integrated back into the model, enhancing the base model with cross-subject capabilities. This iterative process of integrating new subjects enhances the base model's generalisation and zero-shot capabilities by accommodating more subjects. During finetuning, the goal is to minimise the data required from new subjects. Hence, instead of collecting many samples, reliance is placed on a limited set of samples. The pre-trained Individualised Diffusion Probabilistic Model (IDPM) based data augmentation comes into play to generate more training samples, alleviating the data scarcity issue. For evaluation purposes, the model utilises 50% of a session per subject, and for finetuning, the rest 50% are used along with their augmentations. This strategy conserves resources and exemplifies the model's capability to adapt to a modest amount of data, rendering the TRANSIT-EEG framework highly scalable and applicable in real-world scenarios.

## IV. Experiments and Results

This section presents the experimental setup and the results for the various components of TRANSIT-EEG framework.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:26:00 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Cognitive and Developmental Systems. 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/TCDS.2025.3529669



TABLE II: Dataset Statistics for Different Experiments

| **SEED [15]** | | | | | |
|---|---|---|---|---|---|
| **Exp.** | **Num. subjects × sessions** | **Num. features** | **Num. classes** | **Aug. Factor** | **Total sessions** |
| Original Dataset | 15 × 45 | 5 × 265 | 3 | 1x | 675 |
| Single Fold of LOSO training w/o IDPM | 14 × 45 | 5 × 265 | 3 | 1x | 630 |
| Single Fold of LOSO training w IDPM | 14 × 45 | 5 × 265 | 3 | 1.5x | 945 |
| Single Fold of LOSO finetuning w IDPM | 1 × 21 | 5 × 265 | 3 | 5x | 105 |
| **PhyAat [16]** | | | | | |
| **Exp.** | **Num. subjects × sessions** | **Num. features** | **Num. classes** | **Aug. Factor** | **Total sessions** |
| Original Dataset | 25 × 432 | 5 × 384 | 3 | 1x | 10,800 |
| Single Fold of LOSO training w/o IDPM | 24 × 432 | 5 × 384 | 3 | 1x | 10,368 |
| Single Fold of LOSO training w IDPM | 24 × 432 | 5 × 384 | 3 | 1.5x | 15,552 |
| Single Fold of LOSO finetuning w IDPM | 1 × 216 | 5 × 384 | 3 | 5x | 1,080 |

Exp.: Experimental setup, Num.: Number of, Aug.: Augmentation

The proposed model is compared with other state-of-the-art graph models, and additional experiments are conducted to evaluate the effectiveness of different finetuning strategies and augmentation ratios. The framework implementation is available on GitHub.[1] The experiments were conducted on a platform equipped with an Nvidia T4 GPU, running Ubuntu 22.04, PyTorch 2.1, and PyTorch-Geometric 1.5.0.

### A. Experimental setup

This research turns to two prominent and widely acknowledged EEG datasets for empirical validation: Phyaat [16] and SEED [15; 35]. The Phyaat dataset [16] presents EEG signals, Galvanic Skin Response (GSR), and Photoplethysmography (PPG) signals gathered from 25 healthy participants via a 14-channel Emotiv Epoc device at a sampling frequency of 128 Hz. The model segments these signals by specific tasks, notably listening (0), writing (1), and resting (2). In juxtaposition, the SEED dataset [15; 35], tailored for EEG-based emotion recognition, encompasses recordings from 15 participants (7 males, 8 females) using a 62-channel ESI NeuroScan System. The emotional annotations for each session cover three distinct states: happiness (0), sadness (1), and neutrality (2).

The evaluation utilises the Leave-One-Subject-Out (LOSO) validation technique. The model initially undergoes pretraining using data from (N-1) subjects, supplemented with synthetic samples for each subject as shown in Fig. 1 in Section III-B, where $i^{th}$ subject is left out for finetuning. Subsequently, it is finetuned on half of the trials drawn in a stratified manner based on the class label from the omitted subject, including additional synthetic samples. The final model, enhanced with finetuned SOGAT, performs inference on all sessions of the separated subject, thereby thoroughly assessing its cross-subject transfer capabilities. Table II presents the dataset statistics for various experimental setups using the SEED [15] and PhyAat [16] datasets. It outlines four scenarios: the original dataset, LOSO training without IDPM, LOSO training with IDPM, and LOSO finetuning with IDPM. For each scenario, the table shows the number of subjects and sessions, feature count, number of classes, augmentation factor, and total sessions. The SEED dataset starts with 15 subjects and 45 sessions each, while PhyAat [16] has 25 subjects with 432 sessions each. During LOSO training, one subject is left out, reducing the subject count by one. IDPM augmentation increases the total sessions by a factor of 1.5x for training and 5x for finetuning. Notably, finetuning uses fewer sessions per subject: 21 for SEED [15] and 216 for PhyAat [16]. This table provides a clear overview of how data is structured and augmented throughout different experimental phases, helping readers understand the scale and composition of each setup.

[1]https://github.com/cahuja1992/transit-eeg

### B. Experiment 1: Choice of the network architecture

With the aim of cross-subject transfer, state-of-the-art graph models such as EEG-GNN [17], EEG-GAT [18], SOGNN [3] become the incumbent models to evaluate SOGAT. This experimental setup leverages a robust cross-validation strategy, Leave-One-Subject-Out (LOSO), to evaluate the effectiveness without any data leakage and measure cross-subject effectiveness.

LOSO calculates the validation metric across all the subjects one by one while training by keeping the subjects in validation out. For SEED [15], each of the 15 subjects is left out while training happens on 14 of them. Similarly, for PhyAat [16], LOSO validates all 25 subjects iteratively while training only 24 subjects at a time. Hence, LOSO evaluates all the subjects without any data leakage and averages the metric across all the subjects. Macro-F1 is chosen as the metric for final model selection for both datasets due to their almost uniform class balance. For each split of cross-validation, the data is as per the dataset statistics shown in Table II is splitted into training and validation split. The SEED dataset [15] uses 630 sessions (each with 1,325 features) for training and 45 sessions for validation. The PhyAat dataset [16] uses 10,368 sessions (each with 1,920 features) for training and 144 sessions for validation. All validation sessions had the same number of features as their corresponding training sessions. The model is then trained to maximum 200 epochs with learning rate of $1e^{-3}$ using Adam optimizer with stopping criteria of Area Under Curve (AUC) of 0.99 for the training reached.

Table III shows the results for model datasets, with LOSO F1 scores for the corresponding classes along with Macro F1, the results in bold represents the best results in that particular column. Analysis of the results demonstrates that SOGAT exhibits superior performance compared to its predecessor models. Specifically, SOGAT outperforms SOGNN, achieving a Macro F1 score of 87.75% on the SEED dataset [15], which

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:26:00 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Cognitive and Developmental Systems. 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/TCDS.2025.3529669



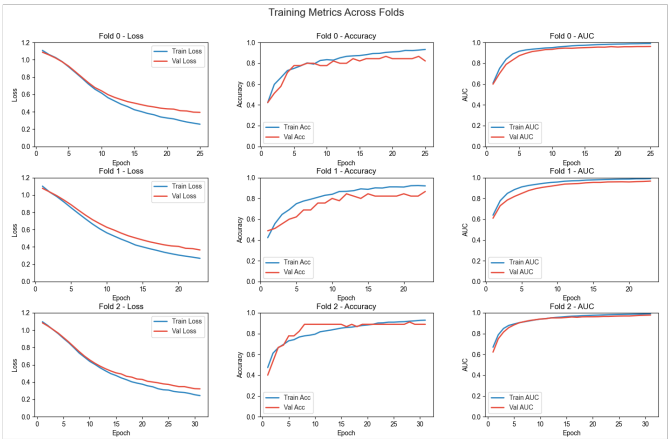



Fig. 5: Training Curves of SOGAT for 3 folds of LOSO

TABLE III: Graph Based Neural Networks performance on various datasets using LOSO based cross validation

| **SEED [15]** | | | | |
|---|---|---|---|---|
| **Model** | **Avg. F1 (%)** | **Hap. F1 (%)** | **Sad F1 (%)** | **Neut. F1 (%)** |
| EEG-GNN [17] | 77.35 | 79.38 | 75.83 | 76.83 |
| EEG-GAT [18] | 78.16 | 81.38 | 75.75 | 77.34 |
| SOGNN [3] | 86.69 | 85.56 | 85.77 | 88.74 |
| SOGAT | **87.75** | **86.82** | **86.52** | **89.92** |
| **PhyAat [16]** | | | | |
| **Model** | **Avg F1 (%)** | **List. F1 (%)** | **Writ. F1 (%)** | **Rest. F1 (%)** |
| EEG-GNN [17] | 69.35 | 72.38 | 66.83 | 68.83 |
| EEG-GAT [18] | 73.16 | 78.38 | 68.75 | 72.34 |
| SOGNN [3] | 80.05 | 80.81 | 77.44 | 81.92 |
| SOGAT | **81.64** | **82.72** | **79.78** | **82.43** |

FT Tech.: Finetuning Technique, Hap.: Happiness, Sad: Sadness, Neut.: Neutral, List.: Listening, Writ.: Writing, Rest.: Resting, Ft. Time: Finetuning Time

represents an absolute improvement of 1.06 percentage points over SOGNN's 86.69%. On the PhyAat dataset [16], SOGAT attains a Macro F1 score of 81.64%, surpassing SOGNN's 80.05% by an absolute margin of 1.59 percentage points. These improvements, while modest, suggest that SOGAT offers enhanced robustness and generalisation capabilities across subjects. Fig. 5 shows the training curve for the training and validation loss and accuracy for two cross-validation folds of LOSO. It shows that with more epochs, the training loss keeps on reducing, whereas validation loss saturates early on and finally stops due to the stopping criteria that checks for training Area Under Curve (AUC) threshold of 0.99 converging to the optimal model without overfitting.

### *C. Experiment 2: Effect of subject-specific data augmentation using IDPM*

Experiment 2 builds upon Experiment 1 by introducing data augmentation using the proposed Individualized Diffusion Probabilistic Model (IDPM) strategy. While Experiment 1 serves as a baseline without any augmentation during model training, this experiment demonstrates the effects of the porposed IDPM augmentation. It implements Phase 1 of our methodology, where IDPM is trained on data from all subjects and their trials, as illustrated in Fig. 2. The trained IDPM then generates new samples to augment the training dataset. To evaluate the effectiveness of this augmentation approach, we employ the same Leave-One-Subject-Out (LOSO) validation strategy used in Experiment 1, allowing for a direct comparison of model performance with and without IDPM augmentation.

TABLE IV: Model performance with proposed IDPM data augmentation

| **SEED [15]** | | | | |
|---|---|---|---|---|
| **Tech.** | **Avg. F1 (%)** | **Hap. F1 (%)** | **Sad F1 (%)** | **Neut. F1 (%)** |
| SOGAT w/o IDPM | 87.75 | 86.82 | 86.52 | 89.92 |
| SOGAT with IDPM (Phase 1) | **88.62** | **87.36** | **88.25** | **90.27** |
| **PhyAat [16]** | | | | |
| **Tech** | **Avg F1 (%)** | **List. F1 (%)** | **Writ. F1 (%)** | **Rest. F1 (%)** |
| SOGAT w/o IDPM | 81.64 | 82.72 | 79.78 | 82.43 |
| SOGAT with IDPM (Phase 1) | **82.50** | **82.72** | **81.62** | **83.18** |

FT Tech.: Finetuning Technique, Hap.: Happiness, Sad: Sadness, Neut.: Neutral, List.: Listening, Writ.: Writing, Rest.: Resting

During each fold of cross-validation, the training data is augmented by 50%. For the SEED dataset [15], this results in 924 training samples and 66 validation samples. Similarly, for the PhyAat dataset [16], the augmented dataset comprises 5,182 training samples and 216 validation samples.

Algorithm 2 in Section III-B then generates new samples for each subject, which increases the dataset size by 5 times for training. To evaluate this methodology, the model without augmentation acts as a baseline. Table IV shows that for both SEED [15] and PhyAat [16] using IDPM-based data augmentation, where, the results in bold represent the best results in that particular column. The results demonstrate that IDPM-based data augmentation consistently enhances model performance. For the SEED dataset, SOGAT with the proposed IDPM achieves a Macro F1 score of 88.62%, representing an absolute improvement of 0.87 percentage points over the non-augmented model's 87.75%. Similarly, for the PhyAat [16] dataset, the augmented model attains a Macro F1 score of 82.50%, surpassing the baseline by 0.86 percentage points. Notably, the performance improvements extend across all individual class F1 scores for both datasets.

These results suggest that the porposed IDPM-based augmentation enhances the model's ability to generalise across subjects, as evidenced by the consistent improvements in LOSO cross-validation metrics. The augmentation technique appears to be particularly effective in addressing class imbalance and capturing inter-subject variability, leading to more robust EEG signal classification.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:26:00 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Cognitive and Developmental Systems. 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/TCDS.2025.3529669



### D. Experiment 3: Effectiveness of Adapter finetuning

*1) Overall Effect of Adapter finetuning with IDPM-based data augmentation:* This experiment builds upon the previous experiment where IDPM aided pretraining. This experiment increases the finetuning data by 5 times by generating synthetic sessions for the new subjects using the IDPM sampler. For the LOSO-based cross-validation, this experiment uses 50% of the new subject samples during finetuning. It then generates the synthetic sessions, and calculates the average F1 score on the other 50% that was left out. This approach retains 21 sessions for finetuning and 24 for evaluation in the SEED [15] dataset across each class equally. For PhyAat [16] dataset, it maintains 216 sessions across the three classes each for finetuning and evaluation. During finetuning, these sessions are further augmented by a factor of 5x, yielding 105 sessions for SEED [15] and 1080 sessions for PhyAat [16]. The model is then finetuned with a reduced learning rate of 1e-6 for both full finetuning and adapter finetuning over 5 epochs. Two finetuning approaches are employed: "TRANSIT-EEG.A" using a Full Rank Adapter and "TRANSIT-EEG.B" using a Low-Rank Adapter (LoRA).

Table V presents the results of the complete TRANSIT-EEG setup. It shows different stages for both datasets. The results in bold represent the best results in that particular column. TRANSIT-EEG.A improves the LOSO Macro-F1 of SEED [15] by 0.61%. It achieves an average F1 of 91.53%. This is an improvement over its finetuning counterpart without IDPM, which achieves 90.92%. Similarly, TRANSIT-EEG.B improved from 89.88% to 90.49%. This is also compared to its finetuning counterpart without IDPM. TRANSIT-EEG.A achieves 91.53% F1. TRANSIT-EEG.B achieves an F1 of 90.49% in just 2082 seconds. This is almost half of TRANSIT-EEG.A's 3929 seconds. Both achieve these results within the same 5 epochs. These results improves significantly over Phase 1 training alone, where Phase 1 training achieves a maximum average F1 of 88.62%.

For PhyAat [16], TRANSIT-EEG.A demonstrates a substantial improvement, increasing the average F1 from 84.81% (finetuning without IDPM) to 87.78%, a gain of 2.97%. TRANSIT-EEG.B shows a similar trend, improving from 83.81% to 87.04%. Compared to SOGAT with IDPM (Phase 1 only) at 82.50%, both TRANSIT-EEG variants show marked improvements: TRANSIT-EEG.A reaches 87.78%, and TRANSIT-EEG.B achieves 87.04%. Notably, TRANSIT-EEG.B accomplishes this in 1184 seconds, significantly faster than TRANSIT-EEG.A's 2089 seconds.

*2) Comparison of Various Finetuning methodologies without IDPM-based data augmentation:* This subsection presents a comparative analysis of different finetuning methodologies for EEG signal classification, emphasising their performance using Leave-One-Subject-Out (LOSO) accuracy [36; 37; 38].

Following methodologies are evaluated for finetuning new samples

i *Full finetuning:* Updates all network parameters using new data, allowing full adaptation but risks overfitting and forgetting previous knowledge.

TABLE V: TRANSIT-EEG performance across different phases on SEED [15] and Phyaat [16] datasets.

| **SEED [15]** | | | | | |
|---|---|---|---|---|---|
| **Tech.** | **Hap. F1 (%)** | **Sad F1 (%)** | **Neut. F1 (%)** | **Avg. F1 (%)** | **Ft. Time (sec)** |
| SOGAT with IDPM (Phase 1) | 87.36 | 88.25 | 90.27 | 88.62 | - |
| SOGAT + Phase 1 + Phase 2 (Only Full Rank Adapter) | 91.00 | 90.20 | 91.56 | 90.92 | 2273 |
| SOGAT + Phase 1 + Phase 2 (Only LoRA) | 89.80 | 90.02 | 89.82 | 89.88 | 1186 |
| **TRANSIT-EEG.A** = SOGAT + Phase 1 + Phase 2 (IDPM + Full Rank Adapter) | **90.89** | **90.81** | **92.89** | **91.53** | 3929 |
| **TRANSIT-EEG.B** = SOGAT + Phase 1 + Phase 2 (IDPM + LoRA) | 90.02 | 89.64 | 91.82 | 90.49 | **2082** |
| **Phyaat [16]** | | | | | |
| **Tech** | **List. F1 (%)** | **Writ. F1 (%)** | **Rest. F1 (%)** | **Avg F1 (%)** | **Ft. Time (sec)** |
| SOGAT with IDPM (Phase 1) | 82.72 | 81.62 | 83.18 | 82.50 | - |
| SOGAT + Phase 1 + Phase 2 (Full Rank Adapter) | 83.51 | 85.50 | 85.42 | 84.81 | 1927 |
| SOGAT + Phase 1 + Phase 2 (LoRA) | 82.50 | 84.00 | 85.00 | 83.81 | 983 |
| **TRANSIT-EEG.A** = SOGAT + Phase 1 + Phase 2 (IDPM + Full Rank Adapter) | **84.70** | **88.82** | **89.92** | **87.78** | 2089 |
| **TRANSIT-EEG.B** = SOGAT + Phase 1 + Phase 2 (IDPM + LoRA) | 85.72 | 87.98 | 87.43 | 87.04 | **1184** |

Tech.: Technology, Avg. F1: Average F1 score, Hap: Happy, Sad: Sadness, Neut: Neutral, List: Listening, Writ: Writing, Rest: Resting, Ft. Time: Finetuning Time.

ii *Classification head finetuning :* Updates only the last layer, preserving deeper pre-trained features, reducing computational load and catastrophic forgetting, but may miss complex new features.

iii *Adapter finetuning :* Introduces adapter blocks in specific layers, balancing comprehensive retraining and minimal updates, enhancing feature representation in the proposed SOGAT model without forgetting prior data.

The evaluation of these finetuning methods used the PhyAat dataset [16], focusing on their performance without data augmentation and with only 50% of the samples per subject. Tests of adapter finetuning with different ranks (full rank, 32, 16, 8) showed that lower ranks correspond to decreased LOSO Macro-F1 scores. However, LoRA with a rank of 32 emerges as the optimal balance between latency and accuracy for applications with constraints such as on-device finetuning. Table VI summarizes these findings for PhyAat dataset [], illustrating that higher ranks correlate with enhanced perfor-

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:26:00 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Cognitive and Developmental Systems. 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/TCDS.2025.3529669



mance but increased computational demands. The full rank addapter achieves the highest Macro F1 score of 84.81%, while LoRA with rank 32 achieves 83.81% with the fastest finetuning time of 983 seconds. In comparison, full finetuning and classification head methods achieve 81.98% and 83.26% Macro F1 scores, respectively. This study also conducts paired t-tests to validate performance differences at a significance level of $\alpha = 0.05$. The test between the LoRA adapter method and the classification head approach yields a t-statistic of 7.0356 and a p-value of 0.0196. Similarly, the comparison between full rank adapter and full finetuning produced a t-statistic of 6.0584 with a p-value of 0.0262. These results statistically substantiate the effectiveness of adapter-based finetuning methods over other techniques in accommodating new subjects.

TABLE VI: Comparative performance of finetuning Techniques for the SOGAT model on the PhyAat dataset [16]

| FT Tech | Rank | Mac. F1 (%) | List F1 (%) | Writ F1 (%) | Rest F1 (%) | Ft. (sec) |
|---|---|---|---|---|---|---|
| Full finetuning | - | 81.98 | 80.90 | 80.90 | 84.14 | 2572 |
| Classification head | - | 83.26 | 81.40 | 83.92 | 84.46 | 1583 |
| Full rank adapter | full | **84.81** | **83.51** | **85.50** | **85.42** | 1927 |
| LoRA | 32 | 83.81 | 82.50 | 84.00 | 85.00 | **983** |
| LoRA | 16 | 84.88 | 83.90 | 83.90 | 86.85 | 690 |
| LoRA | 8 | 84.45 | 83.50 | 83.50 | 86.35 | 430 |

FT Tech.: Finetuning Technique, Mac. F1: Macro F1 score, List: Listening, Writ: Writing, Rest: Resting, Ft: Finetuning time.

*3) Impact of IDPM-based Data Augmentation during finetuning:* This experiment uses the SEED [15] dataset to examine the effect of session augmentation on EEG signal classification. This dataset was chosen due to its balanced nature, which helps avoid potential class bias in the results. The experimental procedure involves randomly selecting 21 sessions for each subject for finetuning. These sessions are then augmented using 1x (no augmentation), 2x, 5x, and 10x. The model is subsequently finetuned on the augmented data for each augmentation factor. Like Leave One Subject Out (LOSO) for cross-subject performance, Leave One Subject's Session Out (LOSSO) cross-validation evaluates cross-session performance. This strategy involves holding out one session as a test set for each subject while using the remaining sessions, along with synthetically generated signals, for finetuning. This process is repeated across all subjects, and the results are averaged to evaluate performance metrics. This methodological approach serves two primary purposes. First, it assesses the model's ability to generalise across different subjects. Second, it tests the model's robustness against variations in session-based training data. By systematically increasing the amount of finetuning data through augmentation, this experiment aims to understand how increased data availability impacts the model's ability to generalise to new, unseen samples. This approach provides insights into the relationship between data quantity and model performance in the context of EEG signal classification.

Fig. 6 summarises the classification accuracy across different session augmentation scales. Finally, performance metrics such as F1-score and accuracy are computed for each subject and averaged to provide a comprehensive overview of the impact of session augmentation on model performance. The results indicate that increasing the number of sessions used in training (through augmentation) potentially enhances the model's generalisation to new sessions. Notably, the improvement in accuracy and F1-score as the augmentation factor increases suggest that the data augmentation is effective and instrumental during finetuning for new subject adaptation, likely due to a more comprehensive representation of the underlying EEG signal variability. For the SEED [15] dataset, accuracy improves from 86.81% at baseline (1x) to 91.53% at 5x augmentation, with a similar trend observed in the F1 Score. However, further augmentation to 10x does not yield additional improvements, indicating a saturation point in the benefit from additional sessions. Similarly, for the PhyAat [16] dataset, the accuracy increases from 84.74% at baseline to 87.78% at 5x augmentation, but a slight decrease is observed at 10x augmentation. It could be due to overfitting or noise introduced by excessive augmentation.

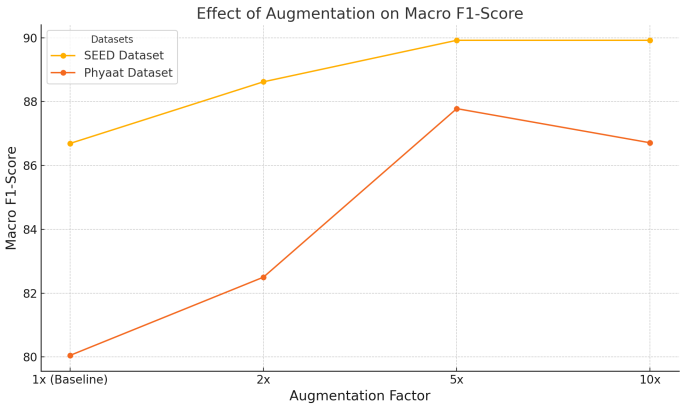



Fig. 6: Impact of IDPM data augmentation on Macro F1-Score for SEED [15] and PhyAat [16] Datasets

## *E. Understanding the relationship between EEG Channels from the learnt model*

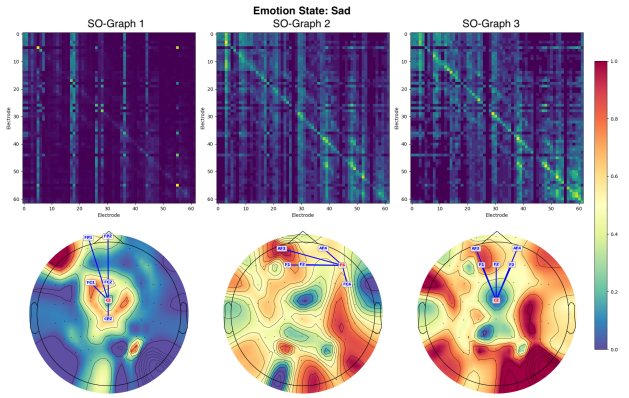
GNNs provide powerful tools for visualizing channel interactions. This study analyze interchannel relationships in the model for SEED dataset [15] Subject 0. It averages and normalizes the adjacency matrices of Self-Organized graphs (SO-graphs 1-3) for each emotional states: sad, neutral, and happy. Using the adjacency matrices, the key channels are calculated and then used to plot the topography maps for each graph.

Figures 7a, 7b, and 7c show adjacency matrix and the topographic maps for each emotional state of Subject 0 of SEED [15]. SO-graph 1 shows diagonal dominance. This indicates that few EEG channels are discriminative in the first graph convolution layer. SO-graphs 2 and 3, however, reveal significant interchannel relationships. These are crucial for classifying different emotional EEG signals. For Sad and Neutral states, prefrontal, centro-parietal, and parietal elec-

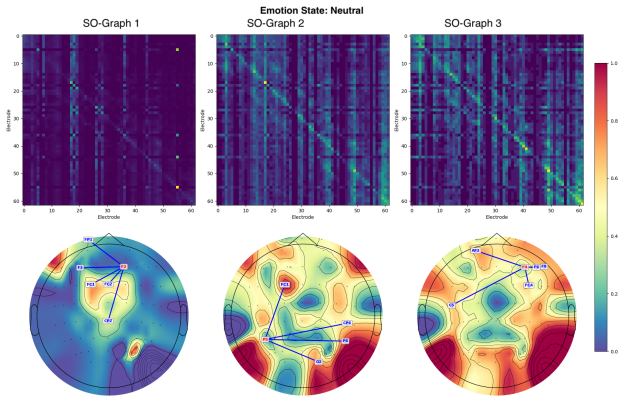
Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:26:00 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Cognitive and Developmental Systems. 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/TCDS.2025.3529669

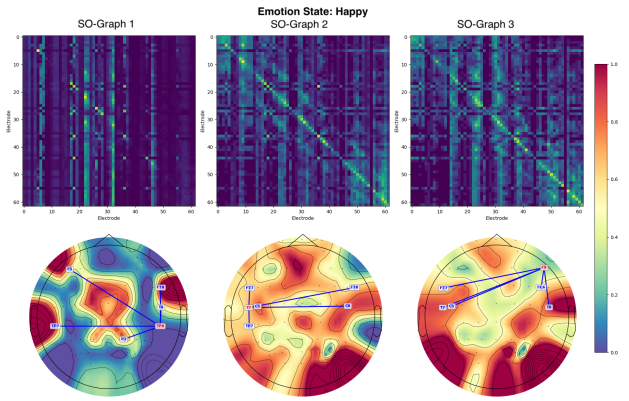





(a) Sad emotional state




(b) Neutral emotional state




(c) Happy emotional state

Fig. 7: Channel Adjacency Matrices and Topographic Maps of each graph learnt in SOGAT for SEED Subject 0 with different emotional states [15]

trodes (e.g., F6, CPZ, P5) have the largest weights. For Happy state, prefrontal, temporal, and temporal-parietal electrodes (e.g., TP8, T7, F8) has the largest weights. These findings align with previous studies [39; 40]. They link prefrontal and centro-parietal activation to emotional processing. These results confirm the importance of discriminative interchannel relationships in EEG-based emotion recognition.

## V. Conclusions and Future Directions

This paper tackles the challenge of EEG signal classification variability across different subjects by introducing the TRANSIT-EEG framework, which incorporates subject-specific modelling, advanced data augmentation, and specialized adapter learning. The framework utilizes the Self-Organizing Graph Attention Transformer (SOGAT) for personalized EEG signal analysis and the Individualized Diffusion Probabilistic Model (IDPM) to generate synthetic data that enhances generalization capabilities. It also includes the Low-Rank Adaptation (LoRA) module, which allows efficient finetuning to new subjects while avoiding catastrophic forgetting, making it preferable for practical applications due to its efficiency and effectiveness.

**Challenges and Future Work:** While IDPM offers end-to-end data augmentation without the explicit need for cleaning the data, this paper does not explore the finetuning of the IDPM for new subjects, a crucial step for maintaining robustness across diverse datasets. When generating synthetic samples from IDPM, there is a high likelihood of similar samples being obtained. Although overall performance increased, devising a mechanism to measure the similarity between signals and prune redundant data to ensure a diverse training set could be useful and worth exploring further. Additionally, this research solely employed the Leave-One-Subject-Out (LOSO) method during finetuning, focusing on a single subject at a time. Theoretically, this approach can be extended to multi-subject finetuning to improve the robustness and generalization of the model; however, this has not been studied in this work and remains an important area for future research. Future work should also assess the effectiveness of the LoRA module in diverse operational environments and with varying degrees of subject variability.

Overall, TRANSIT-EEG offers a comprehensive solution that enhances the accuracy and applicability of cross-subject EEG signal classification in neuroscience and healthcare. By addressing key challenges through a combination of innovative techniques, this research opens new avenues for the practical application of EEG technologies. Future enhancements to the framework will further solidify its utility and broaden its impact in the field.

## Appendix

### *Definition and Purpose*

Individualized Diffusion Probabilistic Model (IDPM) aims to segregate subject-specific noise and clean brain signals from EEG data, essential for applications such as brain-computer interfaces by modelled using the Denoised Diffusion Probabilistic Model (DDPM). A family of generative models that iteratively learns to generate a signal from the noise.

### *List of Notations*

The following symbols are used in the mathematical formulations:

- $x_0, x_1, \ldots, x_T$ - EEG signal at initial state $x_0$ and its noisy states up to $x_T$, the purely noise state.
- $\alpha_t$ - Variance retention coefficient at step $t$ in the diffusion process.
- $\epsilon$ - Standard Gaussian noise.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:26:00 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Cognitive and Developmental Systems. 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/TCDS.2025.3529669



- $\lambda_o, \lambda_r$ - Regularization coefficients for orthogonal and reverse process loss functions, respectively.
- $\mu_\theta(x_t)$ - Mapping of noisy EEG signal $x_t$ to the domain-specific variance space.
- $\mu_\phi(x_t)$ - Mapping of noisy EEG signal $x_t$ to the clean signal space.
- $I$ - Identity matrix.
- $\|\cdot\|_F$ - Frobenius norm.
- $\hat{x}_0$ - Reconstructed EEG signal from the denoised output.
- $s$ - Scaling parameter for the cosine of angles in the Arc-Margin Loss.
- $m$ - Margin added to the angular distance in the Arc-Margin Loss.
- $\theta_{y_i}$ - Angle between the feature vector and the correct class weight.
- $y_i$ - True class label.

*Mathematical Model*

The denoising process involves a diffusion model that incrementally introduces Gaussian noise into an EEG signal $x_0$, culminating in a purely noisy state $x_T$. The model is described as:

$$x_t = \alpha_t x_{t-1} + \sqrt{1-\alpha_t^2}\epsilon, \quad \epsilon \sim \mathcal{N}(0, I)$$

where $\alpha_t$ controls the variance retention at each step.

*Orthogonal and Reverse Process Loss*

*Orthogonal Loss $L_o$:* Orthogonal loss projects domain-specific noise and clean EEG signals into orthogonal subspaces:

$$L_o = \lambda_o \|\mu_\theta(x_t)^\top \mu_\phi(x_t) - I\|_F^2$$

*Reverse Process Loss $L_r$:* This loss function aims to accurately reconstruct the original EEG signal:

$$L_r = \lambda_r \|x_0 - \hat{x}_0\|^2$$

*Separate Subject Domain Variance*

*Dual Stream Generation:* The model uses dual generative paths to optimize for domain-specific variance and invariant content, enhancing adaptability to individual differences.

*Formulating Subject Domain Variance for Human Subjects*

*Arc-Margin Loss $L_{arc}$:* This loss enhances the discriminative power by modifying the angular distance between representations and class labels:

$$L_{arc} = -\log \frac{e^{s\cos(\theta_{y_i}+m)}}{e^{s\cos(\theta_{y_i}+m)} + \sum_{j\neq y_i} e^{s\cos\theta_j}}$$

*Integration of UNet within Denoising Diffusion Probabilistic Models*

Denoising Diffusion Probabilistic Models (DDPM) are a class of generative models that reconstruct data by reversing a diffusion process. This process incrementally adds Gaussian noise to the data until it is indistinguishable from random noise. The task of the generative model is to reverse this noise addition to recover the original data.

*Diffusion and Reverse Process:* The forward diffusion process is defined by a Markov chain where each step incrementally adds noise:

$$x_{t+1} = \sqrt{1-\beta_t}x_t + \sqrt{\beta_t}\epsilon, \quad \epsilon \sim \mathcal{N}(0, I),$$

where $\beta_t$ is a predefined variance schedule, and $\epsilon$ is sampled from a standard Gaussian distribution. The reverse process models the recovery of $x_t$ from $x_{t+1}$ by predicting the noise $\epsilon$ that was added at each step:

$$x_t = \frac{1}{\sqrt{1-\beta_t}}\left(x_{t+1} - \sqrt{\beta_t}\epsilon_\theta(x_{t+1}, t)\right),$$

with $\epsilon_\theta(x_{t+1}, t)$ being the predicted noise by the model parameterized using a neural network.

*UNet Architecture in DDPM:* UNet, an architecture with a symmetric encoder-decoder structure featuring skip connections, is employed to parameterize the noise prediction $\epsilon_\theta$. The architecture's efficacy is due to its ability to maintain spatial hierarchies and detailed feature representation throughout the network layers:

- **Encoder:** Compresses the input data to a compact representation, capturing essential information at each level of noise.
- **Bottleneck:** May include self-attention mechanisms to capture global dependencies, critical for accurate noise prediction in complex patterns.
- **Decoder:** Utilizes the compressed representation to reconstruct data to its less noisy state, with skip connections ensuring the retention of important spatial information.
- **Output:** Produces an estimation of the noise added at each step, used in the iterative denoising during the reverse diffusion process.

*Training and Objective:* The training objective for the UNet within DDPM focuses on minimizing the difference between the actual noise added and the predicted noise:

$$\mathcal{L}(\theta) = \mathbb{E}_{t,x_0,\epsilon}\left[\|\epsilon - \epsilon_\theta(x_t, t)\|^2\right],$$

where $x_t$ is the data at step $t$ of the forward process, and $\epsilon$ is the actual noise added. This training enables the UNet to accurately learn the reverse diffusion dynamics, leading to effective and high-fidelity reconstructions.

## REFERENCES

[1] C. Besthorn, H. Förstl, C. Geiger-Kabisch, H. Sattel, T. Gasser, and U. Schreiter-Gasser, "EEG coherence in alzheimer disease," *Electroencephalography and Clinical Neurophysiology*, vol. 90, no. 3, pp. 242–245, 1994.

[2] K. Värbu, N. Muhammad, and Y. Muhammad, "Past, Present, and Future of EEG-based BCI Applications," *Sensors*, vol. 22, no. 9, p. 3331, 2022.

[3] J. Li, S. Li, J. Pan, and F. Wang, "Cross-subject EEG Emotion Recognition with Self-organized Graph Neural Network," *Frontiers in Neuroscience*, vol. 15, p. 611653, 2021.

[4] W. Hang, W. Feng, R. Du, S. Liang, Y. Chen, Q. Wang, and X. Liu, "Cross-subject EEG Signal Recognition

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:26:00 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Cognitive and Developmental Systems. 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/TCDS.2025.3529669



Using Deep Domain Adaptation Network," *IEEE Access*, vol. 7, pp. 128 273–128 282, 2019.

[5] P. Zhong, D. Wang, and C. Miao, "EEG-based emotion recognition using regularized graph neural networks," *IEEE Transactions on Affective Computing*, vol. 13, no. 3, pp. 1290–1301, 2020.

[6] Y. Li, L. Wang, W. Zheng, Y. Zong, L. Qi, Z. Cui, T. Zhang, and T. Song, "A novel bi-hemispheric discrepancy model for EEG emotion recognition," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 13, no. 2, pp. 354–367, 2020.

[7] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph Attention Networks," in *International Conference on Learning Representations*, 2018. [Online]. Available: https://openreview.net/forum?id=rJXMpikCZ

[8] W. Hang, W. Feng, R. Du, S. Liang, Y. Chen, Q. Wang, and X. Liu, "Cross-subject EEG signal recognition using deep domain adaptation network," *IEEE Access*, vol. 7, pp. 128 273–128 282, 2019.

[9] J. Fan, C. Sun, C. Chen, X. Jiang, X. Liu, X. Zhao, L. Meng, C. Dai, and W. Chen, "EEG data augmentation: towards class imbalance problem in sleep staging tasks," *Journal of Neural Engineering*, vol. 17, no. 5, p. 056017, 2020.

[10] Z. Zhang, Y. Liu, and S.-h. Zhong, "GANSER: A self-supervised data augmentation framework for EEG-based emotion recognition," *IEEE Transactions on Affective Computing*, vol. 14, no. 3, pp. 2048–2063, 2022.

[11] X. Wu, D. Zhang, G. Li, X. Gao, B. Metcalfe, and L. Chen, "Data augmentation for invasive brain–computer interfaces based on stereo-electroencephalography (SEEG)," *Journal of Neural Engineering*, vol. 21, no. 1, p. 016026, 2024.

[12] P. Dhariwal and A. Nichol, "Diffusion models beat gans on image synthesis," *Advances in Neural Information Processing Systems*, vol. 34, pp. 8780–8794, 2021.

[13] S. Torma and L. Szegletes, "Brain Signal Generation and Data Augmentation with a Single-Step Diffusion Probabilistic Model," *arXiv preprint arXiv:2302.03549*, 2023.

[14] T. N. Kipf and M. Welling, "Semi-Supervised Classification with Graph Convolutional Networks," in *International Conference on Learning Representations (ICLR)*, 2017.

[15] "SEED Dataset," http://bcmi.sjtu.edu.cn/ seed/, last accessed on 30th November 2023.

[16] N. Bajaj, J. R. Carrión, and F. Bellotti, "PhyAAt: Physiology of Auditory Attention to Speech Dataset," 2020.

[17] A. Demir, T. Koike-Akino, Y. Wang, M. Haruna, and D. Erdogmus, "EEG-GNN: Graph Neural Networks for Classification of Electroencephalogram (EEG) Signals," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2021, pp. 1061–1067.

[18] A. Demir, T. Koike-Akino, Y. Wang, and D. Erdogmus, "EEG-GAT: A Graph Attention Networks for Classification of Electroencephalogram (EEG) Signals," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2022, pp. 30–35.

[19] N. K. N. Aznan, S. Bonner, J. D. Connolly, N. Al Moubayed, and T. P. Breckon, "Simulating brain signals: Creating synthetic EEG data via neural-based generative models for improved SSVEP classification," 2019.

[20] C. Louizos *et al.*, "A recurrent VAE architecture for capturing temporal dynamics in EEG signals," 2019.

[21] Y. Luo *et al.*, "EEG signal generation using adversarial variational autoencoders for improved realism and diversity," 2020.

[22] R. He and L. et al., "On the Effectiveness of Adapter-Based Tuning for Pretrained Language Model Adaptation," in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Association for Computational Linguistics, August 2021, pp. 2208–2222. [Online]. Available: https://aclanthology.org/2021.acl-long.172

[23] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen, "LoRA: Low-Rank Adaptation of Large Language Models," *arXiv preprint arXiv:2106.09685*, 2021.

[24] Z. K. Hartley, R. J. Lind, M. P. Pound, and A. P. French, "Domain Targeted Synthetic Plant Style Transfer Using Stable Diffusion LoRA and ControlNet," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2024, pp. 5375–5383.

[25] I. Ćulafić, Z. Šćekić, Dejan, Babić, T. Popović, and I. Jovović, "Output Manipulation via LoRA for Generative AI," in *2024 23rd International Symposium INFOTEH-JAHORINA (INFOTEH)*, 2024, pp. 1–4.

[26] Y. Song, L. Yang, X. Jia, and L. Xie, "Common Spatial Generative Adversarial Networks Based EEG Data Augmentation for Cross-subject Brain-Computer Interface," *N/A*, 2021.

[27] Z. Wu, H. Zeng, Y. Zhao, X. Li, J. Zhang, and M. Hattori, "Cross-subject EEG Channel Optimization by Domain Adversarial Sparse Learning Model," in *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2020, pp. 1176–1179.

[28] N. Houlsby, A. Giurgiu, S. Jastrzebski, B. Morrone, Q. De Laroussilhe, A. Gesmundo, M. Attariyan, and S. Gelly, "Parameter-Efficient Transfer Learning for NLP," in *Proceedings of the 36th International Conference on Machine Learning*, ser. ICML, 2019.

[29] A. Q. Nichol and P. Dhariwal, "Improved Denoising Diffusion Probabilistic Models," in *Proceedings of the 38th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, M. Meila and T. Zhang, Eds., vol. 139. PMLR, 18–24 Jul 2021, pp. 8162–8171. [Online]. Available: https://proceedings.mlr.press/v139/nichol21a.html

[30] H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto,

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:26:00 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Cognitive and Developmental Systems. 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/TCDS.2025.3529669



X. Han, Y.-W. Chen, and J. Wu, "UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 1055–1059.

[31] S. Micheloyannis, E. Pachou, C. Stam, M. Breakspear, P. Bitsios, M. Vourkas, S. Erimaki, and M. Zervakis, "Small-World Networks and Disturbed Functional Connectivity in Schizophrenia," *Schizophrenia Research*, vol. 87, no. 1-3, pp. 60–66, 2006.

[32] A. Petrosian, D. Prokhorov, W. Lajara-Nanson, and R. Schiffer, "Recurrent Neural Network Based Prediction of Epileptic Seizures in Intra-and Extracranial EEG," *NeuroImage*, vol. 16, no. 4, pp. 1120–1126, 2000.

[33] Y. Varatharajah, B. Berry, J. Cimbalnik, V. Kremen, J. V. Gompel, M. Stead, B. Brinkmann, and G. Worrell, "Identifying Surgical Targets in Temporal Lobe Epilepsy by Leveraging Multi-Center Data and Machine Learning," *Epilepsy Research*, 2017.

[34] J. Cai, R. Xiao, W. Cui, S. Zhang, and G. Liu, "Application of Electroencephalography-Based Machine Learning in Emotion Recognition: A Review," *Frontiers in Systems Neuroscience*, vol. 15, 2021. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnsys.2021.729707

[35] R. Duan, J. Zhu, and B. Lu, "Differential Entropy Feature for EEG-based Emotion Classification," in *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, 2013, pp. 81–84.

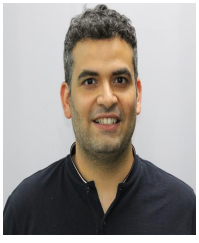
[36] W. Ren, X. Li, L. Wang, T. Zhao, and W. Qin, "Analyzing and Reducing Catastrophic Forgetting in Parameter Efficient Tuning," 2024.

[37] S. V. Eeckt and H. V. hamme, "Using Adapters to Overcome Catastrophic Forgetting in End-to-End Automatic Speech Recognition," 06 2023, pp. 1–5.

[38] S. V. Eeckt *et al.*, "Using Adapters to Overcome Catastrophic Forgetting in End-to-End Automatic Speech Recognition," *arXiv preprint arXiv:2203.16082*, 2022.

[39] C. M. Tyng, H. U. Amin, M. N. M. Saad, and A. S. Malik, "The Influences of Emotion on Learning and Memory. Front Psychol. 2017; 8: 1454," 2017.

[40] P. J. Lang and M. M. Bradley, "Emotion and the Motivational Brain," *Biological Psychology*, vol. 84, no. 3, pp. 437–450, 2010.



**Chirag Ahuja** Chirag Ahuja received the B.Tech. + M.Tech degree in electronics and communication engineering with specialization in intelligent systems and robotics from School of ICT, Gautam Buddha University, Greater Noida, India in 2015. Since 2015, he has been working as an applied scientist in IT industry and has developed various large scale AI and data products. Currently, he is pursuing his PhD from Department of Computer Science and Engineering, Delhi Technological University (DTU), Delhi, India. His area of research is around the physiological signal processing and deep learning.



**Divyashikha Sethia** Divyashikha Sethia received the B.Tech. degree in computer science and engineering from The Maharaja Sayajirao University of Baroda, India, in 1997, and the M.Tech. degree in computer science and engineering from the Indian Institute of Technology Delhi, India, in 2006, with a focus on distributed systems and mobile computing. She has over five years of industry experience in the software development sector. Since 2010, she has been an Assistant Professor with the Department of Computer Science and Engineering, Delhi Technological University (DTU), Delhi. She is the author of more than 13 journal publications and 50 conference papers. Her research interests include telemedicine, security in distributed systems, and the application of brain-computer interfaces in healthcare.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:26:00 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

Check for updates

*Applied Organometallic Chemistry*

WILEY

Applied Organometallic Chemistry

RESEARCH ARTICLE

# Ultrasonic-Assisted Adsorptive Uptake of Malachite Green Dye Onto Zn/Cu-Layered Double Oxide: Spectrophotometric and Smartphone-Based Detection

Bhamini Pandey[1] | Poonam Singh[2] | Ravinder Kumar[3]

[1]Department of Chemistry, GLA University, Mathura, Uttar Pradesh, India | [2]Department of Applied Chemistry, Delhi Technological University, Delhi, India | [3]Department of Chemistry, Gurukula Kangri (Deemed to be University), Haridwar, Uttarakhand, India

**Correspondence:** Poonam Singh (poonam@dtu.ac.in) | Ravinder Kumar (ravinder.kumar@gkv.ac.in)

**Received:** 24 July 2024 | **Revised:** 3 December 2024 | **Accepted:** 18 December 2024

**Funding:** The authors are grateful to UGC-BSR for funding this project (30-416/2018[BSR]). In addition, BP would like to thank CSIR, India, for the financial support it provided.

**Keywords:** layered double oxide | malachite green | smartphone-based colorimetry | ultrasound-assisted adsorption

## ABSTRACT

In this work, efforts have been made to synthesize novel Zn/Cu-layered double oxide by performing calcination of Zn/Cu-layered double hydroxide, which was later used as an adsorbent for the uptake of toxic azo dye. Malachite green (MG) was used as a model cationic dye for assessing the practicability of dye wastewater treatment. From batch experiments study, it was evident that ultrasonication-assisted adsorption enables 90% of the MG to be removed from water within 30 min at neutral pH conditions. The ultrasound-assisted sorption process was driven by the acoustic cavitation process, which facilitated a homogeneous distribution of active adsorption sites on surface of the adsorbent resulting efficient removal of dye molecules. As predicted by the Langmuir isotherm model, the sorbent exhibited maximum adsorption capacity of 128.20 mg/g, and based on kinetic measurements, the pseudo-second-order adsorption model was found to be most accurate. The kinetic studies also revealed that intraparticle diffusion is the rate-limiting step of the adsorption process. The current study also demonstrates smartphone-assisted estimation of adsorbed MG dye in order to provide a simple, convenient, cost-effective, and instrument-free method to quantify dye concentration. A smartphone-based app called "ColorPicker" was used to calculate MG dye adsorption rate in real time based on RGB values and colorimetric responses. The percent removal obtained using smartphone app (87%) was in agreement with the data obtained using UV–vis spectroscopy (90%). As a result of its excellent properties, Zn/Cu-LDO appears to be a potential candidate for wastewater remediation.

## 1 | Introduction

Human survival and marine life depend heavily on water, which is one of the most vital natural resources. However, rapid industrialization in modern world contributes to the release of numerous toxic and harmful pollutants into the environment. The toxic waste generated from paper, textile, leather, food, and pharmaceuticals industries are discharged into water resources which are carcinogenic, mutagenic, and teratogenic, thereby causing serious health and ecological risks [1]. Therefore, to improve the quality of life, there is an urgent need to eliminate toxic wastes from the environment. Various toxic organic/inorganic pollutants are discharged into the environment among these, dyes are complex organic pollutants. The strong adhesion, variable colors, and high stability of anionic and cationic dyes have made them popular for dyeing textile materials [2]. Malachite green (MG) dye is among the commonly utilized cationic dyes in textile, lather, and paper industries. It has been

© 2025 John Wiley & Sons Ltd.

used as food coloring agents as well as pesticides in agriculture applications. Although despite having various applications several groups have reported the toxicity of these dyes. MG dye molecules are absorbed by plant roots, which affects the metabolism of plants. In addition, MG dye is extremely toxic to aquatic organisms due to its strong chelating tendency to metal ions. It is a toxic, mutagenic, carcinogenic contaminant, which also tends to enter the food chain ultimately causing toxic effects on human cells. Therefore, MG dye effluent must be removed or treated to reduce its toxic impact on the ecosystems [3].

Dye effluents have been treated using a diverse range of wastewater treatment methodologies, including coagulation, sorption, membrane filtration, ion exchange, catalysis, and photodegradation [4]. However, the high production of sludge, fouling of membranes, and ineffectiveness for removal of all the dyes are some of the disadvantages of prevalent water treatment methods. A comparison of adsorption method to other wastewater treatment techniques has revealed that it offers greater versatility and superiority owing to its high potential, cost-effectiveness, ease of operation, and eco-friendliness [5]. Moreover, adsorption stands out as an environmentally friendly method with low energy consumption and the potential for adsorbent regeneration, making it particularly suitable for dye removal applications [6]. Adsorption is primarily determined by the surface area of the material, the affinity of the intraparticle interaction, and its strength. In spite of the fact that sorptive removal of industrial dyes is complex, its economic feasibility has made it widely accepted around the world [7]. It is predominantly concerned with developing and selecting new adsorbent materials with high efficiency [8]. Various sorbent materials have been used for the purification of water such as activated charcoal, bioadsorbents, silicates, and polymeric adsorbents. These adsorbents, however, have a limited application owing to high cost, poor adsorption efficiency, and difficulty in regeneration.

During recent years, the layered materials have attracted immense attention from the researchers as a sorbent for the sequestration of organic/inorganic contaminants [9]. LDH are bidimensional materials having general formula $[(M)^{II}_{1-y.}\ M^{III}_{y}(OH)_2]\ (A^{x-})_{y/x}\ .mH_2O$, where $M^{II}$ are divalent cations (e.g., $Zn^{2+}$, $Co^{2+}$, and $Mg^{2+}$), $M^{III}$ are trivalent cations (e.g., $Fe^{3+}$, $Cu^{3+}$, and $Ln^{3+}$), $A^{x-}$ is the charge compensating mobile anion trapped in the interlayer region (e.g., $CO_3^{2-}$ and $NO_3^{-}$), and y is the ratio of $(M^{3+})/(M^{2+}+M^{3+})$ [10]. Availability of high surface area, tunable properties, and ability to exchange anion makes them a desirable material. The treatment of LDH at optimum temperature results in the topotactic transformation to mixed metal oxide (MMO) or layered double oxide (LDO). The LDO obtained upon calcination allows uniform distribution of +2 and +3 metal cations. In case of LDO, the layered structure can be regained by providing similar conditions which is known as the memory effect property that also allows various anions to be intercalated into the layered structure upon rehydration [11]. The LDO can act as catalytic material in various organic transformations as well as adsorbent for removal of pollutants. Claydon et al. performed the hydrogenation of naphthalene using molybdenum-doped Ni/Al MMO and compared its efficiency with Pd-doped aluminum oxide catalyst [12]. Mubarak et al. demonstrated that Mg/Fe-LDO exhibited high adsorption efficiency for the uptake of Congo red and methyl blue dyes [13].

The combination of adsorption technique with other processes like ultrasonication results in increased efficiency of adsorbent. Sonication is a method of agitating particles in a sample by applying sound energy, and the process is called ultrasonication when ultrasonic frequencies are used. The phenomenon of acoustic cavitation is the driving force of the ultrasonication process where the formation and collapse of microbubbles formed near the ultrasound source intensify the mass transfer and breaks the adsorbent–adsorbate affinities [14]. Thus, the combined effect of the ultrasound-assisted adsorption process may result in shorter reaction time and improved efficiency. Separation and purification of contaminants are achieved with ultrasonic-assisted adsorption in research laboratories, industries, and water treatment plants [15]. Acisli et al. revealed that the combination of adsorption process with ultrasonication enhanced the removal efficiency of modified montmorillonite for AR17 dye [16]. Another study by Deb et al. also showed that ultrasound enhanced the adsorptive removal of anionic dyes by hematite/polyaniline nanocomposite [17].

As far as removing dye effluents from water is concerned, two main challenges are faced. The first one is the ease of preparation of efficient adsorbent, and the second is the usage of accessible, easily operated, and convenient method for the detection of dye uptake. An inexpensive, field-portable, and user-friendly platform based on RGB color modeling can be utilized for performing analytical tests, and its ease of use allows for effortless quantification of dye uptake, which has opened new horizons for data analysis. In recent years, smartphones have grown in popularity as a laboratory tool for chemical analysis and biological experiments (as biomarkers) [18]. The advent of smartphones with high-resolution cameras and large storage capacities allows users to capture clear and detailed pictures. Smartphones can be used for the colorimetric detection of organic/inorganic effluents depending upon the color received from the analyte solution. In a recent study by Bahrami et al., mercury ions were removed using Fe-Cr LDH and detected using phone-based colorimetry, and results were in agreement with those obtained using UV–visible spectroscopy [19].

Despite significant advances in wastewater treatment technologies, the efficient removal of toxic dyes, such as MG, remains a challenge due to the limitations of conventional adsorbents in terms of cost, efficiency, and environmental sustainability. This study presents a novel Zn/Cu-LDO adsorbent, synthesized via calcination of Zn/Cu LDH, as a promising solution for dye removal. The ultrasonication-assisted adsorption process demonstrated exceptional efficiency under neutral pH conditions, achieving 90% removal of MG within 30 min. The impacts of various factors including pH, time of ultrasonication, concentration of dye, and adsorbent dosage have also been carried out. To propose a system that is applicable under all conditions, adsorption isotherms and kinetic models were examined. A blank experiment was also conducted without the sorbent to study the influence of ultrasonication on dye uptake behavior. Additionally, the study introduces a smartphone-assisted approach using the "ColorPicker" app for real-time monitoring of dye adsorption, providing a cost-effective, portable, and accessible method for quantifying dye concentration without the need for sophisticated instrumentation. The comparison between the results obtained using



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7981 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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 smartphone app and traditional UV–vis spectroscopy, showing close agreement (87% vs. 90%), further validates the practicality of this method. The integration of these innovative methodologies not only enhances the practicability of dye removal but also expands the scope of this research by offering a versatile platform for the development of low-cost, user-friendly monitoring tools for broader environmental applications. This work lays the foundation for future studies to explore the applicability of Zn/Cu-LDO in treating a variety of contaminants and underscores the potential of integrating smart technologies in environmental monitoring.

## 2 | Experimental Details

### 2.1 | Chemical and Reagents

Zinc (II) acetate dihydrate (99.5%, Merck), copper (II) acetate monohydrate (≥99.0%, Merck), and hydrogen peroxide were procured from Merck (Germany). The MG dye was purchased from CDH and was dissolved in 500 mL of deionized (DI) water to prepare the stock solution (100 mg $L^{-1}$).

### 2.2 | Synthesis of Zn/cu-LDO

First, Zn/Cu LDH was prepared by following the hydrolysis method outlined in our previously reported work [20], where the corresponding metal salt solutions were prepared using 0.9877 g of Zn $(CH_3COO)_2.2H_2O$ and 0.0998 g of Cu $(CH_3COO).H_2O$ in DI water. After that, the hydrogen peroxide solution (5 mL of 30% $H_2O_2$ + 30 mL DI water) was added to the above metal salt solutions. Later resultant solution was ultrasonicated for 30 min followed by 24 h of stirring. Finally, the resultant mixture was dried at 60°C in the oven. The above synthesized Zn/Cu LDH (0.5 g) was then calcined at 350°C for 6 h in a muffle furnace to obtain Zn/Cu-LDO. The schematic representation of Zn/Cu-LDO synthesis is depicted in Figure 1.

### 2.3 | Characterization

A Bruker D8 advance diffractometer operating at Cu K radiations at a step size of 0.02° was used in the $2\theta = 5°–70°$ range to record powder X-ray diffraction patterns at room temperature. The sample was analyzed using a PerkinElmer Fourier-transform infrared spectrometer (version 10.5.3). Cary Series (Cary 300) UV–Visible Spectrophotometer was used to analyze the sample concentration. Nitrogen adsorption–desorption isotherm was measured using a Quantachrome, USA, Novatouch $LX^4$ to determine the Brunauer–Emmett–Teller (BET) surface area at 77 K. A Zeiss GeminiSEM was used for FE-SEM with EDX analysis of Zn/Cu-LDO. XPS measurements were recorded with Al Kα photoelectron excitation using a PHI 5000 VersaProbe III. Transmission electron microscope (TEM) and selected area electron diffraction (SAED) were performed with a Talos f200x microscope.

### 2.4 | Ultrasound-Assisted Adsorption Studies

An ultrasonic bath having 40 ± 3 KHz ultrasonic frequency and 150-W ultrasonic wattage (main supply: 220 V, AC 50 Hz, single phase) was used to conduct adsorption experiments using 0.2 g/L of Zn/Cu-LDO at different concentrations (10–40 mg/L) of MG dye for 30 min in the temperature range of 30°C–35°C. After adsorption, the sample was centrifuged for 10 min, and a UV–visible spectrophotometer at the lambda maximum of 617 nm (for MG dye) was utilized to determine the unknown concentration of dye. The removal efficiency and adsorption capacity were estimated using Equations (1) and (2), respectively.

$$Removal\ efficiency\ (\%) = \frac{Co - Ce}{Co} \times 100\% \quad (1)$$

$$q_e = \frac{C_o - C_e}{W} X\ V \quad (2)$$

$$q_t = \frac{C_o - C_t}{W} X\ V \quad (3)$$




**FIGURE 1** | Synthesis of Zn/Cu-LDO.



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7981 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

where $C_o$, $C_e$, and $C_t$ are dye concentrations at the initial, equilibrium, and final stage, respectively. $q_e$ and $q_t$ are the adsorption capacities at the equilibrium and time $t$, respectively. $V$ and $w$ are the volume of solution and amount of the adsorbent, respectively.

## 3 | Results and Discussion

### 3.1 | Characterization of Synthesized LDO

The PXRD patterns of both Zn/Cu LDH and Zn/Cu-LDO are given in Figure 2a. The formation of LDH is corroborated by the presence of a sharp diffraction peak at 2$\theta$ values of 11.96° followed by low intense reflections at higher 2$\theta$. However, upon calcination, prominent reflections of LDH disappear in the PXRD pattern indicating the successful conversion of LDH to LDO. The characteristic peaks of LDO were observed at 31.99°, 34.64°, and 36.43° with a slight shift to lower angles from pure ZnO because of the presence of copper ion in the lattice. The average particle size of synthesized Zn/Cu-LDO calculated from Debye Scherrer formula was found to be 53 nm [21].

The FTIR spectra of the Zn/Cu LDH and LDO (Figure 2b) suggested the removal of interlayer acetate ion on calcination of LDH. The spectra of LDH and LDO exhibit a broad band at 3418 and 3382 $cm^{-1}$ corresponding to the existence of hydroxyl groups. The strong bands corresponding to acetate ions of the interlamellar region of LDH cannot be observed in the spectrum of Zn/Cu-LDO due to the removal of interlayer ions upon calcination. The bands in the range of 600–400 $cm^{-1}$ are assigned to the metal oxygen (M—O and M—O—H) linkages present in the lattices. In addition, the disappearance of a broad band around 3000 $cm^{-1}$ due to the hydrogen bonding in LDH also corresponds to the removal of interlayer water molecules and acetate ions [22].

SEM/EDX analysis was performed on prepared LDO for determining its morphology and composition, as shown in Figure 2c–e. The average particle size was also calculated from SEM image and was found to be 58 nm. The EDX analysis demonstrated the homogeneous distribution of zinc, copper, and oxygen elements. The obtained results confirmed the successful synthesis of Zn/Cu-LDO.

The morphology of the synthesized material was examined using TEM imaging, which revealed multiple shapes including hexagonal, rod-like, spherical, and irregular shapes (Figure 2f). The figure also reveals some degree of agglomeration along with the polycrystallinity due to the presence of distinct boundaries among nanoparticles. The SAED pattern of Zn/Cu-LDO is presented in Figure 2g, which shows the distinct diffraction patterns corresponding to (100), (002), (101), (102), (110), (200), and (203) diffractions lines of zinc oxide. Moreover, the presence of multiple diffraction rings also confirms the polycrystallinity of the synthesized material. This pattern confirms that there occurs a structural consistency with PXRD data because the crystallographic planes observed in XRD analysis align well with those observed in SAED pattern.

From XPS data, oxidation states and elemental distributions of ions in the lattice were estimated. Zn/Cu-LDO predominantly comprised of Zn, Cu, O, and C, based on the survey spectrum (Figure 3a). The O *1s* spectrum consists of three peaks at 529.92, 530.21, 531.90 eV corresponding to the presence of M—O and M—OH bonds and absorbed $H_2O$ molecules, respectively (Figure 3b) [23]. In the core level spectrum of Zn *2p* (Figure 3c), two prominent peaks can be seen at 1021.27 and 1044.37 eV, which confirmed that Zn (II) was present within the lattice. The high-resolution spectrum of copper *2p* (Figure 3d) exhibited peaks of binding energies of 932.59 (Cu $2p_{3/2}$) and 952.37 (Cu $2p_{1/2}$) eV, confirming copper presence in the sample. From XPS analysis, it is clear that the Zn/Cu-LDO sample consists of bimetal mixed oxides, which is in agreement with the aforementioned EDS analysis.

By analyzing the surface of the adsorbents, the size and pore volume, and the specific surface area were assessed. The data for surface properties of prepared LDO in comparison with parent LDH are given in Table 1 [10]. In accordance with the IUPAC classification system, the nitrogen adsorption–desorption isotherm falls under type IV category with $H_4$ hysteresis loop (Figure 4a), which indicates microporous nature of material (because the pore size is below 2 nm). Pore volumes of Zn/Cu-LDO (0.122 $cm^3/g$) are greater than that of pristine LDH (0.030 $cm^3/g$), suggesting that the collapse of layered arrangement after calcination (Figure 4b). Moreover, the removal of acetate ions and $H_2O$ molecules from the interlamellar region resulted in an increased surface area of LDO in comparison with LDH. It can also be summarized that the presence of interlayer anion in LDH resulted in smaller and larger pore volume and diameter, respectively, than that of Zn/Cu LDH [24]. It can be stated that the large pore diameter and abundance of micropores are also the major contributors to the fast diffusion of MG dye molecules.

BET analysis also revealed that, following adsorption, the specific surface area of the adsorbent decreased significantly. This reduction can be attributed to the blockage of surface-active sites of the adsorbent by the dye molecules, which limits the availability of the active sites for further adsorption [25]. Furthermore, the adsorption of dye molecules also leads to a decrease in both BJH pore diameter and pore volume of the adsorbent upon adsorption, as illustrated in Table 1. This reduction advocates that there is penetration of dye molecules and occupation of pores of the adsorbent, thus effectively blocking the active sites of the sorbent. The BET analysis results also imply that the MG dye adsorption involves two key mechanisms; first is surface adsorption, which occurs due to the binding of adsorbate molecules with the active sites of Zn/Cu-LDO, and second is the pore adsorption mechanism responsible for the diffusion of dye molecules into the internal pore network of the sorbent.

### 3.2 | Uptake of MG dye

#### 3.2.1 | Influence of Sonication Time

Figure 5a,b presents the influence of sonication time on removal efficiency of synthesized sorbent at 30 mg/L concentration of MG dye and 0.2 g/L of adsorbent dosage. A significant and sharp increase in dye uptake can be observed in the initial 5 min of



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7981 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

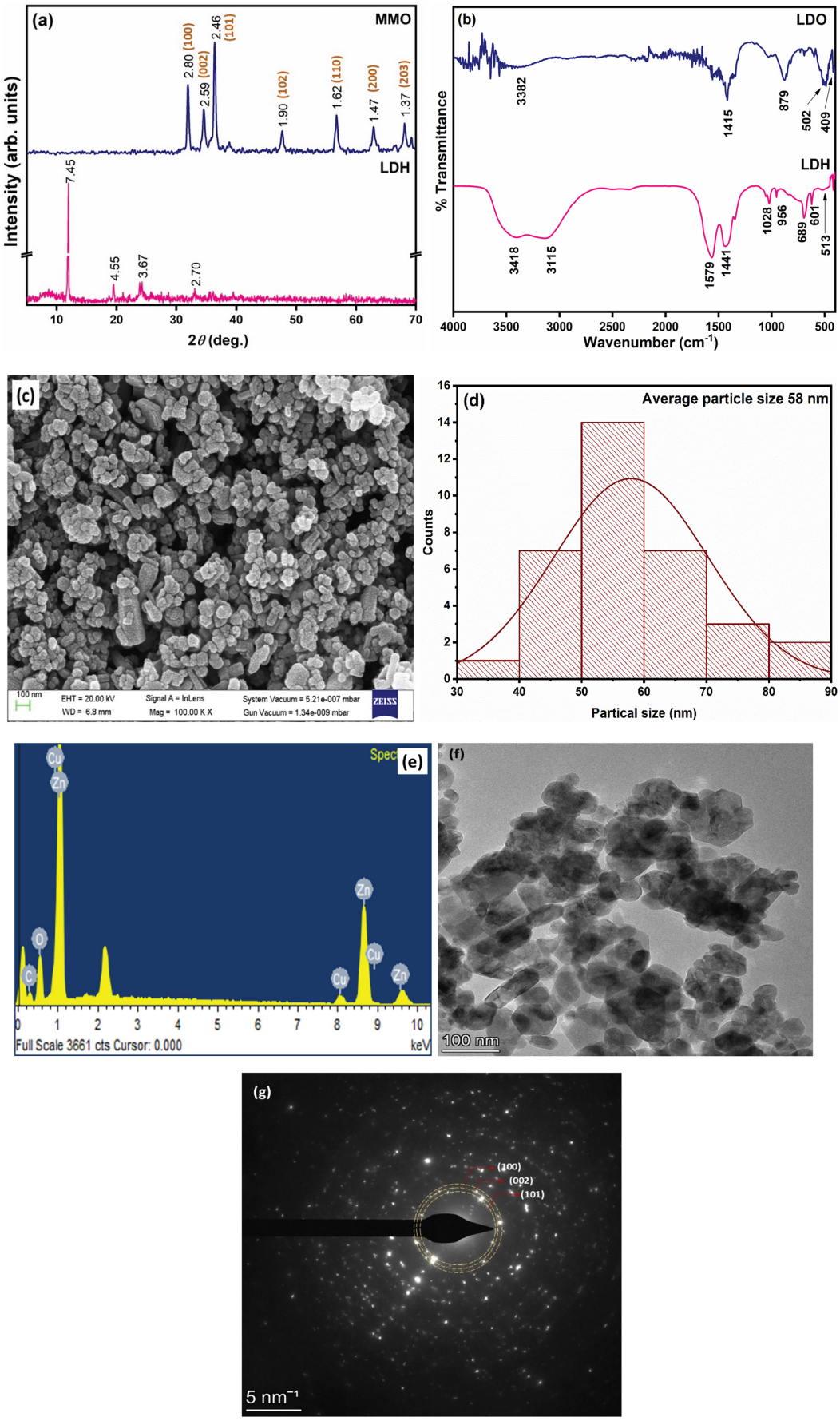



**FIGURE 2** | (a) PXRD pattern, (b) FTIR spectra of LDH and LDO, (c) SEM image, (d) average particle size distribution histogram, (e) EDX plot, (f) TEM image, and (g) SAED pattern of Zn/Cu-LDO.



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7981 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

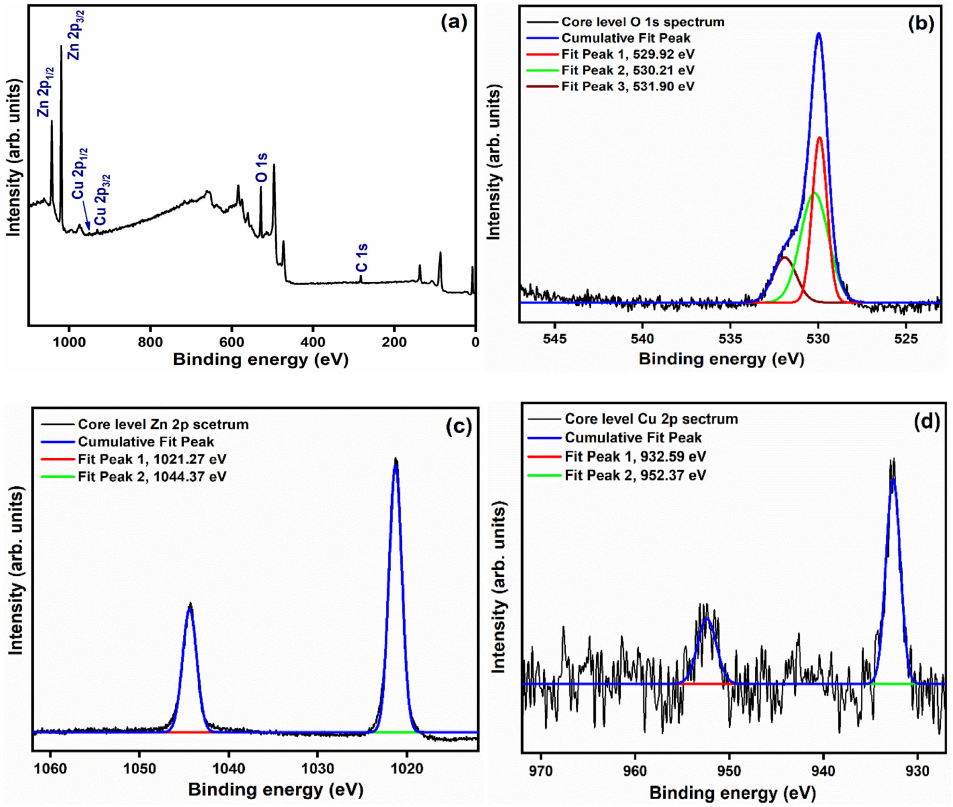



**FIGURE 3** | (a) XPS survey spectrum and high resolution XPS spectra of (b) O *1s*, (c) Zn *2p*, and (d) Cu *2p*.

**TABLE 1** | Surface properties of synthesized adsorbent.

| Adsorbent | BET surface area ($m^2/g$) | Pore diameter (nm) | Pore volume ($cm^3/g$) |
|---|---|---|---|
| Zn/Cu LDH | 61.841 | 3.658 | 0.030 |
| Zn/Cu-LDO (before adsorption) | 84.217 | 1.171 | 0.122 |
| Zn/Cu-LDO (after adsorption) | 13.567 | 0.547 | 0.068 |

ultrasonic irradiation, that is, reaching > 85% after which equilibrium was achieved gradually. The adsorbent's large surface area and active sites contribute to a high initial adsorption rate. In addition, a large amount of mass is transferred from the adsorbate to the adsorbent during sonication which contributes to the high MG adsorption.

### 3.2.2 | Influence of Concentration of MG

For the purpose of examining the influence of the initial MG concentration on the removal efficiency of the Zn/Cu-LDO, experiments were performed with different MG concentrations (Figure 6a). After 30 min of ultrasonically treating MG dye solutions, a specific amount of adsorbent was added to different concentrations of dye (5–30 mg/L). According to the findings of the study, a higher initial dye concentration resulted in increased adsorption of MG dye at equilibrium. When concentrations are high, molecular collisions are more likely to occur with adsorbent surfaces, leading to greater transfer of masses from solution to solid. In response to increasing collisions, molecules are removed from the solution phase at a faster rate.

### 3.2.3 | Influence of pH and $pH_{PZC}$ Determination

The surface charge of sorbent is greatly influenced by pH of solution, which is important for the sequestration of pollutants the aqueous phase. It is notable that detection of MG dye is difficult above pH 8 due to the interaction of hydroxy ions and cationic dye molecules resulting in the precipitation of MG, which discolors the dye solution [26]. In addition, MG dye is unstable and decolorizes at high pH; therefore, the impact of pH on the uptake of dye was studied in 3 to 7 pH range where other parameters were kept constant. There is an increased amount of hydrogen ions at low pH conditions, resulting in competition between the positively charged dye molecules for sorption sites on the sorbent. A correlation exists between dye removal efficiency and pH value (Figure 6b), which is likely to be attributed to the negative surface charge of the adsorbent at high pH conditions, interacting with the positive dye molecules.

The $pH_{pzc}$ was estimated to be 5.86 suggesting that adsorbent's surface is positive and negative below and above 5.86 pH value, respectively (Figure 6c). A negative charge on adsorbent surface and a positive charge on the dye molecules result in a strong electrostatic interaction above pH 5.86. As a result, above pH 5.86, there occurs an electrostatic interaction between adsorbent surface (negatively charged) and dye molecules (positively



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7981 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

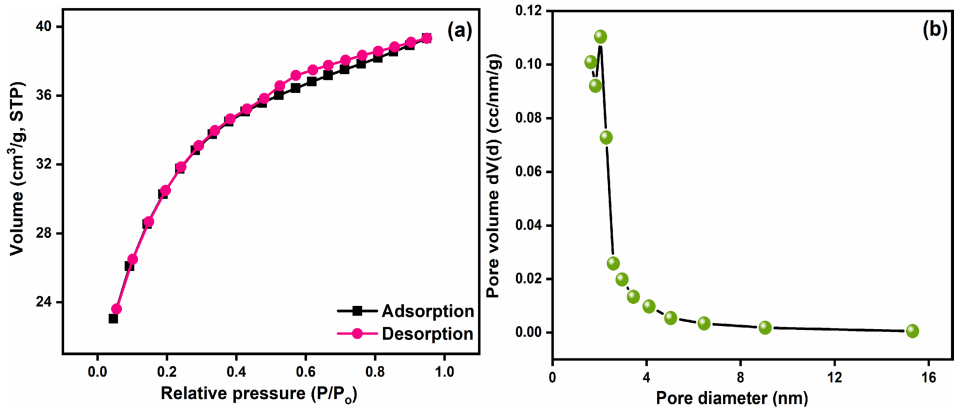



**FIGURE 4** | (a) $N_2$ adsorption–desorption isotherm and (b) pore size distribution plot of Zn/Cu-LDO.

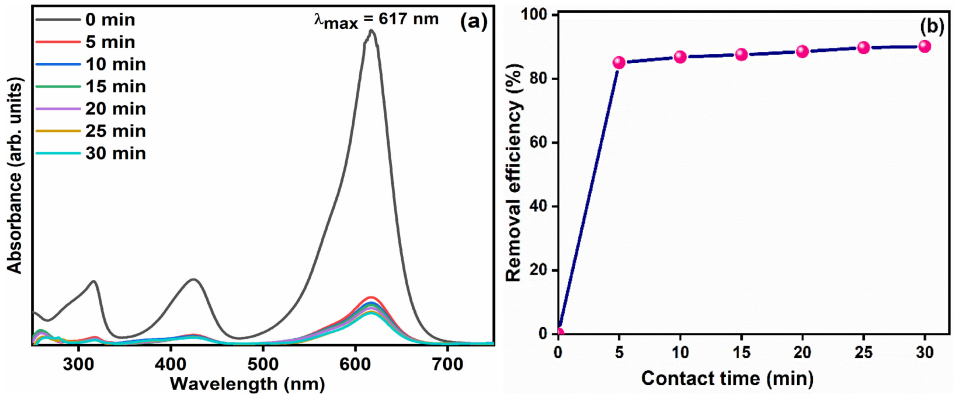



**FIGURE 5** | (a) Wavelength versus absorbance and (b) contact time versus percentage of dye removal plots.

charged). Additionally, LDO has the advantage of being able to carry out the adsorption process at a pH range that is compatible with drinking water.

### 3.2.4 | Influence of Adsorbent Amount

There is a direct correlation between the amount of adsorbent in aqueous solutions and the rate of contaminant removal. As depicted in Figure 6d, the MG removal efficiency increased from 74% to 90% as the dosage rate increased from 0.1 to 0.4 g $L^{-1}$. As the adsorbent dosage is increased, more active sites are available for adsorption thus, resulting in increased removal efficiency.

### 3.2.5 | Influence of Competitive Ions

A significant amount of inorganic salts such as sodium chloride, potassium chloride, and calcium chloride are commonly used during textile dyeing in order to help the dye fixation process. However, the adsorption performance of the sorbents can be significantly influenced by the composition of the dye solution, including the presence of inorganic salts or competitive ions [27]. Because MG is a cationic dye, the presence of ions like $Na^+$, $K^+$, and $Ca^{2+}$ may also compete for the active sites of the adsorbent. This competition results in decreased adsorption efficiency of sorbent for the cationic dye. In order to assess the stability or selectivity of sorption performance, the adsorption capacity of MMO was tested in the presence of competing ions (having 0.01-M concentration of NaCl, KCl, and $CaCl_2$) [28]. As expected, a decrease in removal efficiency of MMO was observed in the presence of interfering ions. The impact of competitor ions on removal efficiency follows the series $Na^+$ (89%) < $K^+$ (87%) < $Ca^{2+}$ (72%) (illustrated in Figure 6e).

The observed trend can be attributed to the electrostatic shielding caused by competitive cations, which reduces the electrostatic attraction between dye molecules and the adsorbent. Moreover, the large positive charge and high polarizability of $Ca^{2+}$ ions enhance their ability to shield the adsorbent's active sites more effectively than $Na^+$ and $K^+$ ions. This phenomenon advocates that high charge density of the interfering ions is the major contributor to the reduced removal efficiency of MMO. The observed reduction highlights the importance of considering the presence of salts or competing ions in real-world applications, as these can significantly impact the adsorbent's performance and efficiency [29].

### 3.2.6 | Kinetic Study

Adsorption kinetics plays a substantial role in predicting the sorption rate in batch adsorption processes. Adsorption kinetic study was conducted to understand possible adsorption mechanisms and predict adsorption rates. The adsorbents' physical and chemical properties are crucial to chemical kinetic modeling as they give valuable information about the factors that



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7981 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

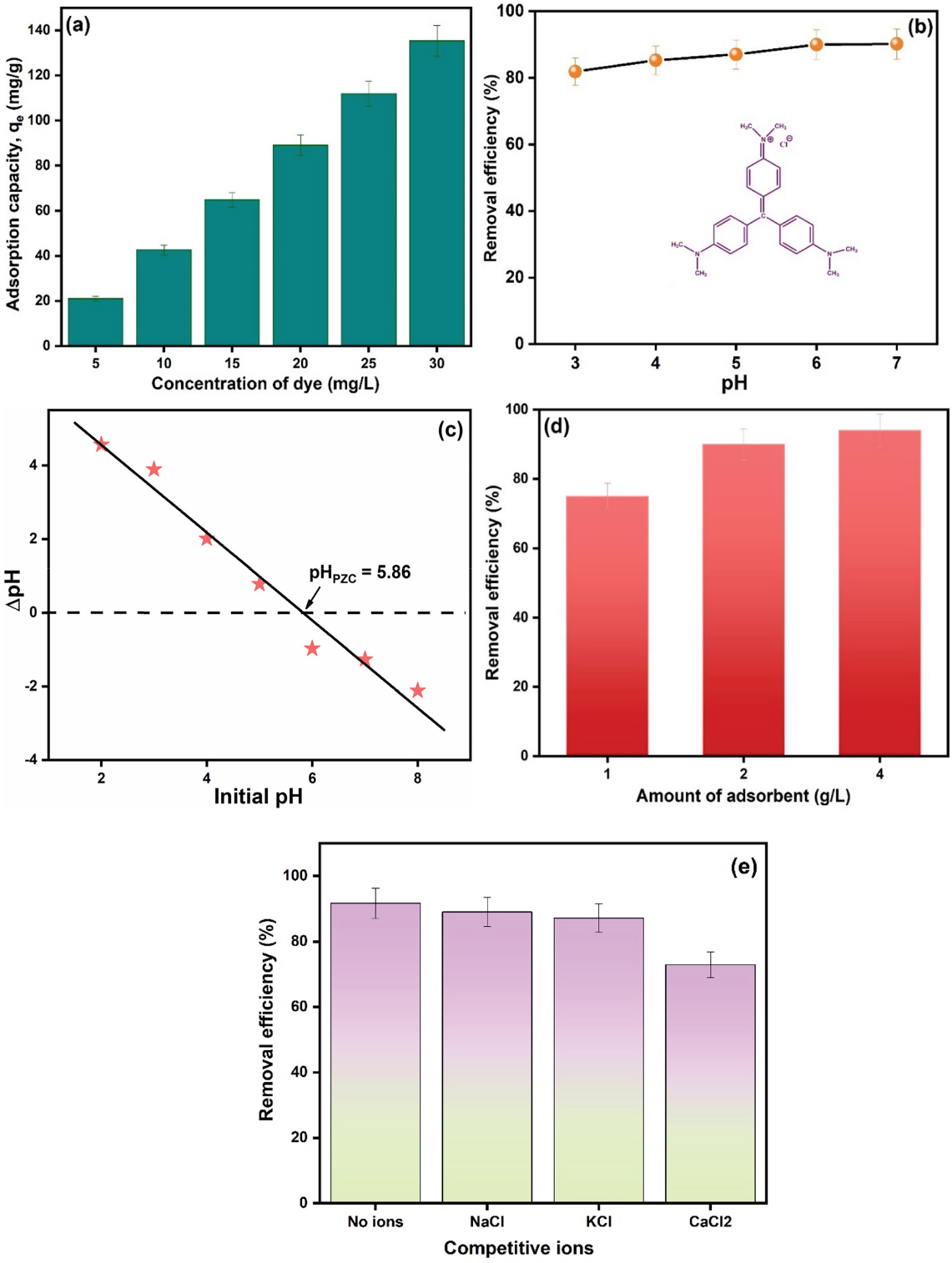



**FIGURE 6** | Plot of (a) concentration variation of MG dye versus adsorption capacity, (b) % removal versus pH of dye solution, (c) $pH_{pzc}$ versus ΔpH, (d) effect of adsorbent amount on MG dye sequestration, and (e) effect of competitive ions on removal efficiency.

affect reaction rates. In the present work, pseudo-first-order (PFO), pseudo-second-order (PSO), and Elovich models were employed to simulate the experimental kinetic parameters of the adsorbent.

The general expression of the linearized form of PFO is given as

$$ln(q_e - q_t) = \ln q_e - K_1 t \quad (4)$$

where $q_e$ ($mg\,g^{-1}$) and $q_t$ ($mg\,g^{-1}$) are adsorption capacities at equilibrium and at time $t$, respectively and $K_1$ ($min^{-1}$) is the PFO rate constant.

Unlike other adsorption models, the PSO model predicts the adsorption behavior over the entire absorption range because it is based on the supposition that adsorption capacity depends on the available active sites (Equation 5).



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7981 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

$$\frac{t}{q_t} = \frac{1}{K_2 q_e^2} + \frac{t}{q_e} \quad (5)$$

where $K_2$ ($g\,mg^{-1}\,min^{-1}$) is the second-order rate constant that can be estimated from the linear plot of $t/q_t$ versus $t$.

It has been successfully demonstrated that the Elovich equation in linear form can be used to describe a second-order kinetic equation based on adsorption capacity, assuming the solid surfaces are energetically heterogeneous. There is, however, no specific mechanism conveyed in the equation to explain the adsorbate–adsorbent interaction. In general, it is widely agreed that the following semi-empirical equation can accurately describe the chemisorption process:

$$q_t = \beta(ln\alpha\beta) + \beta\, lnt \quad (6)$$

where $\alpha$ ($mg\,g^{-1}\,min^{-1}$) is the adsorption rate at initial stage and $\beta$ ($g\,mg^{-1}$) is the desorption constant. The applicability of intraparticle diffusion model was also studied in order to evaluate the contribution of pore diffusion in the sorption process. The model assumes that sorption occurs in multiple steps and intraparticle diffusion is the rate-controlling step. Weber and Morris formulated the linearized form of intraparticle diffusion model which is given in Equation (7).

$$q_t = k_{id} t^{1/2} + C \quad (7)$$

where $q_t$ (mg/g) denotes dye uptake capacity at time $t$ and $k_{id}$ denotes intraparticle diffusion rate constant. Based on the models, the kinetic parameters and correlation coefficients were estimated and are given in Table 2, and the plots are displayed in Figure 7.

The adsorption capacity evaluated from the PSO kinetic model ($q_e$ calculated) was observed to be 136.99 mg/g, which was in close correlation with the value obtained from UV–vis measurements ($q_e$ experimental = 135.32 mg/g). Based on $R^2$ and $q_e$ values, PSO model better explains the kinetics of adsorption, suggesting chemisorption is involved. Additionally, the linear nature of the $q_t$ versus $t^{1/2}$ plot indicates that intraparticle diffusion plays a significant role in the adsorption process and acts as the rate-limiting step for the sorptive removal of MG dye. The plot passes through the origin, suggesting that intraparticle diffusion contributes to the rate-controlling mechanism. These results suggests that the adsorption process may involve a direct interaction between pores of sorbent and dye molecules.

### 3.2.7 | Isotherm Study

For the purpose of determining adsorbent maximum adsorption capacity and nature of adsorption, Langmuir, Freundlich, and Temkin models were applied to analysis adsorption isotherms. First, Langmuir adsorption isotherm was used to fit the ultrasound-assisted MG adsorption data. In the Langmuir model (Equation 8), adsorbent molecules are assumed to have no interaction with one another adsorption that occurs in homogenous sites on the adsorbent. Furthermore, a finite amount of adsorbate can be absorbed by the adsorbent (at equilibrium), and all adsorption sites have the same energetic properties [30].

**TABLE 2** | Kinetic parameters for MG dye sequestration.

| Kinetic models | Kinetic parameters | | $R^2$ |
|---|---|---|---|
| PFO | $K_1$ ($min^{-1}$) | $q_e$ ($mg\,g^{-1}$) | 0.8938 |
| | 0.14 | 23.07 | |
| PSO | $K_2$ ($g\,mg^{-1}\,min^{-1}$) | $q_e$ ($mg\,g^{-1}$) | 0.9981 |
| | 0.015 | 136.99 | |
| Elovich | $\beta$ ($mg\,g^{-1}\,min^{-1}$) | $\alpha$ ($g\,mg^{-1}$) | 0.9771 |
| | 0.24 | 9.32 | |
| Intraparticle diffusion | $K_{id}$ ($g^{-1}mg\,min^{-0.5}$) | $C$ | 0.98 |
| | 2.35 | 122.39 | |

$$\frac{1}{q_e} = \frac{1}{q_m K_L C_e} + \frac{1}{q_m} \quad (8)$$

where $q_e$ ($mg\,g^{-1}$) and $C_e$ ($mg\,L^{-1}$) are the adsorption capacity and concentration of adsorbate at equilibrium, respectively. The maximum adsorption capacity is denoted by $q_m$ ($mg\,g^{-1}$), and $K_L$ ($L\,mg^{-1}$) is termed as the Langmuir constant.

$R_L$ is a dimensionless separation parameter that explains the main characteristics of the Langmuir isotherm that is favorability of adsorption and is given by formula: $R_L = 1/(1 + K_L C_0)$. When the $R_L$ value lies in the range of 0–1 the process is favorable and is unfavorable, linear, and irreversible when the value of $R_L$ is more than 1, equal to 1, and equal to 0, respectively.

In Freundlich isotherm, the adsorption nonideal and irreversible and deals with multilayer adsorption on heterogeneous surfaces having distinct adsorption energies and affinities.

$$lnq_e = lnK_F + \frac{1}{n} lnC_e \quad (9)$$

The Freundlich constants $K_F$ and $1/n$ refer to the adsorbent's adsorption capacity and the favorability of isotherm, respectively, as follows:

$$0 < 1/n < 1 \text{ favorable}$$

$$1/n = 0 \text{ irreversible}$$

$$1/n > 1 \text{ unfavorable.}$$

In the Temkin isotherm model, all molecules in a layer are subjected to a linear decrease in heat of adsorption with coverage as a result of interactions between adsorbent and adsorbate molecules, and their binding energies are uniformly distributed, up to a certain maximum.

$$q_e = BlnK_T + BlnC_e \quad (10)$$



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7981 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

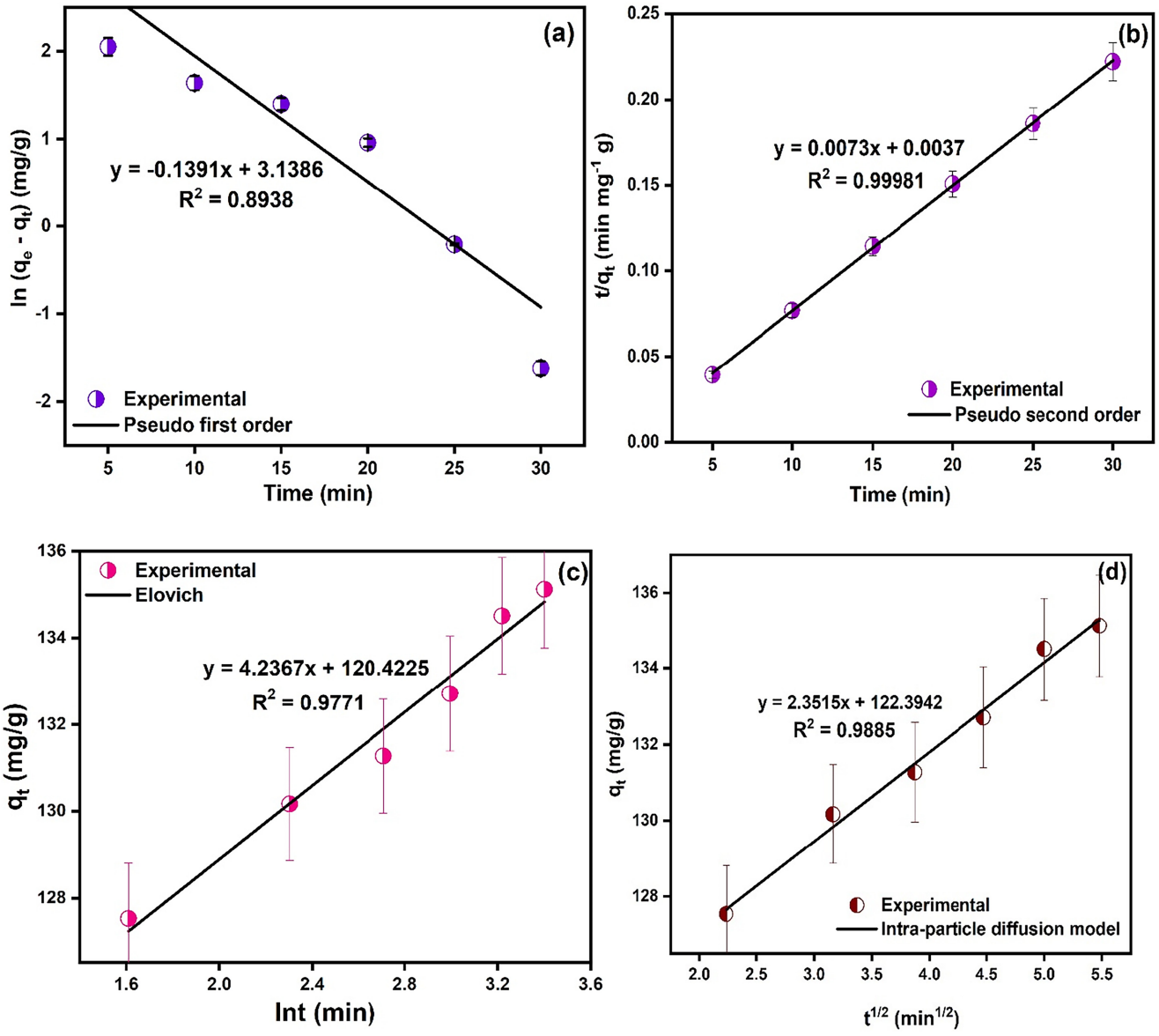



**FIGURE 7** | Plots of (a) PFO, (b) PSO, (c) Elovich, and (d) intraparticle diffusion models.

Here $K_T$ and $B$ are Temkin constants termed as equilibrium binding constant ($L\,mg^{-1}$) and heat of adsorption ($J\,mol^{-1}$), respectively.

The plots for MG adsorption isotherm are given in Figure 8 and Table 3 summarizes the values of constants and correlation coefficients ($R^2$) obtained from different models. The Langmuir equation fitted the experimental data well, thus suggesting monolayer adsorption. It also appears that MG adsorption onto Zn/Cu-LDO is a favorable process as the calculated value of $R_L$ lies between 0 and 1, that is, 0.1591.

### 3.2.8 | Regeneration Study

Results from regeneration experiments were obtained under optimum conditions and are presented in Figure 9a. A solution of 0.1-M NaOH was used to desorb dye from adsorbent surfaces. The removal efficiency was found to decrease from 90% to 73% after four regeneration cycles. Due to strong interactions between adsorbate and adsorbent, after each cycle removal, efficiency decreases, resulting in fewer active sites for adsorption.

### 3.2.9 | Assessment of Ultrasound-Assisted Method for MG Dye Uptake

The adsorption efficiency of Zn/Cu-LDO using ultrasound-assisted adsorption was compared with traditional adsorption methods such as magnetic stirring and incubator shaking. Figure 9b shows the percentage removal of MG with different adsorption methods. The % removal of MG in 30 min of reaction time followed the order: ultrasonication (90%) > magnetic stirring (74%) > incubator shaker (64%). The cavitation effect induced by ultrasonication plays a crucial role in enhancing sequestration of toxic effluents. The mechanical forces and high-energy shock waves produced during cavitation increase



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7981 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

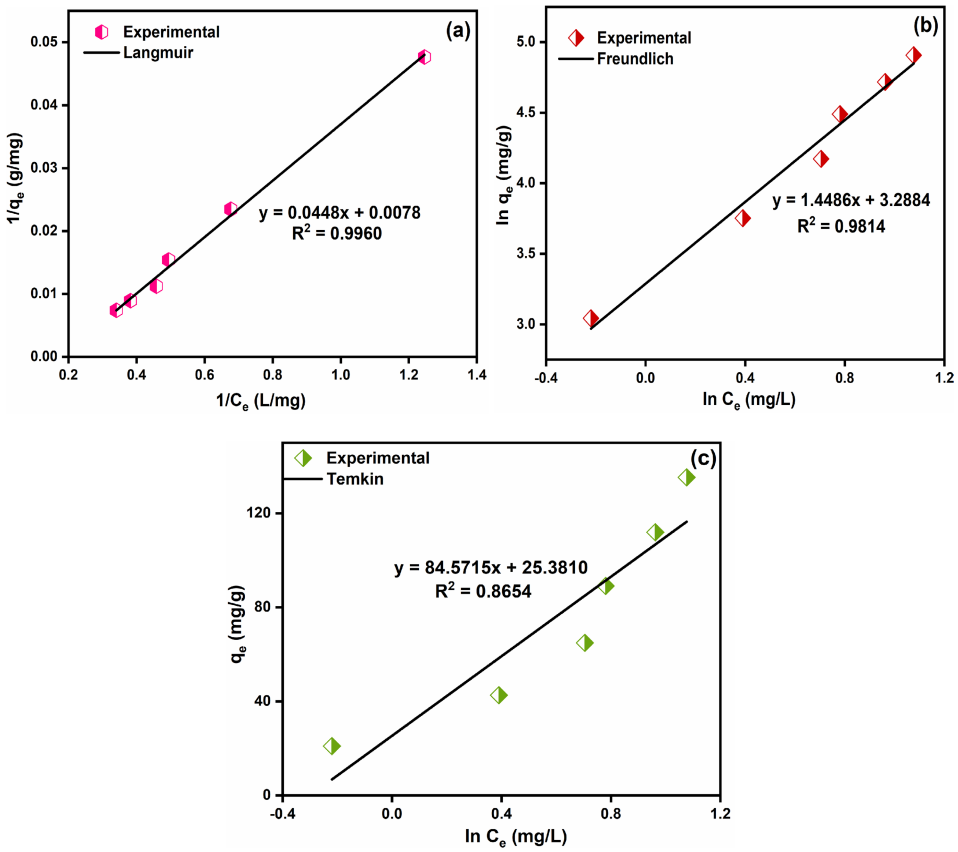



**FIGURE 8** | Plots of (a) Langmuir, (b) Freundlich, and (c) Temkin isotherms.

**TABLE 3** | Equilibrium parameters for MG dye removal.

| Langmuir | | | Freundlich | | | Temkin | | |
|---|---|---|---|---|---|---|---|---|
| $K_L$ ($min^{-1}$) 0.18 | $q_m$ (mg/g) 128.20 | $R^2$ 0.9960 | $K_F$ ($g\,mg^{-1}\,min^{-1}$) 26.80 | *1/n* 1.45 | $R^2$ 0.9814 | $K_T$ ($L\,mg^{-1}$) 1.35 | *B 84.57* | $R^2$ 0.8654 |

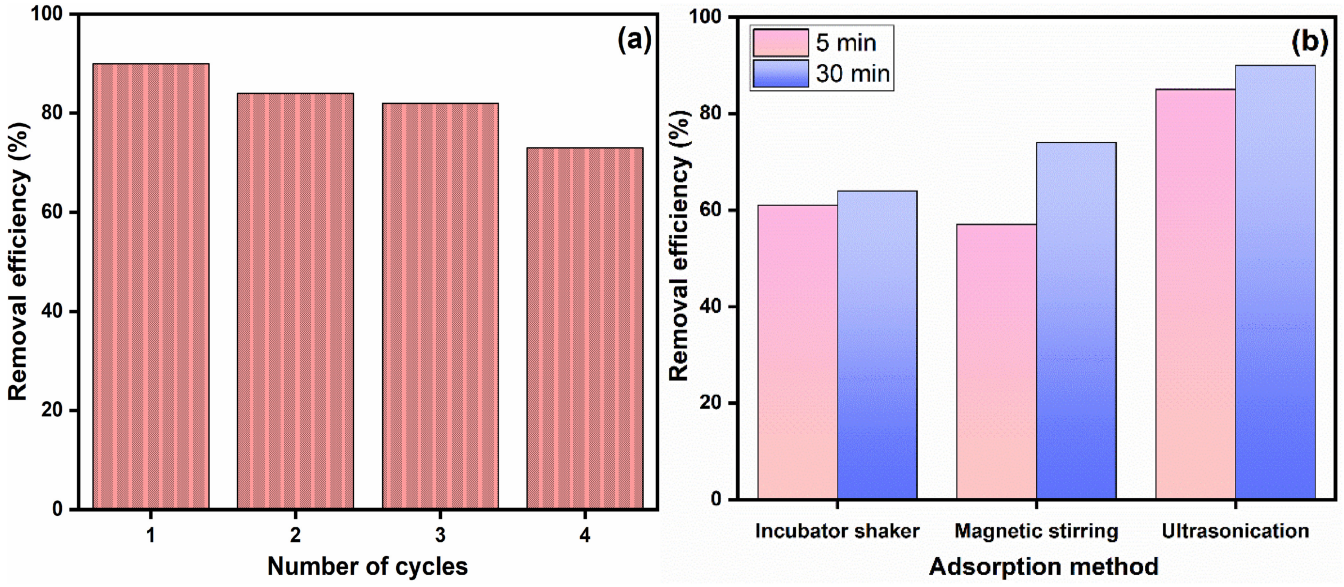



**FIGURE 9** | (a) Dye uptake efficiency of Zn/Cu-LDO up to four cycles and (b) assessment of adsorption methods for MG dye uptake.



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7981 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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** | Different sorbents reported in literature for MG dye uptake.

| Sorbent | Removal efficiency (%) | Adsorption capacity (mg/g) | Time (min) | pH | References |
|---|---|---|---|---|---|
| $Fe_3O_4$@PANI | 98 | 240 | 240 | 7 | [31] |
| Acid treated coffee husk | 96 | 264.81 | 120 | 6.8 | [32] |
| GumT-cl-HEMA-$TiO_2$ hydrogel | 99.3 | 103.09 | 110 | 7 | [26] |
| Zeolitic imidazole framework-8 (ZIF-8) | 95 | 224.14 | 90 | 7 | [15] |
| Cu-MOFs/$Fe_3O_4$ | 90 | 113.67 | 70 | — | [33] |
| Coal-assisted soil | > 95 | 89.97 | 60 | 6 | [3] |
| Zn/Cu-LDO | > 90 | 128.20 | 30 | 7 | Present work |

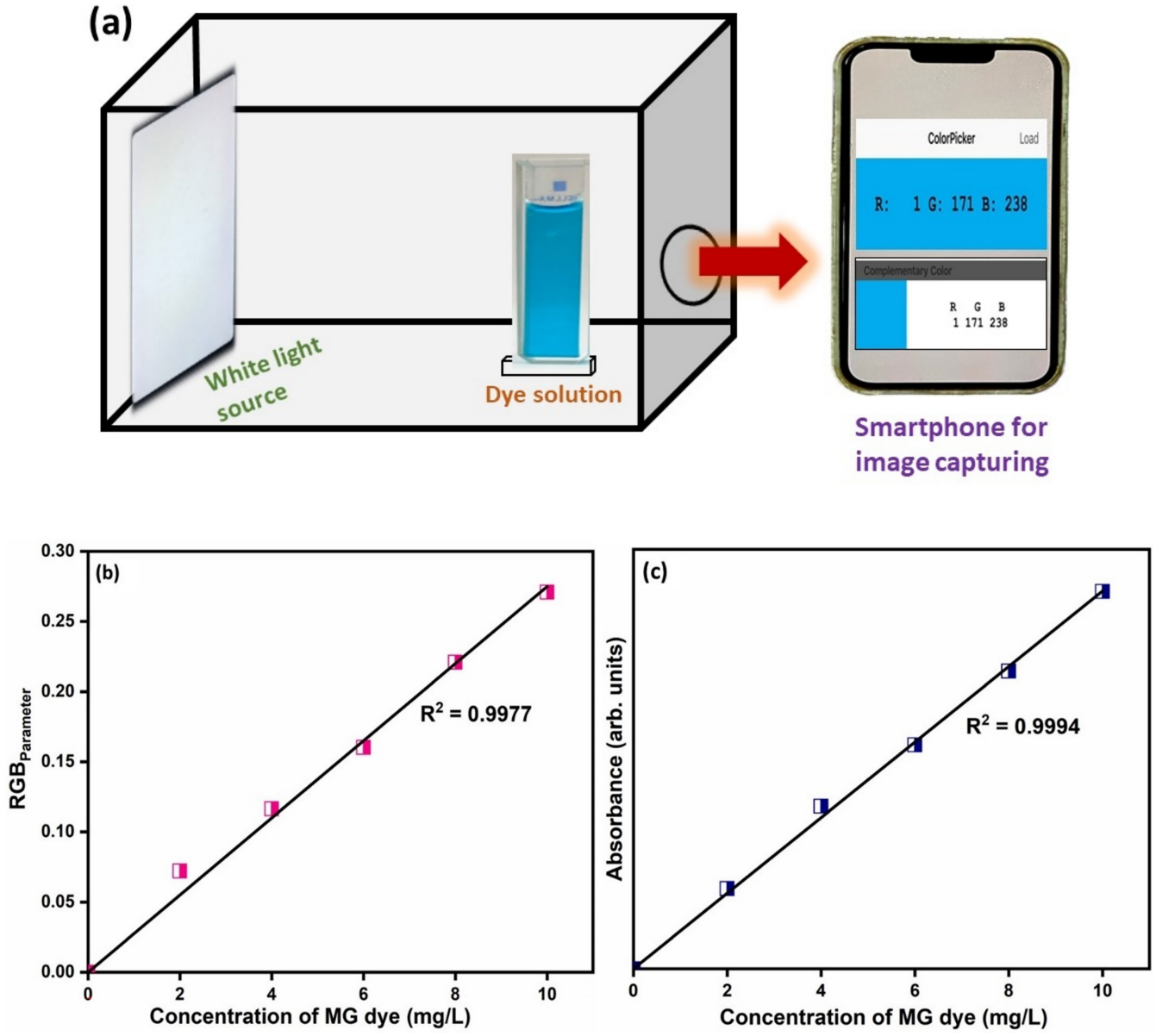



**FIGURE 10** | (a) Setup for smartphone-based determination of RGB parameters, calibration curves for MG dye concentration versus (b) $RGB_{parameter}$, and (c) UV–vis spectrophotometry.



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7981 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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 surface roughness of the adsorbent and create additional active sites (porous structure). This results in an improved adsorption capacity as more dye molecules can be captured on the activated surface. It prevents the aggregation of adsorbent particles by breaking up clusters, thereby maintaining a high surface area available for adsorption. This is especially important in maintaining the effectiveness of the adsorbent throughout the process [5]. Consequently, ultrasonication and sorption have synergistically reduced reaction time while enhancing dye removal efficiency as a result. Ultrasonic treatment of an adsorbent–adsorbate system enhances the surface porosity, thereby causing dye molecules to be distributed evenly inside the porous holes of the sorbent, which accelerates the mass transfer and sorption process and reduces the equilibrium time considerably.

The adsorption capability of synthesized lattice was also compared with previously reported sorbents (Table 4). According to the proposed method, MG showed superior removal performance compared with other reported sorbents in terms of adsorption capacity and time, as well as superior performance compared with all other adsorbents tested. As a consequence, the synthesized adsorbent can be utilized as an efficient material for the uptake of MG dye.

### 3.3 | Smartphone-Assisted Determination of Percentage dye Uptake

An approach based on smartphone colorimetry was utilized for the analysis of dye removal by using a simple and portable setup. The use of a smartphone to estimate dye adsorption rates using RGB (red, green, blue) color intensity is a rapid and efficient way to assess dye removal on-site [18]. Smartphones have an inbuilt camera and are capable of processing images, they eliminate the need for specialized instruments, and because they are readily available and simple to use, they do not require professional training. The setup uses two smartphones where one smartphone's screen is used to detect RGB values, whereas the other is used to generate white light, mounted on a fixed surface separated by a fixed distance (Figure 10a). In the current study, MG dye color intensity was analyzed using ColorPicker app (Smartphone: iPhone 13 possessing 12-MP dual camera). Due to the nonlinear relationship between RGB values and solution concentrations, a logarithmic scale was chosen and a calibration curve has been constructed for MG dye concentrations using $RGB_{parameter}$, which can be computed using the following equations:

$$RGB_{reference/sample} = \frac{(R+G+B)}{3} \tag{11}$$

$$RGB_{parameter} = Log\frac{(RGB_{reference})}{(RGB_{sample})} \tag{12}$$

Table 5 clearly shows that RGB increases with increasing MG concentration, corresponding to the solution's ability to absorb light. The RGB parameters for known concentration of MG dye are given in Table 5, and the calibration curves obtained from smartphone app and UV–vis spectrophotometer are displayed in Figure 10b,c, respectively.

**TABLE 5** | Color parameters of known concentration of malachite green dye.

| Concentration (mg/L) | R | G | B | $RGB_{parameter}$ |
|---|---|---|---|---|
| 0 | 255 | 255 | 255 | 0.0 |
| 2 | 184 | 227 | 237 | 0.07 |
| 4 | 130 | 216 | 239 | 0.12 |
| 6 | 85 | 205 | 239 | 0.16 |
| 8 | 37 | 187 | 236 | 0.22 |
| 10 | 1 | 171 | 238 | 0.27 |

It is possible to estimate the unknown dye concentration through calibration of the $RGB_{parameter}$ with the dye concentration. It has been demonstrated that the smartphone-assisted colorimetric determination for dye uptake is validated by the good agreement between the $\%R$ and regression coefficient values obtained from smartphone colorimetry ($\%R = 87\%$, $R^2 = 0.9977$) and UV–vis spectrophotometry ($\%R = 90\%$, $R^2 = 0.99$).

## 4 | Conclusions

Herein, Zn/Cu-LDO was synthesized by performing calcination of Zn/Cu LDH and was then used as adsorbent for ultrasonic-assisted adsorption of MG dye from aqueous solutions. The morphology structure and physicochemical and surface characteristics of Zn/Cu-LDO was studied using series of techniques, that is, PXRD, FTIR, SEM, EDX mapping, XPS, and BET. The results demonstrated that synthesized adsorbent exhibited high removal efficiency towards cationic dye. In the adsorption process, principally the effects of ultrasonication time, pH, adsorbent dosage, and adsorbate concentration on the removal efficiency of cationic MG dye were studied. Zn/Cu-LDO demonstrated a good fit to Langmuir isotherm and pseudo-second-order kinetic models for MG dye and exhibited a maximum monolayer adsorption capacity of 128.20 mg/g. Additionally, smartphone-assisted colorimetry based on RGB parameters proved to be a rapid, cost-effective, precise, and quantitative method for estimation of % removal of dye (87%) as compared with the conventional method of detection, that is, UV–vis spectroscopy (90%). Thus, owing to their ease of synthesis, porous structure, and high adsorption efficiency, these LDOs can be used as efficient adsorbent in wastewater remediation.

#### Author Contributions

**Bhamini Pandey:** conceptualization, validation, methodology, data curation, formal analysis, software. **Poonam Singh:** conceptualization, funding acquisition, validation, methodology, formal analysis, project administration, visualization, resources, supervision, writing–review and editing. **Ravinder Kumar:** writing–review and editing, project administration, supervision, visualization.



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7981 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

## Acknowledgments

The authors are grateful to Delhi Technological University for supporting their research. The authors are grateful to UGC-BSR for funding this project (30-416/2018[BSR]). In addition, BP would like to thank CSIR, India, for the financial support it provided.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## References

1. D. H. Kumar Reddy and S. M. Lee, "Water Pollution and Treatment Technologies," *Journal of Environmental & Analytical Toxicology* 02 (2012): 2–4.

2. B. S. Rathi, P. S. Kumar, and D. N. Vo, "Critical Review on Hazardous Pollutants in Water Environment: Occurrence, Monitoring, Fate, Removal Technologies and Risk Assessment," *Science of the Total Environment* 797 (2021): 149134.

3. T. R. Sundararaman, A. Saravanan, P. S. Kumar, et al., "Adsorptive Removal of Malachite Green Dye Onto Coal-Associated Soil and Conditions Optimization," *Adsorption Science and Technology* 2021 (2021): 1–11.

4. B. Pandey, P. Singh, and V. Kumar, "Photocatalytic-Sorption Processes for the Removal of Pollutants From Wastewater Using Polymer Metal Oxide Nanocomposites and Associated Environmental Risks," *Environmental Nanotechnology, Monitoring & Management* 16 (2021): 100596.

5. S. Das, S. Rudra Paul, and A. Debnath, "Fabrication of Biochar From Jarul (*Lagerstroemia speciosa*) Seed Hull for Ultrasound Aided Sequestration of Ofloxacin From Water: Phytotoxic Assessments and Cost Analysis," *Journal of Molecular Liquids* 387 (2023): 122610, https://doi.org/10.1016/j.molliq.2023.122610.

6. A. Deb, S. Das, and A. Debnath, "Fabrication and Characterization of Organometallic Nanocomposite for Efficient Abatement of Dye Laden Wastewater: CCD Optimization, Adsorption Mechanism, Co-Existing Ions, and Cost Analysis," *Chemical Physics Letters* 830 (2023): 22–25.

7. R. Abraham, S. Mathew, S. Kurian, M. P. Saravanakumar, A. Mary Ealias, and G. George, "Facile Synthesis, Growth Process, Characterisation of a Nanourchin-Structured α-$MnO_2$ and Their Application on Ultrasonic-Assisted Adsorptive Removal of Cationic Dyes: A Half-Life and Half-Capacity Concentration Approach," *Ultrasonics Sonochemistry* 49 (2018): 175–189.

8. K. Y. Foo and B. H. Hameed, "An Overview of Dye Removal via Activated Carbon Adsorption Process," *Desalination and Water Treatment* 19 (2010): 255–274.

9. J. Pathak and P. Singh, "Adsorptive Removal of Congo Red Using Organically Modified Zinc–Copper–Nickel Ternary Metal Hydroxide: Kinetics, Isotherms and Adsorption Studies," *Journal of Polymers and the Environment* 31, no. 1 (2022): 1–18, https://doi.org/10.1007/s10924-022-02612-0.

10. B. Pandey and P. Singh, "Statistical Optimization of Process Parameters for Ultrafast Uptake of Anionic Azo Dyes by Efficient Sorbent: Zn/Cu Layered Double Hydroxide," *Applied Organometallic Chemistry* 37 (2023): 1–18.

11. J. Pathak and P. Singh, "Synthesis and Characterization of Ternary Layered Double Hydroxide Containing Zinc/Copper/Nickel and Its PANI Composite," *Polymer Composites* 43, no. 11 (2022): 7836–7844.

12. R. M. Claydon, L. A. Roman-Ramirez, and J. Wood, "Comparative Study on the Hydrogenation of Naphthalene Over Both $Al_2O_3$-Supported Pd and NiMo Catalysts Against a Novel LDH-Derived Ni-MMO-Supported Mo Catalyst," *ACS Omega* 6 (2021): 20053–20067.

13. M. Mubarak, M. S. Islam, D. Y. Yoon, et al., "Flower-Like Mg/Fe-Layered Double Oxide Nanospheres With Ultrahigh Adsorption Efficiency for Anionic Organic Dyes," *Colloids and Surfaces A: Physicochemical and Engineering Aspects* 618 (2021): 1–8.

14. A. Mary Ealias and M. P. Saravanakumar, "A Critical Review on Ultrasonic-Assisted Dye Adsorption: Mass Transfer, Half-Life and Half-Capacity Concentration Approach With Future Industrial Perspectives," *Critical Reviews in Environmental Science and Technology* 49 (2019): 1959–2015.

15. N. Khoshnamvand, A. Jafari, B. Kamarehie, A. Mohammadi, and M. Faraji, "Removal of Malachite Green Dye From Aqueous Solutions Using Zeolitic Imidazole Framework-8," *Environmental Processes* 6 (2019): 757–772.

16. O. Acisli, A. Khataee, S. Karaca, and M. Sheydaei, "Modification of Nanosized Natural Montmorillonite for Ultrasound-Enhanced Adsorption of Acid Red 17," *Ultrasonics Sonochemistry* 31 (2016): 116–121.

17. A. Deb, A. Debnath, and B. Saha, "Ultrasound-Aided Rapid and Enhanced Adsorption of Anionic Dyes From Binary Dye Matrix Onto Novel Hematite/Polyaniline Nanocomposite: Response Surface Methodology Optimization," *Applied Organometallic Chemistry* 34 (2020): 1–20.

18. R. Koohkan, M. Kaykhaii, M. Sasani, and B. Paull, "Fabrication of a Smartphone-Based Spectrophotometer and Its Application in Monitoring Concentrations of Organic Dyes," *ACS Omega* 5 (2020): 31450–31455.

19. H. Bahrami, M. Mejmarian, H. Jafari, and F. Rasekh, "Synthesis of New Fe/Cr LDH for Removal of $Hg^{2+}$ Ion From Aqueous Solutions and Adsorption Study Using Smartphone-Based Colorimetry," *Environmental Nanotechnology, Monitoring & Management* 18 (2022): 100702.

20. B. Pandey, P. Singh, and V. Kumar, "Facile Synthesis of Zn/Cu and PANI Modified Layered Double Hydroxides With Insights Into Structural Memory Effect Property," *Bulletin of Materials Science* 45 (2022): 169.

21. A. L. Patterson, "The Scherrer Formula for X-Ray Particle Size Determination," *Physics Review* 56 (1939): 978–982.

22. H. Zhang, H. Chen, S. Azat, et al., "Super Adsorption Capability of Rhombic Dodecahedral Ca-Al Layered Double Oxides for Congo Red Removal," *Journal of Alloys and Compounds* 768 (2018): 572–581.

23. Y. Song, J. Bai, S. Jiang, et al., "Co-Fe-Mo Mixed Metal Oxides Derived From Layered Double Hydroxides for Deep Aerobic Oxidative Desulfurization," *Fuel* 306 (2021): 121751.

24. S. Intachai, P. Tongchoo, P. Sumanatrakul, P. Pankam, and N. Khaorapapong, "Efficient and Practical Adsorption of Mixed Anionic Dyes in Aqueous Solution by Magnetic NiFe-Layered Double Oxide," *Korean Journal of Chemical Engineering* 39 (2022): 2675–2684.

25. A. M. Ealias and M. P. Saravanakumar, "Application of Protein-Functionalised Aluminium Nanosheets Synthesised From Sewage Sludge for Dye Removal in a Fixed-Bed Column: Investigation on Design Parameters and Kinetic Models," *Environmental Science and Pollution Research* 27 (2020): 2955–2976.

26. B. Sharma, S. Thakur, G. Mamba, et al., "Titania Modified Gum Tragacanth Based Hydrogel Nanocomposite for Water Remediation," *Journal of Environmental Chemical Engineering* 9 (2021): 104608.

27. B. Pandey, P. Singh, and V. Kumar, "Adsorption of Anionic Dye From Aqueous Environment Using Surface-Engineered Zn/Cu Hydroxy Double Salt-Based Material: Mechanistic, Equilibrium and Kinetic Studies," *Journal of Environmental Science and Health.—Part A Toxic/Hazardous Substances and Environmental Engineering* 58 (2023): 869–884.

28. W. Jiang, L. Zhang, X. Guo, et al., "Adsorption of Cationic Dye From Water Using an Iron Oxide/Activated Carbon Magnetic Composites



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7981 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

Prepared From Sugarcane Bagasse by Microwave Method," *Environmental Technology (United Kingdom)* 42 (2021): 337–350.

29. X. Han, X. Niu, and X. Ma, "Adsorption Characteristics of Methylene Blue on Poplar Leaf in Batch Mode: Equilibrium, Kinetics and Thermodynamics," *Korean Journal of Chemical Engineering* 29 (2012): 494–502.

30. I. Langmuir, "The Constitution and Fundamental Properties of Solids and Liquids. Part II—Liquids," *Journal of the Franklin Institute* 184 (1917): 721.

31. T. K. Mahto, A. R. Chowdhuri, and S. K. Sahu, "Polyaniline-Functionalized Magnetic Nanoparticles for the Removal of Toxic Dye From Wastewater," *Journal of Applied Polymer Science* 131 (2014): 1–9.

32. T. P. Krishna Murthy, B. S. Gowrishankar, M. N. Chandra Prabha, M. Kruthi, and R. Hari Krishna, "Studies on Batch Adsorptive Removal of Malachite Green From Synthetic Wastewater Using Acid Treated Coffee Husk: Equilibrium, Kinetics and Thermodynamic Studies," *Microchemical Journal* 146 (2019): 192–201.

33. Z. Shi, C. Xu, H. Guan, et al., "Magnetic Metal Organic Frameworks (MOFs) Composite for Removal of Lead and Malachite Green in Wastewater," *Colloids and Surfaces A: Physicochemical and Engineering Aspects* 539 (2018): 382–390.



10990739, 2025, 2, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/aoc.7981 by Delhi Technological University, Wiley Online Library on [26/01/2025]. 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

# Unified dimensionality reduction techniques in chronic liver disease detection

Anand Karna[1], Naina Khan[1], Rahul Rauniyar[1], Prashant Giridhar Shambharkar[2]

[1]B. Tech Student, Computer Engineering Department, Delhi Technological University
[2]Assistant Professor, Computer Engineering Department, Delhi Technological University
*Corresponding Author: prashant.shambharkar@dtu.ac.in*

arXiv:2412.21156v1 [cs.LG] 30 Dec 2024

**Abstract**

*Globally, chronic liver disease continues to be a major health concern that requires precise predictive models for prompt detection and treatment. Using the Indian Liver Patient Dataset (ILPD) from the University of California at Irvine's UCI Machine Learning Repository, a number of machine learning algorithms are investigated in this study. The main focus of our research is this dataset, which includes the medical records of 583 patients, 416 of whom have been diagnosed with liver disease and 167 of whom have not. There are several aspects to this work, including feature extraction and dimensionality reduction methods like Linear Discriminant Analysis (LDA), Factor Analysis (FA), t-distributed Stochastic Neighbour Embedding (t-SNE), and Uniform Manifold Approximation and Projection (UMAP). The purpose of the study is to investigate how well these approaches work for converting high-dimensional datasets and improving prediction accuracy. To assess the prediction ability of the improved models, a number of classification methods were used, such as Multi-layer Perceptron, Random Forest, K-nearest neighbours, and Logistic Regression. Remarkably, the improved models performed admirably, with Random Forest having the highest accuracy of 98.31% in 10-fold cross-validation and 95.79% in train-test split evaluation. Findings offer important new perspectives on the choice and use of customized feature extraction and dimensionality reduction methods, which improve predictive models for patients with chronic liver disease.*

**Keywords:** Machine Learning, Preprocessing, Feature Extraction, LDA, t-SNE, Random Forest, Feature Integration, Liver Disease

## 1. Introduction

Over 1.7 million people globally pass away from a liver disease every year. The 11th most common cause of death and morbidity worldwide is cirrhosis (Cheemerla & Balakrishnan, 2021). Prominent diseases of the liver include liver cirrhosis, hepatitis A, B, and C, fatty liver disease, and liver cancer. Over 75 million persons worldwide suffer from alcohol use disorders associated with liver disease out of the 2 billion drinkers (approx.) worldwide. In addition, 400 million adults have diabetes, 2 billion adults are overweight, and these conditions are risk factors for cancer and non-alcoholic fatty liver disease. Drug-induced liver problems are on the rise, and viral hepatitis is still very common (Asrani et al., 2019). Liver transplants, the second most common procedure, meet less than 10% of global needs.

These are definitely fatal illnesses, but if caught early enough, they can be cured. The signs of liver disease usually appear only in the later stages of the illness so, early detection is difficult (Umbare et al., 2023). This is where machine learning techniques come in handy. They are widely employed in the healthcare industry, especially when it comes to the diagnosis and classification of certain diseases based on information about their features (Javaid et al., 2022). The increasing concern over chronic liver disease is what motivated this research. We employ cutting-edge methods to enhance prediction models on the Indian Liver Patient Dataset. In order to make timely treatment decisions for patients with chronic liver diseases, we are focusing on understanding how these techniques can improve diagnostic accuracy.

**Contribution**

- The study introduces a novel and sophisticated methodology, incorporating IQR-based outlier replacement, oversampling, and a unique combination of dimensionality reduction techniques (LDA, FA, t-SNE, UMAP) followed by scaling.
- The study demonstrates outstanding results, particularly with the Random Forest algorithm, showcasing accuracy of 98.31% and improved diagnostic precision.
- The study emphasizes the practical implications for clinical decision-making, providing nuanced insights for physicians to effectively apply sophisticated techniques in chronic liver disease diagnosis
- This research also contributes a valuable resource for practitioners and researchers, offering insights to enhance model performance in comparable healthcare settings.

The research paper's second section offers a thorough analysis of the body of existing literature as well as a thorough summary of all the pertinent studies.

The investigation's methodology is explained in Section 3, where important details like dataset information, preprocessing methods, and feature extraction method integration are covered in detail. Section 4 presents the research outcomes and findings, providing an understanding of the conclusions drawn from the study process. The study's major findings and conclusions, which have been discussed throughout the paper, are summed up in the final section.

## 2. Literature Survey

A significant burden of chronic liver disease (CLD) exists worldwide; in 2017, there were 5.2 million cases, which resulted in 1.48 million deaths in 2019. The prevalence of cirrhosis has been rising since 1990. In 2019, disability-adjusted life-years came in at number 16 out of all diseases and at number 7 among those aged 50-74 (Liu & Chen, 2022). A growing trend in alcohol consumption and cirrhosis linked to NAFLD contrasts (Younossi et al., 2016) with declining viral-associated burdens is seen. The use of predictive models to forecast the course of the disease, demonstrating the efficacy of ML in early identification. Additionally, effectiveness of machine learning in detecting problems has been shown, namely in differentiating between cirrhosis stages and forecasting the likelihood of hepatic decompensation (Popa et al., 2023). ML algorithms are essential for customizing treatment plans for individuals with CLD. Through the analysis of large datasets, specific patient features, and therapeutic strategy optimization based on expected outcomes, they offer individualized interventions. Patient results in the management of CLD have been greatly impacted by the use of machine learning, which has resulted in more focused and efficient treatments (Ahn et al., 2021).

When the high death rates associated with liver disorders were noticed, Bhupathi et al. (2022) investigated machine learning models such as Support Vector Machines (SVM), Naïve Bayes, K-Nearest Neighbors (KNN), Linear Discriminant Analysis (LDA), and Classification and Regression Trees (CART). Notably, KNN achieved 91.7% accuracy, while an autoencoder network demonstrated an even higher accuracy of 92.1%. Dhayanand et al. (2015) found that when SVM (79.66%) was applied to the Indian Liver Patient Records dataset, which contained 583 instances and 11 characteristics, it outperformed Naïve Bayes (61.28%). Wu et al. (2019) highlighted in their contribution the features that are present in datasets for fatty liver disease (FLD), such as patient demographics including age, gender, blood pressure, glucose levels, and liver enzyme levels. When data is collected using feature generation techniques in traditional machine learning systems, the input raw feature space is often highly dimensional and saturated with a large amount of useless feature information (Amin et al., 2023). The goal of linear discriminant analysis (LDA), a statistical technique for dimensionality reduction and classification, is to identify the linear feature combinations that most effectively divide data into distinct classes or groups (Batarseh & Freeman, 2022). A statistical method called factor analysis (FA) is used to determine the latent components that underlie the correlations between observed data to comprehend the underlying structure among them (Tavakol & Wetzel, 2020). A non-linear dimensionality reduction method called t-Distributed Stochastic Neighbour Embedding (t-SNE) is frequently used to visualize high-dimensional data in lower dimensions while maintaining local structures and emphasizing clusters or patterns in the data (Cai & Ma, 2021). Another nonlinear dimensionality reduction method, noted for its capacity to successfully capture both local and global features in high-dimensional data, is called Uniform Manifold Approximation and Projection (UMAP). It is frequently used for the analysis and visualization of complex datasets.

Singh et al. (2020) performed a thorough assessment of classifier performance using ILPD and compared results with and without feature selection strategies. This study focused specifically on how feature selection affects classifier performance. Interestingly, when feature selection was removed, Logistic Regression showed the highest accuracy (72.50%), but Random Forest outperformed with an accuracy of 74.36% after feature selection was added. In order to clarify how feature selection strategies aid in the improvement of ILPD models, the evaluation comprised a variety of metrics, such as the number of examples that were successfully classified, the Kappa statistic, mean absolute error, and execution time. The study aimed to guide the development of more effective and efficient diagnostic models for liver health assessment by comparing classifiers with and without feature selection, thereby offering insights into the subtle effects of this approach. Wang et al. (2015) examined pulse signal analysis in cirrhosis patients, fatty liver disease (FLD) patients, and healthy individuals. Using sophisticated methods like harmonic fitting and learning algorithms, such as Principal Component Analysis (PCA), Least Squares (LS), and Least Absolute Shrinkage and Selection Operator (LASSO), the research sought to identify distinguishing patterns that would emphasize differences between the FLD/cirrhosis and healthy groups. With seven factors identified, this computer-aided diagnostic method showed great promise in the field of Traditional Chinese Medicine (TCM), with an ac-

curacy rate of over 93%. This highlights its potential to greatly improve clinical diagnostics and demonstrates its applicability as a useful instrument in the evaluation of disorders related to the liver.

Kumar and Sahoo (2013) used a dataset with 583 individuals to present a Rule-Based Classification Model (RBCM) that combines rule-based and data mining methods for liver disease prediction. Several ML, such as Support Vector Machines (SVM), Rule Induction (RI), Decision Trees (DT), Naïve Bayes (NB), and Artificial Neural Networks (ANN), were applied during their investigation. The Decision Tree (DT) demonstrated superior performance compared to the other methods, with remarkable results of 98.46% accuracy, 95.7% sensitivity, 95.28% specificity, and a Kappa coefficient of 0.983. The study underscored the efficacy of the Rule-Based Classification Model, highlighting its superiority over non-rule-based models. The remarkable efficacy of DT highlighted the importance of applying rule-based methods when discussing the prognosis of liver disease. In addition to demonstrating the Rule-Based Classification Model's strong predictive abilities, this research established the model as a useful resource for medical decision-making regarding the prognosis of liver disease. Trigka et al. (2023) distinguished between "Liver-Disease" (LD) and "Non-Liver-Disease" (Non-LD) classes in their carefully constructed classification problem on liver disease prediction. A balanced dataset of 828 participants was produced by using the oversampling method (SMOTE) to address class imbalance. Afterwards, the significance of each feature was assessed using the three feature ranking techniques of Pearson Correlation, Gain Ratio, and Random Forest. Direct bilirubin (DB) was found to be the most significant feature after a Pearson correlation analysis showed significant relationships between a few features. In summary, the most crucial features—DB, TB, and SGOT—were used for the models' training and validation. With the use of multiple ML models and ensemble approaches, the developed predictive model showed impressive results in terms of accuracy (80.1%), precision (80.4%), recall (80.1%), and area under the curve (AUC) (88.4% post-SMOTE with 10-fold cross-validation).

Muthuselvan et al. (2018) investigated the use of machine learning algorithms like Naïve Bayes, K-Star, J48, and Random Tree for the classification of liver disease. The study concentrated on a dataset that included 11 liver health-related parameters and was obtained from the northeastern Indian state of Andhra Pradesh. In order to prepare the dataset for machine learning analyses, it underwent extensive exploration and pre-processing in WEKA. This included actions like handling missing values, data discretization, and converting numeric values to nominal ones. The unique methodology of the study consisted of analyzing various algorithms to identify the subtleties of their performances. The results showed that K-Star executed faster (less than 1 second) and had a reasonably competitive accuracy, while Random Tree had the highest accuracy (74.2%) but required a longer execution time (0.05 secs). The study's focus on a variety of algorithms and real-world applicability, along with WEKA's user-friendly environment, provided insightful information for researchers examining machine learning within the framework of liver disease prediction as well as educators.

Several classification algorithms, such as the Naive Bayesian Classifier (NBC), K-Nearest Neighbors (K-NN), and C4.5, were used in the study by Babu et al. (2016) to assess how well they predicted liver disease. These algorithms obtained classification accuracies of 0.56, 0.64, and 0.69 on the original ILPD dataset, respectively. But the researchers used a thorough approach that included feature selection and clustering, which resulted in the development of an updated dataset (NDS). The updated dataset showed significant gains in classification accuracy: 0.90, 0.95, and 0.93 for K-NN, C4.5, and NBC, respectively. The need to address issues with inconsistent data and low accuracy in clustering and classification techniques led to the selection of this specific methodology. The number of attributes or dimensions is a key factor that affects how well clustering algorithms perform, as the researchers discovered. They used feature selection techniques to lower dimension and then put the K-Means clustering algorithm into practice for unsupervised partitioning in order to lessen these difficulties. A refined training dataset was produced by validating the resulting clusters using cross-validation and expert intuition. The improved results on the NDS dataset highlight how well this methodology works to optimize liver disease prediction models.

In a recent study, Sujith et al. (2023) used the ILPD dataset in a Jupyter Notebook environment to classify liver diseases using a variety of machine learning techniques, such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Artificial Neural Networks (ANN), and Logistic Regression. Among the methods investigated, CNN achieved the highest accuracy (67%) and precision (71%) when optimized with Adamax, making it the most successful model. The choice to employ Adamax is in line with its established efficacy in neural network model optimization, specifically in managing sparse gradients and fostering stability throughout the training process. The researchers' selection of Adamax is consistent with its broad use in machine learning applications and demonstrates its suitability for

the particular difficulties associated with classifying liver disease in their study. In order to predict liver disease based on different attributes, Sivasangari et al. (2020) investigated the use of three different machine learning algorithms: SVM, DT and RF. The suggested approach called for gathering data from the UCI machine learning library, replacing missing values through preprocessing, and validating the model through a split train-test approach. The experimental results, which included confusion matrices and quantitative assessments of accuracy, precision, and recall, demonstrated the usefulness of the models. SVM achieved the highest accuracy (95.18%) according to the comparative analysis, while RF showed perfect precision. The study emphasized the significance of algorithmic selection and evaluation metrics, and it provided insightful information about liver disease prediction through ML.

# 3. Methodology

In the course of this investigation, the Indian Liver Disease Dataset (ILPD) was subjected to a thorough preprocessing regimen, encompassing imputation and oversampling, data simulation, concatenation, and outlier replacement. Using dimensionality reduction methods, key patterns supporting the classification of liver disease were extracted. A strong protocol for efficient data handling was formed by this methodological approach, which also included operations like standard scaling, potential data integration, and outlier management. A rigorous train-test split and cross-validation approach was used during the evaluation phase to ensure the reliability of predictive models and reduce the biases associated with the investigation. This methodical and condensed approach made it easier to create a thorough framework for efficient data processing and model building. Figure 1 visually represents the research methodology.

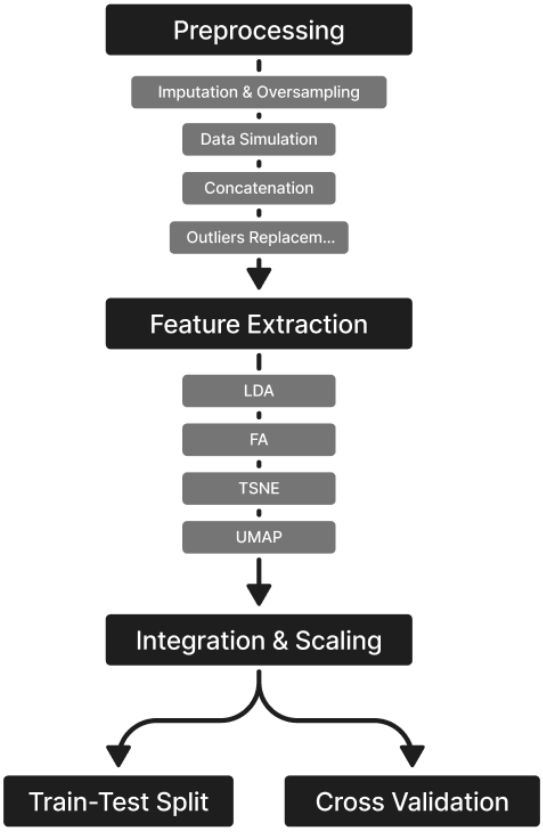



**Figure 1:** *A flow diagram depicting the methodology used in the study*

## 3.1. Dataset

A useful resource that is often obtained from the UCI Machine Learning Repository is the Indian Liver Patient Dataset (ILPD), which has patient medical attributes mostly pertaining to liver health. Many characteristics are usually included in this dataset, including age, gender, liver function tests, symptoms, alcohol intake, hepatitis history, blood test findings (bilirubin levels, liver enzymes), and symptoms. Its target variable shows whether a diagnosis of liver disease has been made (liver condition: 1) or not (liver condition: 2 or 0). The ILPD dataset, a moderately substantial collection of patient records, is used as a standard for machine learning research, facilitating the creation and assessment of models intended to predict the presence or absence of liver disease. We converted the 'Target' variable into a binary format, which typically indicates whether or not liver disease has been diagnosed.With this modification, the values 0 and 1 were mapped to represent the presence and absence of liver conditions, respectively.

## 3.2. Preprocessing

### 3.2.1 Imputation and Oversampling

We methodically preprocessed the data to reduce class imbalance within the dataset. This included removing unnecessary columns i.e., Gender as it is categorical, handling missing values, and uniformly converting data types. We used the mean of the 'ALB' column to impute the missing values in the 'A/G' column (McKnight et al., 2007). With RandomOverSampler, we were able to effectively balance the distribution of classes. After oversampling, the class distribution was 1334 (0) and 1330 (1). In order to reduce bias resulting from missing values and improve the accuracy of statistical measures, imputation was performed before outlier replacement. In order to mitigate biases in minority class representation, oversampling following imputation guaranteed that

**Table 1:** *Indian Liver Patient Dataset Feature Description*

| Variable Name | Type | Description | Missing Values |
|---|---|---|---|
| Age | Integer | Age | no |
| Gender | Categorical | Gender | no |
| TB | Real Number | Total Bilirubin | no |
| DB | Real Number | Direct Bilirubin | no |
| Alkphos | Integer | Alkaline Phosphotase | no |
| Sgpt | Integer | Alamine Aminotransferase | no |
| Sgot | Integer | Aspartate Aminotransferase | no |
| TP | Real Number | Total Proteins | no |
| ALB | Real Number | Albumin | no |
| A/G Ratio | Real Number | Albumin and Globulin Ratio | 4 |
| Target | Binary | Absence or presence of disease | no |

synthetic instances reflected a more comprehensive dataset.

### 3.2.2. Data Simulation

We generated a synthetic dataset with 1000 samples and 10 features, where 5 features provided relevant information, 3 were redundant, and 2 were repeated. This dataset was produced specifically for a binary classification assignment. The synthetic dataset was then concatenated with the original dataset.

Let, the dataset matrix is represented by $X$, where a feature is represented by a column and a sample by a row. This is represented mathematically as $X \in \mathbb{R}^{1000\times 10}$. $X_{\text{info}}$ denotes the informative features, which are intended to hold pertinent data, while $X_{\text{red}}$ stands for the redundant features. The formulation in can be written as follows:

$$X_{\text{info}} \in \mathbb{R}^{1000\times 5}, \quad X_{\text{red}} \in \mathbb{R}^{1000\times 3}$$

Additionally, $X_{\text{rep}}$ represents the repeated features, which are duplicated to create variation in the dataset. Mathematically, this is represented as:

$$X_{\text{rep}} \in \mathbb{R}^{1000\times 2}$$

This mathematical framework allowed for careful exploration of the behavior of machine learning models. Informative features aim to capture meaningful patterns, whereas redundant and repeated features pose challenges for models to process. Our understanding of how various types of information impact algorithms has improved as a result of this methodical approach, which has helped with model design and training methods. Furthermore, the learned lessons have useful ramifications for improving model applications in various real-world contexts.

### 3.2.3. Concatenation

This operation was beneficial as it vertically stacked the samples from both datasets, creating a larger dataset that incorporated information from both sources. By utilizing the combined knowledge from the simulated and original datasets, it made it possible to conduct a more thorough analysis and train the model.

$$X_{\text{original}} \in \mathbb{R}^{n_{\text{original}}\times m}$$

$$X_{\text{simulated}} \in \mathbb{R}^{n_{\text{simulated}}\times m}$$

$$X_{\text{combined}} = \begin{bmatrix} X_{\text{original}} \\ X_{\text{simulated}} \end{bmatrix} \tag{1}$$

$$X_{\text{combined}} \in \mathbb{R}^{(n_{\text{original}}+n_{\text{simulated}})\times m}$$

$$y_{\text{combined}} = \begin{bmatrix} y_{\text{original}} \\ y_{\text{simulated}} \end{bmatrix} \tag{2}$$

$$y_{\text{combined}} \in \mathbb{R}^{n_{\text{original}}+n_{\text{simulated}}}$$

Aligning the target labels with the combined dataset was essential for accurate analysis and model training. The subsequent steps in the data processing pipeline were made easier by the mathematical process that guaranteed consistency in the dataset structure.

### 3.2.4. Outliers Replacement

The Interquartile Range (IQR) is a robust statistical measure defined as the difference between the third quartile ($Q3$) and the first quartile ($Q1$):

$$IQR = Q3 - Q1 \tag{3}$$

The values known as quartiles split a dataset into four equal sections. The 25th percentile of the data is represented by the first quartile ($Q1$), and the 75th percentile by the third quartile ($Q3$). Because it is insensitive to extreme values or dataset outliers, the IQR is a reliable indicator of spread. By identifying the distribution of the central 50% of the data, it can withstand outliers. Lower and upper bounds are computed in order to identify outliers:

$$\text{Lower Bound: } = Q1 - k \times IQR \tag{4}$$

$$\text{Upper Bound: } = Q3 + k \times IQR \tag{5}$$

where $k$ (1.5 or 3) determines sensitivity.

To ensure reliability, outliers were then found and replaced with second quartile values using the IQR approach. At the same time, second quartile values were used to imputation for missing values. These methodical processing steps produced a refined dataset with IQR-based outlier detection, establishing the foundation for further analytical and modeling efforts.

## 3.3. Feature Extraction

Feature extraction stands out as a crucial step in analysis as it means taking better, more significant characteristics out of the raw data in order to reduce dimension. This method seeks to improve data management and increase the effectiveness of machine learning models by eliminating unnecessary or superfluous features while keeping relevant information (Alpaydin, 2014). There are several approaches in the field of feature extraction, which include both linear and non-linear dimensionality reduction methods. When integrated, they provide a holistic method for improving data representation. Combining linear approaches for dimensionality reduction first and then using non-linear approaches enables a cooperative combination that captures both global and local data structures (Muixí et al., 2023). The combination of Linear Discriminant Analysis and Factor Analysis with non-linear methods like t-distributed Stochastic Neighbor Embedding and Uniform Manifold Approximation and Projection, which is the driving factor of the investigation, is reflected.

### 3.3.1. Linear Discriminant Analysis

The goal of LDA is to maximize class separability while producing a lower-dimensional space (Tharwat et al., 2017). This is accomplished by determining the axes for classification jobs that retain the most amount of information. LDA prioritizes two important goals: increasing the between-class scatter, which guarantees that distinct class clusters are farther apart, and limiting the within-class scatter, which ensures that data points within the same class are closer together. The number of projected axes is determined by either the number of features or the number of unique classes reduced by one in order to prevent overfitting and capture the most discriminative information. LDA maximizes the performance of classification algorithms by preserving important class-specific information while identifying a transformation that best divides classes in the smaller feature space. This is done by computing the scatter matrices reflecting within-class and between-class variances.

The within-class scatter matrix is defined as:

$$S_w = \sum_{i=1}^{K} \sum_{\mathbf{x} \in X_i} (\mathbf{x} - \boldsymbol{\mu}_i)(\mathbf{x} - \boldsymbol{\mu}_i)^T \tag{6}$$

The between-class scatter matrix is defined as:

$$S_b = \sum_{i=1}^{K} N_i (\boldsymbol{\mu}_i - \boldsymbol{\mu})(\boldsymbol{\mu}_i - \boldsymbol{\mu})^T \tag{7}$$

Where:
- $N$ = Total number of samples
- $K$ = Number of classes
- $\mathbf{x}_i$ = Samples from the $i^{th}$ class
- $\boldsymbol{\mu}_i$ = Mean of samples from the $i^{th}$ class
- $S_w$ = Within-class scatter matrix
- $S_b$ = Between-class scatter matrix
- $S_t = S_w + S_b$ (Total scatter matrix)

By resolving the generalized eigenvalue problem, LDA seeks to determine the projection vector $\mathbf{w}$ that maximizes the ratio of between-class scatter to within-class scatter:

$$S_w^{-1} S_b \cdot \mathbf{w} = \lambda \mathbf{w} \tag{8}$$

Where $\mathbf{w}$ is the projection vector and $\lambda$ is the eigenvalue associated with that vector. A lower-dimensional subspace that offers the greatest separation between distinct classes is created from the data using this projection vector, $\mathbf{w}$.

### 3.3.2. Factor Analysis

By identifying latent variables—referred to as factors—that explain correlations between observed variables, FA is a statistical technique used to reduce the dimension of a dataset (Smelser & Baltes, 2001). Its operation is predicated on the idea that fewer unobserved factors have an impact on observed variables. FA seeks to identify a lower-dimensional representation that captures most of the variability in the data and to summarize the underlying structure by estimating the associations between the observed variables using the covariance matrix.The number of factors to extract is pre-specified, in our case 3, in order to reveal the key elements of the dataset. Factor loadings, which show the degree of correlation between latent factors and observed variables, are calculated by FA. FA also estimates an error term, which is the unexplained variance or noise in the dataset that cannot be explained by the factors that have been found. This error term offers insights into the variability not captured by the extracted factors and aids in evaluating the model fit and understanding the residual variance in the dataset.

Using the covariance matrix, FA calculates the relationships between the observed variables in an effort

to find a lower-dimensional representation that summarizes the underlying structure and captures the majority of the variability in the data (Johnson & Wichern, 2007). The observed variables' covariance matrix ($S$) can be decomposed into two components representing common variance due to factors and unique variance:

$$S = \mathbf{L}\mathbf{L}^T + \mathbf{\Psi} \tag{9}$$

Where:

- $S$ is the observed variables' covariance matrix.
- $\mathbf{L}$ represents factor loadings.
- $\mathbf{\Psi}$ signifies the diagonal matrix representing unique variances.

FA also estimates an error term, which is the unexplained variance or noise in the dataset that cannot be explained by the factors that have been found. This error term helps to clarify the residual variance in the dataset and evaluate the model fit. The estimation of the error term ($\mathbf{\Psi}$) is derived from the diagonal elements of the residual covariance matrix:

$$\mathbf{\Psi} = \text{diag}(\mathbf{\Sigma} - \mathbf{L}\mathbf{L}^T) \tag{10}$$

Where:

- $\mathbf{\Sigma}$ is the sample covariance matrix of the observed variables.
- $\mathbf{L}\mathbf{L}^T$ represents the estimated common variance due to factors.

### 3.3.3. t-Distributed Stochastic Neighbour Embedding

We applied t-SNE, a dimensionality reduction technique, to combined features that were obtained from Linear Discriminant Analysis (LDA) and Factor Analysis (FA). To create a combined feature set, the reduced features that were extracted independently from LDA (X_lda) and FA (X_fa) were first concatenated and then, t-SNE was then used. It was instantiated with three output dimensions and a fixed random state. t-SNE models high-dimensional data into a lower-dimensional space while maintaining the data points' local structure (Zhou et al., 2018). In order to facilitate visualization and investigation of underlying patterns or clusters within the data, it attempts to represent each high-dimensional data point in a lower-dimensional space, with dissimilar data points placed farther away and comparable data points closer together. As a result, the reduced features from the concatenated FA and LDA features obtained via t-SNE weres stored in the resulting X_tsne. The joint probability $p_{ij}$ that a data point $x_i$ would pick $x_j$ as its neighbor in the high-dimensional space is calculated using a Gaussian distribution with the similarities $\text{sim}_i(x_j)$ (van der Maaten et al., 2008):

$$p_{ij} = \frac{\exp(-\|x_i - x_j\|^2/2\sigma_i^2)}{\sum_{k\neq i}\exp(-\|x_i - x_k\|^2/2\sigma_i^2)} \tag{11}$$

where $\sigma_i$ is the Gaussian distribution's variance computed using the high-dimensional space's data point distances.

A Student's t-distribution with degrees of freedom df and the Euclidean distances $\|y_i - y_j\|^2$ between the low-dimensional representations are used to determine the joint probability $q_{ij}$ that $x_i$ would select $x_j$ as its neighbor in the low-dimensional space:

$$q_{ij} = \frac{(1 + \|y_i - y_j\|^2)^{-1}}{\sum_{k\neq i}(1 + \|y_i - y_k\|^2)^{-1}} \tag{12}$$

The goal of t-SNE optimization is to minimize the KL divergence between the distributions. $P = \{p_{ij}\}$ and $Q = \{q_{ij}\}$, which is given by:

$$C = \text{KL}(P||Q) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}} \tag{13}$$

The algorithm iteratively adjusts the low-dimensional representations $Y = \{y_i\}$ to minimize this cost function by employing gradient descent. By exposing underlying patterns or clusters, this procedure aids in maintaining the local structure of the data points in the low-dimensional space.

### 3.3.4. Uniform Manifold Approximation and Projection

UMAP effectively extracts local and global structures from high-dimensional data while providing better data relationship preservation in lower-dimensional spaces (McInnes et al., 2018). After LDA, FA, and t-SNE, we presented a unique computational step wherein we configured a three-component UMAP model, which differs from the parameterizations used in the previous methods. UMAP sought to maximize the preservation based on the ideas of manifold learning. In contrast to the linear separability emphasis of LDA, the latent variable extraction of FA, or the probabilistic modeling of t-SNE, which focuses on local structures, UMAP used a topological approach, giving complex manifold relationships within the data priority. This customized method produced an improved three-dimensional model that provided a fresh viewpoint different from previous methods. The objective function minimized by UMAP is given by:

$$\underset{Y}{\arg\min}\left(\sum_{i=1}^{N}\sum_{j=1}^{N} f(d_{ij}) \log\left(\frac{f(d_{ij})}{f(\hat{d}_{ij})}\right) + \lambda \sum_{i=1}^{N}\sum_{j=1}^{N}(1 - w_{ij})(\hat{d}_{ij} - d_{ij})^2\right) \tag{14}$$

Where:

- $Y$ represents the low-dimensional embedding that UMAP tries to optimize.

- $N$ denotes the number of data points or instances in the dataset.

- $f()$ is a function that operates on pairwise distances $d_{ij}$ between data points $i$ and $j$ in the original high-dimensional space.

- $d_{ij}$ and $\hat{d}_{ij}$ are the actual and approximated distances, respectively, between data points $i$ and $j$ in the high-dimensional space.

- $\lambda$ is a parameter that controls the trade-off between preserving local and global structure in the lower-dimensional embedding.

-$w_{ij}$ represents weights associated with the pairwise distances between data points

## 3.4. Feature Integration

### 3.4.1. Integration

We intentionally diverged from the proposal put forth by Vaidya and Vaidya (2022), which recommended replacing outliers following dimensionality reduction. This choice was carefully considered in light of the need to preserve data integrity and lessen the effect of outliers, particularly when employing t-SNE. t-SNE is intrinsically sensitive to outliers because it seeks to preserve local structures in the data. By giving outlier replacement precedence over dimensionality reduction, the intrinsic structure of the data is faithfully captured in the lower-dimensional space by taking care to shield data points from the perturbing effects of extreme values.

By reducing the divergence between probability distributions, the t-SNE algorithm highlights the similarities between nearby points and largely ignores the global structure. According to Halladin-Dabrowska et al. (2019), outliers can significantly impact the local structure and result in skewed representations in the lower-dimensional space by drawing or repelling neighboring points. The t-SNE cost function uses pairwise similarities, and the extreme values of outliers can disproportionately impact the optimization process and have a discernible impact on the visualization that is generated. The t-SNE cost function is defined by the Kullback-Leibler divergence between the conditional probabilities of pairs of data points in the high-dimensional space and their equal pairs in the low-dimensional space (van der Maaten et al., 2008). This can be expressed mathematically as equation (13). In contrast, LDA searches for a lower-dimensional space where the maximum separation between different classes exists. LDA is less prone to outliers since it is a linear method that focuses on the global structure and class separability. Outliers have less of an impact on the determination of the discriminant axes because LDA is based on the overall distribution of data points and their class labels. Class separability in the feature space is emphasized by the LDA's objective.

### 3.4.2. Scaling

The `StandardScaler` was used to normalize the dataset's features. A common technique used to make sure features are on a similar scale is standard scaling. In the underlying mathematics, each feature is transformed independently by dividing by its standard deviation and subtracting its mean. In mathematical terms, the transformation for a feature $X_i$ in the dataset is $X_i' = \frac{X_i - \text{mean}(X_i)}{\text{std}(X_i)}$ (15).

The scaling guarantees that all features are on an equalized scale, which is important for models that depend on distance metrics (e.g., support vector machines or k-nearest neighbors) or other features that are sensitive to feature scales. By taking the mean out of every feature, it accomplishes mean centering, which makes features easier to compare and understand. Additionally, `StandardScaler` maintains the data's distributional shape, which makes it a good option when the analysis depends on preserving the original distribution. Although there are other scaling techniques such as Robust and Min-Max scaling, `StandardScaler` is especially suitable in cases where the features have a distribution that is roughly normal.

## 3.5. Model Assessment

### 3.5.1. Train-Test Split

We used stratified splitting to maintain class proportions when separating the dataset into training (75%) and testing (25%) subsets. Reproducibility was guaranteed by the random state of 42. This division made it possible to train and assess models independently for an objective appraisal of performance.

### 3.5.2. Cross Validation

10-fold cross-validation was also used to assess the of the model. Using this method, the dataset was divided into 10 subsets, 9 of which were used for training and one for validation, alternately.

# 4. Results and Discussion

We used multiple dimensionality reduction techniques on our dataset and sought to compress the feature space: LDA concentrated on class separability, FA captured latent variables, t-SNE prioritized local structures, and UMAP preserved both local and

global structures in lower dimensions. The ensuing discourse contrasts and evaluates these methodologies, accentuating their individual merits and demerits in proficiently portraying intricate data structures. The algorithms used were LR, MLP, KNN and RF.

## 4.1. Visualization and Evaluation

### 4.1.1. ROC Curves

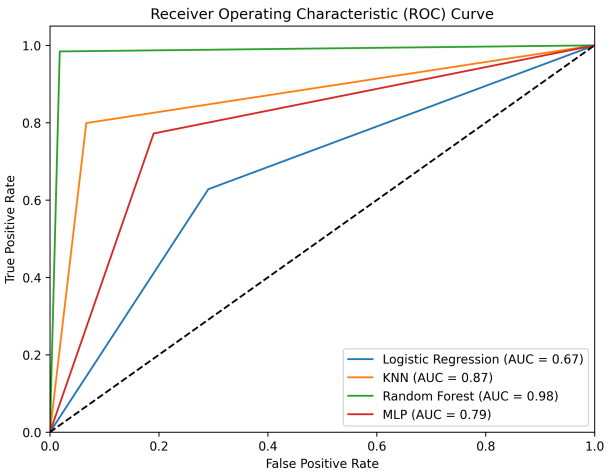



**Figure 2:** *ROC curves in 10-fold CV*

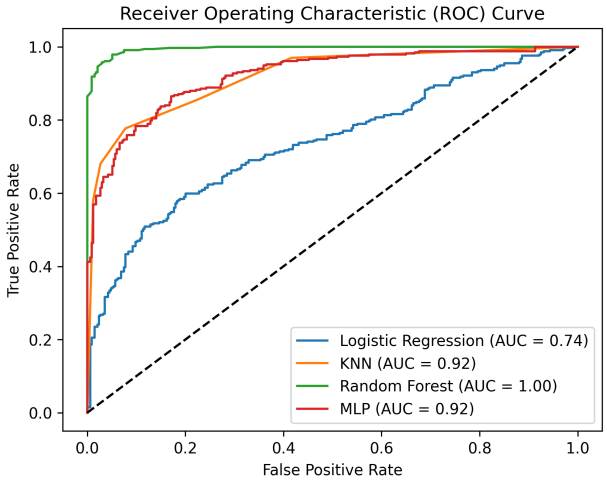



**Figure 3:** *ROC curves in Train-Test Split*

The ability of each model to differentiate between classes was clearly shown by figures 2 and 3. Above all False Positive Rate thresholds, Random Forest maintained higher True Positive Rates than the others. KNN and MLP showed moderate discriminatory abilities, while Logistic Regression exhibited relatively lower discriminatory power.

### 4.1.2. Precision-Recall Curves

The curve for precision and recall for each algorithm was figures 4 and 5. Across a range of recall thresholds, Random Forest and Multi-Layer Perceptron demonstrated superior precision levels and outperformed Logistic Regression and K-Nearest Neighbors.

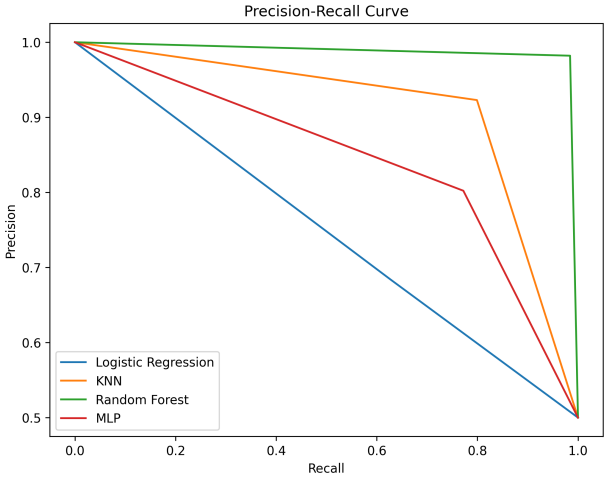



**Figure 4:** *Precision-Recall curves in 10-fold cross validation*

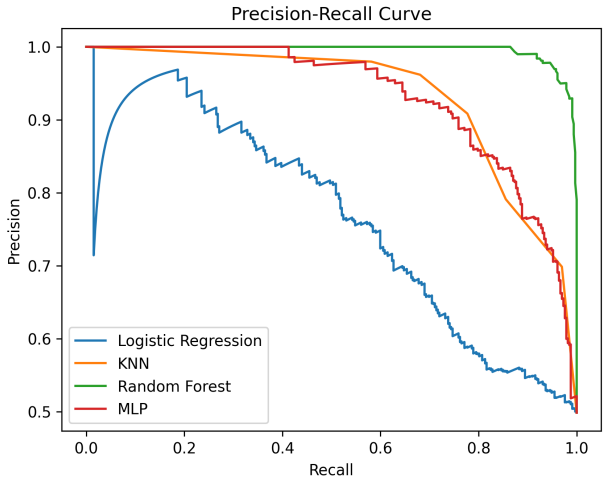



**Figure 5:** *Precision-Recall curves in Train-Test Split*

Random Forest demonstrated its superior ability to retrieve relevant instances without sacrificing precision in the train-test split analysis, as evidenced by its ability to maintain higher precision levels across a range of recall thresholds. While KNN demonstrated remarkable precision but marginally worse recall performance, MLP demonstrated balanced but moderate performance whereas Logistic Regression fell out.

### 4.1.3. Evaluation Metrics Bar chart

Evaluation metrics were used to gauge model performance, depicted in a bar chart represented in figures 6 and 7. Random Forest led with 98.31% accuracy, while KNN followed at 86.62%. Recall ranged from 62.81% (Logistic Regression) to 98.42% (Random Forest), with precision at 98.20% (Random

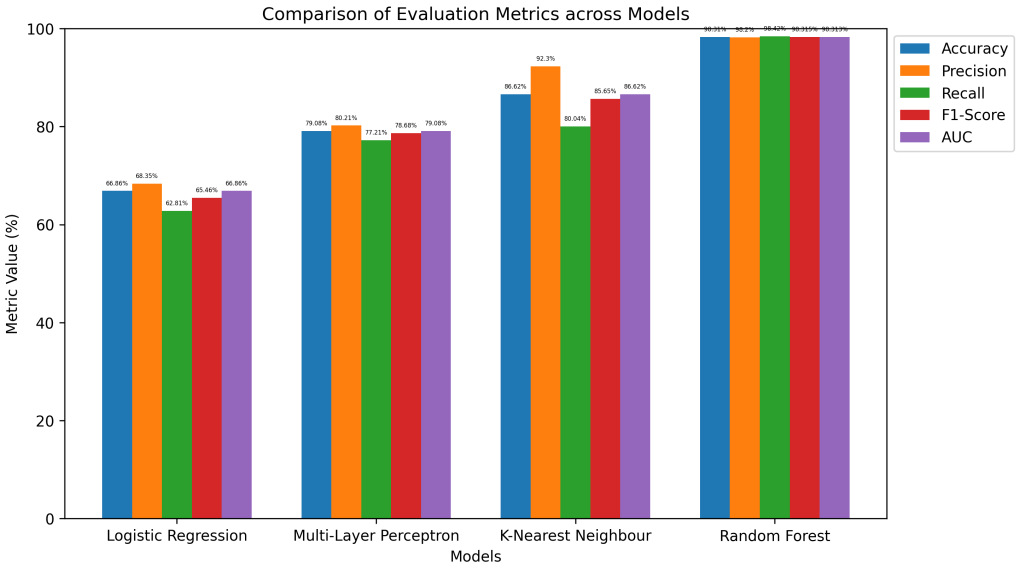



**Figure 6:** *Bar chart representing evaluation metrics for each algorithm in 10-fold cross validation*

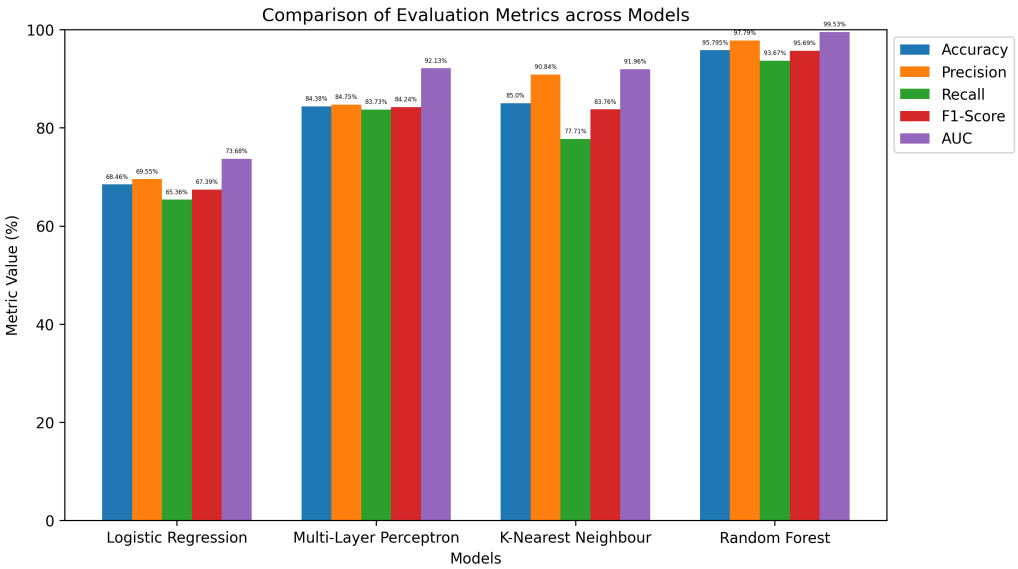



**Figure 7:** *Bar chart representing evaluation metrics for each algorithm in Train-Test Split*

Forest) and 92.30% (KNN). In cross-validation, Random Forest peaked at 98.315% F1-Score, matching its AUC and accuracy. In the train-test split analysis, Random Forest excelled across all metrics: 95.79% accuracy, 97.79% precision, and 99.53% AUC. KNN followed closely with strong precision (90.84%) and accuracy (85%). MLP showed moderate, balanced predictive capabilities whereas Logistic Regression was lagging behind.

#### 4.1.4. Runtime Bar Chart

According to the runtime analysis, the models' computational requirements varied, with KNN requiring the least amount of time (0.05 seconds) and MLP requiring the longest (1.74 seconds) as shown in the figure 8. This difference emphasizes how crucial

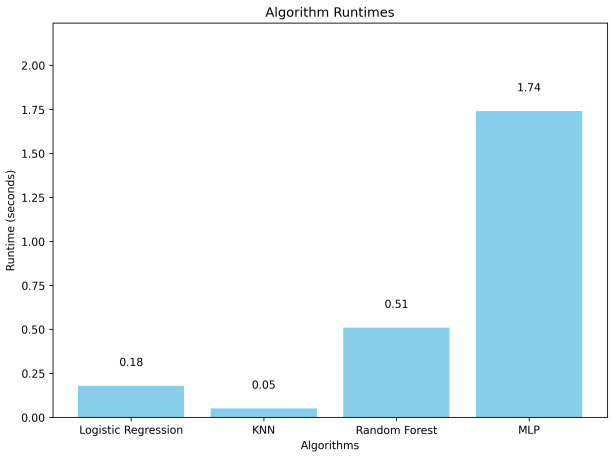



**Figure 8:** *Runtime of different algorithms*

it is to take computational efficiency into account when choosing a model. Our study was carried out on a system that had an Intel Core i-7 with 16 GB RAM and a 10th generation CPU, and it provided insight into how well the model performed in different computing environments. Making informed decisions about which models to use for specific computational configurations is made easier by this short runtime curve.

#### 4.1.5. Calibration Curve

Brier Scores, which measure the precision of predicted probabilities, were carefully used to evaluate the calibration efficiency of the models. The calibration of the logistic regression was deemed acceptable, as evidenced by its moderate Brier Score of 0.2028. In contrast, with a score of 0.1113, KNN demonstrated better calibration. Among them, Random Forest performed the best out of all of them as shown in figure 9, with a considerably lower Brier Score of 0.0351, indicating higher probability prediction accuracy. Comparable to KNN in terms of performance was the MLP, with a Brier Score of 0.1157.

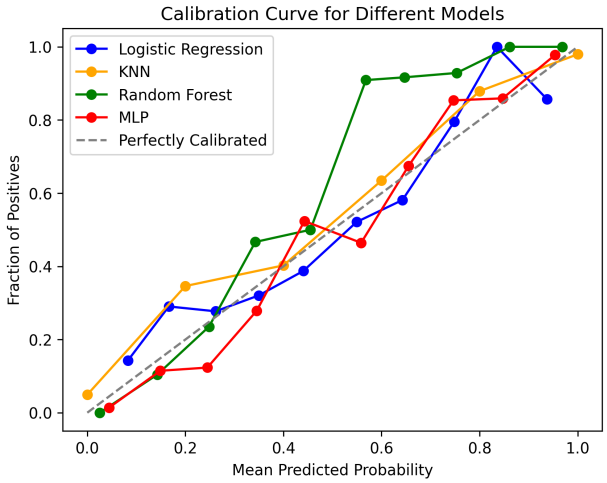



**Figure 9:** *Learning curve for different algorithms in train-test split*

The use of a calibration curve is justified by its capacity to provide information about model reliability by illuminating the relationship between expected probabilities and actual results. Isotonic Regression refines the accuracy of probability estimates by transforming predicted probabilities to closely match observed frequencies, thus improving calibration even further. Interestingly, Random Forest turns out to be the best calibration technique, outperforming MLP, KNN, and Logistic Regression. This assessment provides important information about the models' ability to produce accurate probability estimates, which is crucial for applications that depend on well-calibrated forecasts. The results of the study highlight the need of using reliable calibration methods, with Random Forest demonstrating particularly good results.

#### 4.1.6. Learning Curve for Random Forest

Analyzing learning curves to investigate the Random Forest classifier's performance revealed fascinating dynamics as per figures 10 and 11. The training score, represented by the blue line, demonstrated impressive stability, indicating a constant high accuracy over the course of exposure to the training dataset. By comparison, the cross-validation (CV) score (orange line) started at around 76% and rises steadily as the training dataset grows to 95.79%, suggesting improved generalization. The CV score's upward trajectory indicated that the model skillfully integrated more training data to improve its predictive abilities on cases that had not yet been observed.

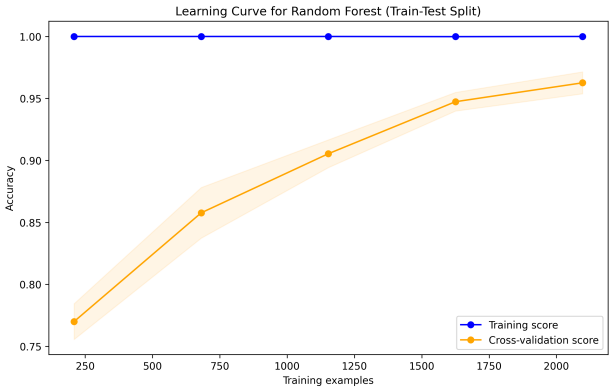



**Figure 10:** *Learning curve for random forest in train-test split*

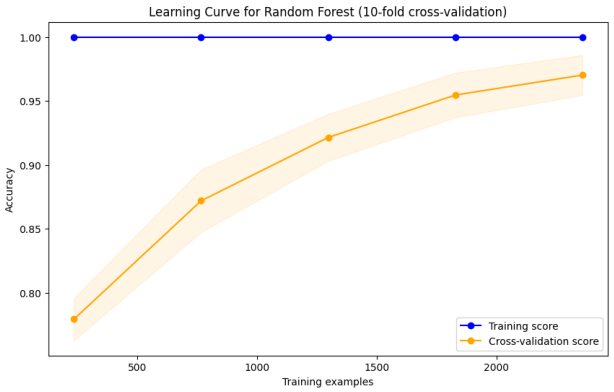



**Figure 11:** *Learning curve for random forest in 10-fold cross validation*

The CV score increased dramatically to 98.31% during the learning curve plotting with a 10-fold cross-validation, starting at about 77%. This additional knowledge strengthens the model's robustness by highlighting its ability to generalize across various folds. A well-generalized model is implied by

the training score's consistent stability and the CV score's convergence towards it. The decreasing difference between the two scores is noteworthy because it highlights the model's resistance to overfitting. The model is especially well-suited for the task at hand because of its balanced learning process, which places it in an adaptive position to handle different training dataset sizes. In conclusion, the learning curve story illustrated a Random Forest model that balanced stability and flexibility, demonstrating complex learning dynamics for successful generalization.

#### 4.1.7. Confusion Matrix for Random Forest

When assessing the RF model for disease prediction, the use of a confusion matrix produced informative results. The matrix, in figure 12, demonstrated the model's ability to correctly identify 313 cases of liver disease (LD) and 329 cases of non-liver disease (NLD). In addition to exposing five instances of misclassifying NLD as LD and 19 cases of misclassifying LD as NLD, this impressive performance was tainted by errors. A confusion matrix, which provides a thorough breakdown of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), was crucial in conducting a thorough evaluation of the model's predictive abilities.

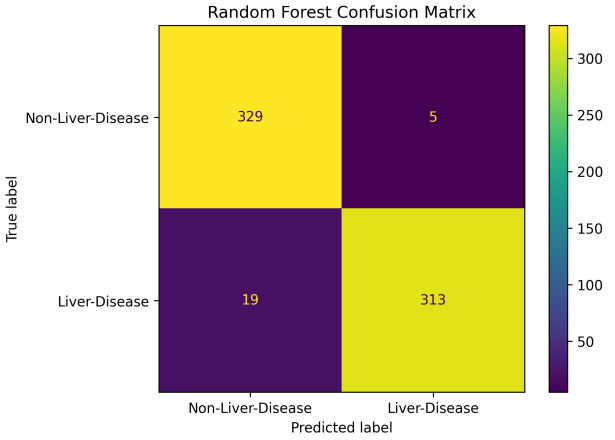



**Figure 12:** *Confusion matrix for random forest in train-test split*

These results not only showed that the model could produce accurate predictions, but they also highlighted some areas that still required work. The confusion matrix was a useful tool for a more accurate and nuanced assessment of the RF model's performance in liver disease prediction because it presented these nuanced metrics.

## 4.2. Performance of different algorithms

When we evaluated the models' performance using cross-validation and train-test splits, the Random Forest model consistently performed better in both cases as shown in tables 2 and 3. With a remarkable accuracy of 95.79% in the train-test split, the Random Forest model outperformed Logistic Regression (68.46%), KNN (85%), and MLP (84.38%). In the same way, the Random Forest continued to outperform the other models in cross-validation, with a mean accuracy of 98.31%. Because of its consistent high performance, the Random Forest model is positioned as dependable and strong for the particular dataset, indicating its potential for precise predictions in real-world applications.

**Table 2:** *Performance of different algorithms in 10-fold cross validation*

| Model | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | AUC (%) |
|---|---|---|---|---|---|
| Logistic Regression | 66.86 | 68.35 | 62.81 | 65.46 | 66.86 |
| MLP | 79.08 | 80.21 | 77.21 | 78.68 | 79.08 |
| KNN | 86.62 | 92.30 | 80.04 | 85.65 | 86.62 |
| **Random Forest** | **98.31** | **98.20** | **98.42** | **98.315** | **98.313** |

**Table 3:** *Performance of different algorithms in train-test split*

| Model | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | AUC (%) |
|---|---|---|---|---|---|
| Logistic Regression | 68.46 | 69.55 | 65.36 | 67.39 | 73.68 |
| MLP | 84.38 | 84.75 | 83.73 | 84.24 | 92.13 |
| KNN | 85.00 | 90.84 | 77.71 | 83.76 | 91.96 |
| **Random Forest** | **95.79** | **97.79** | **93.67** | **95.69** | **99.53** |

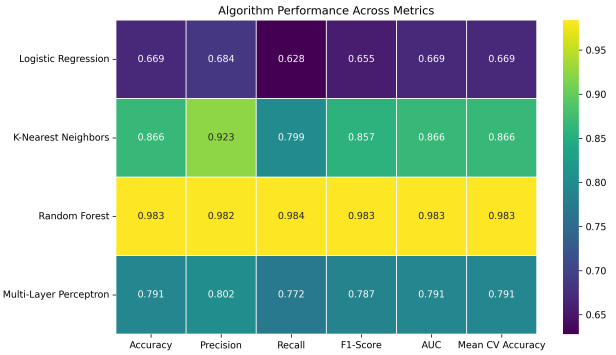



**Figure 13:** *Heatmap for different algorithms in 10-fold cross validation*

The Random Forest's consistent performance in a variety of assessment methods highlights both its adaptability to generalize well on unknown data and its resistance to overfitting. Because of its resilience, it is a sensible option for predictive modeling in situations similar to the dataset under study.Our analysis's results highlight the Random Forest's potential for precise forecasts, promoting its use in real-world applications that demand strong predictive abilities.

## 4.3. Comparative Analysis

In comparison to recent state-of-the-art studies, this comparative evaluation in table 4 highlights the superior performance of our proposed method on the ILPD dataset and demonstrates its efficacy in achieving high predictive accuracy. The table shows that our strategy—which used Random Forest (RF) in particular—achieved the highest accuracy of 98.31%, outperforming all other approaches. Interestingly,

**Table 4:** *Performance comparison of recent state-of-the-art studies with proposed method on the ILPD dataset*

| Author(s) | Year | Protocol | Classifiers | Accuracy (%) |
|---|---|---|---|---|
| Amin et al. | 2023 | Cross Validation | RF<br>MLP<br>Ensemble<br>KNN<br>SVM<br>LR | 88.10<br>83.53<br>82.09<br>67.90<br>67.90<br>55.40 |
| Niha et al. | 2023 | Train-test Split | LR<br>RF<br>KNN | 76.575<br>74.28<br>70.20 |
| Anthonysamy and Babu | 2023 | Train-test Split | HVC<br>MLP<br>SVM<br>KNN | 78.62<br>77.24<br>76.55<br>73.10 |
| Bhupathi et al. | 2022 | CRISP-DM | Auto encoders<br>KNN<br>CART<br>SVM<br>LDA<br>Naive Bayes | 92.1<br>91.7<br>83.6<br>78.1<br>70.9<br>65.1 |
| Sivasangari et al. | 2022 | Train-test Split, Cross Validation | SVM<br>RF<br>DT | 95.18<br>92.77<br>87.95 |
| Ghosh et al. | 2021 | Train-test Split | RF<br>XGBoost<br>SVM<br>KNN<br>LR<br>DT<br>AdaBoost | 83.76<br>82.05<br>81.20<br>81.20<br>79.49<br>79.49<br>76.07 |
| Panwar et al. | 2021 | Train-test Split | SVM<br>DT<br>RF<br>LR<br>Naive Bayes | 74.09<br>72.54<br>72.02<br>71.50<br>57.51 |
| Kuzhippallil et al. | 2020 | Genetic Algorithm | RF<br>XGBoost<br>Light GBM<br>SE<br>DT<br>GB<br>AdaBoost<br>MLP<br>KNN<br>LR | 88<br>86<br>86<br>85<br>84<br>84<br>83<br>82<br>79<br>76 |
| **Proposed Method** | | **Cross Validation** | **RF**<br>KNN<br>MLP<br>LR | **98.31**<br>86.62<br>79.08<br>66.86 |

KNN and MLP also showed significant accuracy at 86.62% and 79.08%, respectively, proving the effectiveness of our approach.

RF demonstrated the highest accuracy at 88.10% when used in a cross-validation protocol by Amin et al. (2023), followed by MLP at 83.53%. Niha et al. (2023) examined RF, KNN and LR in train-test split. LR obtained an 87.16% F1 score, 76.57% accuracy, 75.00% precision, and 99.50% recall. KNN showed 84.32% F1 score, 84.90% recall, 83.10% precision, and 70.20% accuracy. RF displayed an 80.32% F1 score, 77.90% recall, 84.40% precision, and 74.28% accuracy. The study evaluated TP, TN, FP, and FN rates with a focus on liver disease classification. These outcomes demonstrated the effectiveness of the algorithms and offered insightful information about how they might be used for machine learning tasks involving classification and predictive modeling.

In order to predict liver disease, Anthonysamy and Babu (2023) used machine learning algorithms such as RF, KNN, and LR. Scaling, column elimination, and KNN-imputer for null values were all part of the data preprocessing. The Hard Voting Classifier (HVC) and MLP were introduced. HVC came out to be better than other classifiers, according to the results, with 78.62% accuracy, 80% specificity, 87% F-score, 78% precision, and 78% recall. 77% accuracy, 100% specificity, 87% F-Score, 77% precision, and 100% recall were attained by MLP. 96% recall, 76% accuracy, and 96% specificity were attained by SVM using an RBF kernel. KNN showed 89% recall, 73% accuracy, and 89% specificity. Bhupathi et al. (2022) carried out a study in 2022 to evaluate different classifiers in compliance with the CRISP-DM (Cross-Industry Standard Process for Data Mining) protocol. Notably, their results showed that autoencoders, with an astounding accuracy score of 92.1%, emerged as the most accurate classifier. Using a train-test split and cross-validation, Sivasangari et al. (2022) found that SVM had the highest accuracy, at 95.18%. Ghosh et al. (2021) investigated liver disease prediction using RF, KNN, SVM, Decision Tree, and other ensemble methods. The methodology included feature scaling as a necessary preprocessing step prior to training datasets with sizes ranging from 50% to 90%. RF was clearly the most successful algorithm, with accuracy of 83.76%, precision of 87%, recall of 93.5%, and an F1 score of 90.1%. RF's AUC was 81.3%.

Panwar et al. (2021) used advanced pre-processing techniques, such as replacing missing values with null values and their corresponding instances, in order to detect chronic liver disease. An approach that combined filter and wrapper techniques was used to carefully select features. Through correlation analysis, attributes with a correlation of greater than

70% were first disregarded. Training (70%) and test (30%) sets were created from the dataset through randomization. SVM achieved an impressive accuracy of 74.09%. Various classification algorithms were implemented, including Naïve Bayes, LR, RF, and Support Vector Machine. Kuzhipallil et al. (2020) looked at preliminary data analysis, using data preprocessing methods like imputation, label encoding, duplicate value removal, resampling, and outlier detection, as well as data exploration to summarize and visualize patterns. Feature selection was made easier by genetic algorithms, which optimized input features for predictive models. Several machine learning algorithms, including shortened versions like RF and XGBoost, were used for classification. LightGBM and Stacking Estimator showed the highest accuracy (88%) in the study, highlighting enhanced performance following feature selection and outlier removal.

Using a cross-validation protocol, proposed method outperformed all other studies with RF obtaining an astounding accuracy of 98.31%. Significantly, KNN and MLP also showed good accuracy at 86.62% and 79.08%, respectively, proving the effectiveness of the suggested methodology.

## 5. Conclusion

Our thorough investigation of ML algorithms for the diagnosis of CLD highlights the efficacy of various strategies. Notably, with consistently high accuracy rates, the RF model proved to be the most successful. Although preprocessing steps like replacing outliers and oversampling are important, the main focus of this study is the unified approach that combined non-linear (t-SNE, UMAP) and linear (LDA, FA) techniques for dimensionality reduction and feature integration. The reliability of our evaluations was guaranteed by carefully verifying and improving model performance using train-test splits and 10-fold cross-validation.

Subsequent studies in the field of chronic liver disease detection may explore sophisticated anomaly detection strategies and ensemble approaches, thereby expanding the potential to detect subtle indicators of the disease. The potential to strengthen the effectiveness of predictive models by adding extra clinical data beyond non-standard laboratory testing is promising. Furthermore, incorporating more dimensionality reduction techniques like PCA, and incremental feature selection could provide a more in-depth understanding of the features of the disease. Deep Belief Networks (DBNs), autoencoders, Non-negative Matrix Factorization (NMF), Random Forest using Particle Swarm Optimization (RF with PSO) could also be used. There is a chance that the detection of chronic liver disease will advance significantly with the use of these nuanced data analysis techniques, which seek to improve diagnostic precision and broaden the coverage of detection algorithms.

**Declaration**

No subject testing or data collection procedures were taken into consideration for this study because it uses data that is publicly available or data from published sources. The authors declare unequivocally that all the work reported in this paper could not have been influenced by any known competing financial interest or personal relationship .

**Authors contribution**
All authors contributed equally.

**Availability of data and materials**
Data will be made available on request.

## References


1. Cheemerla, S., & Balakrishnan, M. (2021b). Global Epidemiology of Chronic Liver Disease. Clinical Liver Disease, 17(5) https://doi.org/10.1002/cld.1061
2. Asrani, S. K., Devarbhavi, H., Eaton, J. E., & Kamath, P. S. (2019). Burden of liver diseases in the world. Journal of Hepatology, 70(1), 151–171. https://doi.org/10.1016/j.jhep.2018.09.014
3. Umbare, R., Ashtekar, O., Nikhal, A., Pagar, B., & Zare, O. (2023). Prediction & Detection of Liver Diseases using Machine Learning. IEEE 3rd International Conference on TEMSMET. https://doi.org/10.1109/temsmet56707.2023.10150135
4. Javaid, M., Haleem, A., Singh, R. P., Suman, R., & Rab, S. (2022). Significance of machine learning in healthcare: Features, pillars and applications. International Journal of Intelligent Networks, 3, 58–73. https://doi.org/10.1016/j.ijin.2022.05.002
5. Liu, Y., & Chen, M. (2022). Epidemiology of liver cirrhosis and associated complications: Current knowledge and future directions. World Journal of Gastroenterology, 28(41), 5910–5930. https://doi.org/10.3748/wjg.v28.i41.5910
6. Younossi, Z. M., Blissett, D. B., Blissett, R., Henry, L., Stepanova, M., Younossi, Y., Racila, A., Hunt, S., & Beckerman, R. (2016). The economic and clinical burden of nonalcoholic fatty liver disease in the United States and Europe. Hepatology, 64(5), 1577–1586. https://doi.org/10.1002/hep.28785

7. Popa, S., Ismaiel, A., Abenavoli, L., Padureanu, A. M., Dita, M. O., Bolchis, R., Munteanu, M., Brata, V. D., Pop, C., Bosneag, A., Dumitraşcu, D. I., Bârsan, M., & David, L. (2023). Diagnosis of liver fibrosis using Artificial Intelligence: A Systematic review. Medicina-lithuania, 59(5), 992. https://doi.org/10.3390/medicina59050992
8. Ahn, J., Connell, A., Simonetto, D. A., Hughes, C., & Shah, V. H. (2021). Application of artificial intelligence for the diagnosis and treatment of liver diseases. Hepatology, 73(6), 2546–2563. https://doi.org/10.1002/hep.31603
9. Bhupathi D, Tan C N-L, Tirumula S.S, Ray S.K. (2022) Liver Disease Detection using Machine Learning Techniques. The Computing and Information Technology Research and Education New Zealand (CITRENZ) Springer
10. Mohan, V., Dhayanand, S. (2015). Liver Disease Prediction using SVM and Naïve Bayes Algorithms.
11. Wu, J., Lin, S., Wan, B., Velani, B., & Zhu, Y. (2019). Pyroptosis in Liver Disease: New Insights into Disease Mechanisms. Aging and Disease, 10(5), 1094. https://doi.org/10.14336/ad.2019.0116
12. Amin, R., Yasmin, R., Ruhi, S., Rahman, M. H., & Reza, M. S. (2023). Prediction of chronic liver disease patients using integrated projection based statistical feature extraction with machine learning algorithms. Informatics in Medicine Unlocked, 36, 101155. https://doi.org/10.1016/j.imu.2022.101155
13. Batarseh, F. A., & Freeman, L. (2022). AI Assurance: Towards Trustworthy, Explainable, Safe, and Ethical AI. Elsevier. ISBN 978-0-323-91919-7.
14. Tavakol, M., & Wetzel, A. P. (2020). Factor Analysis: a means for theory and instrument development in support of construct validity. International Journal of Medical Education, 11, 245–247. https://doi.org/10.5116/ijme.5f96.0f4a
15. Cai, T., & Ma, R. (2021). Theoretical Foundations of T-SNE for Visualizing High-Dimensional Clustered Data. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2105.07536
16. Singh, J., Bagga, S., & Kaur, R. (2020). Software-based Prediction of Liver Disease with Feature Selection and Classification Techniques. Procedia Computer Science, 167, 1970–1980. https://doi.org/10.1016/j.procs.2020.03.226
17. Wang, N., Yu, Y., Huang, D., Xu, B., Liu, J., Li, T., Xue, L., Zengyu, S., Chen, Y., & Wang, J. (2015). Pulse Diagnosis signals analysis of fatty liver disease and cirrhosis patients by using machine learning. The Scientific World Journal, 2015, 1–9. https://doi.org/10.1155/2015/859192
18. Kumar, Y., & Sahoo, G. (2013). Prediction of different types of liver diseases using rule based classification model. Technology and Health Care, 21(5), 417–432. https://doi.org/10.3233/thc-130742
19. Dritsas, I., & Trigka, M. (2023). Supervised Machine learning models for liver disease risk prediction. Computers, 12(1), 19. https://doi.org/10.3390/computers12010019
20. Muthuselvan, S., Rajapraksh, S., Somasundaram, K., & Karthik, K. (2018). Classification of liver patient dataset using machine learning algorithms. International Journal of Engineering & Technology, 7(3.34), 323. https://doi.org/10.14419/ijet.v7i3.34.19217
21. Babu, M. S. P., Ramjee, M., Katta, S., & Swapna, K. (2016). Implementation of partitional clustering on ILPD dataset to predict liver disorders. ICSESS. https://doi.org/10.1109/icsess.2016.7883256
22. Sujith, J. G., Kumar, P., Reddy, S. J. M., & Kanhe, A. (2023). Computative analysis of various techniques for classification of liver disease. Journal of Physics: Conference Series, 2466(1), 012035. https://doi.org/10.1088/1742-6596/2466/1/012035
23. Sivasangari, A., Reddy, B. J. K., Kiran, A., & Ajitha, P. (2020). Diagnosis of Liver Disease using Machine Learning Models. 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC). https://doi.org/10.1109/i-smac49090.2020.9243375
24. McKnight, P. E., McKnight, K. M., Sidani, S., & Figueredo, A. J. (2007). Missing data: A Gentle Introduction. Guilford Press
25. Alpaydin, E. (2014). Introduction to machine learning. MIT Press.
26. Muixí, A., Garcia-Gonzalez, A., Zlotnik, S., & DiEz, P. (2023). Linear and nonlinear dimensionality reduction of biomechanical models. In Elsevier eBooks (pp. 23–44). https://doi.org/10.1016/b978-0-32-389967-3.00004-4
27. Tharwat, A., Gaber, T., Ibrahim, A., & Hassanien, A. E. (2017). Linear discriminant analysis: A detailed tutorial. Ai Communications, 30(2), 169–190. https://doi.org/10.3233/aic-170729
28. Smelser, N. J., & Baltes, P. B. (2001). International Encyclopedia of the Social & Behavioral Sciences.
29. Johnson, R. A., & Wichern, D. W. (2007). Applied Multivariate Statistical analysis. Prentice Hall.
30. Zhou, H., Wang, F., & Tao, P. (2018). t-Distributed Stochastic Neighbor Embedding Method with the Least Information Loss for Macromolecular Simulations. Journal of Chemical Theory and Computation, 14(11), 5499–5510.

https://doi.org/10.1021/acs.jctc.8b00652
31. van der Maaten, L. J. P., & Hinton, G. E. (2008). Visualizing High-Dimensional Data Using t-SNE. Journal of Machine Learning Research, 9(nov), 2579-2605.
32. McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform Manifold Approximation and Projection for dimensionality reduction. arXiv preprint arXiv:1802.03426.
33. Vaidya, V., & Vaidya, J. (2022). Impact of dimensionality reduction on Outlier Detection: an Empirical Study. IEEE 4th Int Conf Trust Priv Secur Intell Syst Appl (2022). https://doi.org/10.1109/tps-isa56441.2022.00028
34. Halladin-Dabrowska, A., Kania, A., & Kopeć, D. (2019). The T-SNE algorithm as a tool to improve the quality of reference data used in accurate mapping of heterogeneous Non-Forest vegetation. Remote Sensing, 12(1), 39. https://doi.org/10.3390/rs12010039
35. Niha, S. A., Lakshmi, K. J., Blessi, P. G., Lakshmi, T. S., Chowdary, Y. M., & Rao, M. P. (2023). A comparison of machine learning algorithms for predicting liver disease. International Journal for Research in Applied Science and Engineering Technology, 11(3), 2221–2230. https://doi.org/10.22214/ijraset.2023.49954
36. Anthonysamy, V., & Babu, S. K. K. (2023). Multi perceptron neural network and voting classifier for liver disease dataset. IEEE Access, 11, 102149–102156. https://doi.org/10.1109/access.2023.3316515
37. Ghosh, M., Raihan, M. M. S., Raihan, M., Akter, L., Bairagi, A. K., Alshamrani, S. S., & Masud, M. (2021). A comparative analysis of machine learning algorithms to predict liver disease. Intelligent Automation and Soft Computing, 30(3), 917–928. https://doi.org/10.32604/iasc.2021.017989
38. Panwar, V., Choudhary, N., Mittal, S., & Sahu, G. (2021). Review of liver disease prediction using machine learning algorithm. Journal of Emerging Technology and Innovative Research (JETIR), 8(2), 2349–5162.
39. Kuzhippallil, M. A., Joseph, C., & Kannan, A. (2020). Comparative Analysis of Machine Learning Techniques for Indian Liver Disease Patients. ICACCS. https://doi.org/10.1109/icaccs48705.2020.9074368

This article has been accepted for publication in IEEE Transactions on Nanotechnology. 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/TNANO.2025.3531937



# Unveiling the Impact of Interfacial Trap Charges on Strained VS-FeFinFETs for Improved Reliability:Device to Circuit level Assessment

Kajal Verma, *Student Member, IEEE,* Rishu Chaujar, *Senior Member, IEEE,*

***Abstract*—This paper focuses on the device to circuit level assessment of Si/SiGe strained vertically stacked ferroelectric based FinFETs (VS-FeFinFETs) for improved reliability under the influence of interfacial trap charges (ITCs) at the semiconductor/oxide interface. The device is designed with the amalgamation of several advanced technologies such as SOI, strained tri-layered silicon channel system along with the integration of ferroelectric material in superior gate controlled FinFET. Gate engineering has also been incorporated to further improve the device's reliability against ITCs, forming hetero dielectric vertically stacked ferroelectric based FinFET (HD-VS-FeFinFET) and it is found to possess superior analog, linearity, and harmonic distortion performance. It shows 91.48% reduction in leakage current resulting in 13 times increment in switching ratio along with improvement in quality factor by 46.01%, transconductance by 32.77%, and device efficiency by 26.54% with negligible variations due to ITCs as compare to VS-FeFinFET. Various linearity and harmonic parameters also improved and showed negligible average variations like 4.72% (177.15%) in VIP2 and 6.525% (25.3%) in 1-dB compression point for HD-VS-FeFinFET (VS-FeFinFET) against different ITCs polarity making it more reliable for low power microwave and distortion less wireless communication applications. Further logic circuit application of HD-VS-FeFinFET based CMOS inverter has been analysed and it shows improvement by 17.9% in transition range, 51.674% in voltage gain along with ITCs induced average variation of 3.66% (15.88%) in noise margin for HD-VS-FeFinFET(VS-FeFinFET) based circuit thus led to its development with enhanced functionality, reliability, and performance, poised to shape the landscape of modern electronics.**

***Index Terms*—FinFET, Ferroelectric, Strain, Si/SiGe, CMOS, HD-VS-FeFinFET, Interfacial trap charges, Reliability, Inverter**

## I. Introduction

TECHNOLOGY advancements like Internet of Things, mobile smart devices, and logic in memory (LiM) computing have increased the demand for high-performance and low-power electronic devices[1]. To meet this increasing demand, devices using CMOS technology are being rapidly scaled down and they are approaching their fundamental limits now. In such a scenario, various short channel effects (SCEs) begin to predominate at submicron dimensions and degrade the device performance[2].

As documented in different works of literature, several engineering techniques and device designs, including dual gates, recessed channels, trenched gates, trigate FinFETs, FeFETs, etc., have been used to get around these restrictions[3]. Among these structures, FeFinFET(Ferroelectric based FinFET) is considered a viable contender for the ultimate CMOS device structure due to its resilience against SCEs and enhanced transistor performance. By blending the advantages of FinFET technology with the unique qualities of ferroelectric materials, the FeFinFET marks a burgeoning field within semiconductor technology[4]. Further to lessen the vulnerability to negative effects of parasitic capacitance, SOI (Silicon On Insulator) technology is adopted in the current structure to create the SOI FeFinFET. In the era of energy-efficient computing, this isolation improves the transistor's electrostatic control, enabling more effective switching and lowering the power consumption[5].

Silicon has been the cornerstone of the semiconductor industry for decades, but its limitations in meeting the demand of progressing technology make it necessary to look beyond silicon channel materials like SiGe, GaAs, group III-IV semiconductors, and alloy-based materials to further enhance the performance of SOI FeFinFET. In our previous work, we have analyzed the performance of vertically stacked ferroelectric-based FinFET involving a strained-silicon channel system using a compound semiconductor material (SiGe). The results showed that the device has showcased improved performance against SCEs and can be further scaled down as compared to FeFinFET[6].

At the heart of FeFinFETs lies the interface between oxide and semiconductor layers, where the dynamics of interfacial trap charges (ITCs) exert a profound influence on device performance and functionality. The complexity introduced by ITCs at the semiconductor-oxide (S/O) interface in FeFinFETs calls for careful research. The dynamics of these ITCs are affected by various elements such as material composition, processing techniques, and operational environment. The OFF-state performance deteriorates and the threshold voltage is shifted by these trap charges[7]. To maximize the performance of device and realize their full potential for a variety of applications, it is imperative to comprehend the behavior of ITCs. There are various factors responsible for occurrence of these ITCs. Firstly, the creation of dangling bonds due to disruption of atomic lattice at the S/O interface which leads to creation of energy states with in the bandgap of the semiconductor. Secondly, contamination of the surface with impurities can introduce localized states that trap charges[8]. Further, other factors may be thermal oxidation process during

This work was funded in part by the Council of Scientific and Industrial Research under Grant 08/0133(13434)/2022-EMR-I for Kajal Verma.

Kajal Verma and Rishu Chaujar are with Department of Applied Physics, Delhi Technological University, Delhi, 110042, India (e-mail: kajalverma_2k21phdap507@dtu.ac.in; chaujar.rishu@dtu.ac.in).

Copyright (c) 2025 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:48:23 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Nanotechnology. 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/TNANO.2025.3531937



growth of oxide layer, high energy radiation during ion implantation, electrical stress during operation or testing, mechanical stress during deposition or thermal expansion can break bonds or introduce defects, ultimately creating ITCs. Controlling ITCs is crucial for enhancing the performance and reliability of semiconductor devices. Various potential methods can be incorporated in the devices to mitigate these ITCs such as passivation techniques, optimized thermal oxidation, surface cleaning before deposition, high-k dielectric materials, low-damage deposition techniques like atomic layer deposition (ALD), etc[9]. In this paper, gate stack engineering with heterogeneous gate dielectric ($Al_2O_3$+$HfO_2$) is incorporated in vertically stacked ferroelectric based FinFET to form HD-VS-FeFinFET, a novel combination of Si/SiGe strained tri-layered channel-engineered, heterostructure on insulator FeFinFET, and has been examined to minimize the effect of ITCs. The combination leverages the passivation properties of $Al_2O_3$ and the high-k benefits of $HfO_2$[10]. $Al_2O_3$ is used as an interfacial layer ($t_{ox1}$), which is known for high-quality interface with semiconductors leading to reduced dangling bonds and ITCs. $Al_2O_3$ has a high bandgap ($\sim$9 eV), which minimizes creation of energy states, further reducing trap density at the interface[11]. Then, the high-k nature of $HfO_2$ reduces the electric field stress at the interface, thereby minimizing the generation of new traps. $HfO_2$ also exhibits good thermal stability, reducing defect formation during high-temperature processing. Thus, the stacked structure minimizes the density of trap charges at both the interface and within the dielectric layers. These dielectrics are also compatible with atomic layer deposition (ALD) processes, allowing precise thickness and composition control[12]. Thus, the combination of $Al_2O_3$ and $HfO_2$ uses their complementary properties, such as $Al_2O_3$'s superior passivation and $HfO_2$'s high-k dielectric performance, make them ideal for next-generation electronic devices requiring high efficiency and robust reliability[10].

In this article, the detailed analysis of static, analog, linearity, and harmonic distortion performance under the influence of ITCs is done for both the devices. Further, the work is extended to discuss the influence of ITCs at the circuit level by comparing the performance of HD-VS-FeFinFET based CMOS inverter with VS-FeFinFET based CMOS inverter for improved efficiency and reliability. This work is divided into five sections: Section II includes all the specifications regarding the device structure; Section III explains the simulation framework and fabrication feasibility; Section IV provides the results and discussion; and Section V concludes the paper.

## II. Device Structure

Figure 1(a) depicts the 3-dimensional structure of HD-VS-FeFinFET, while Fig. 1(b) and Fig. 1(c) respectively shows the device's vertical and lateral 2-dimensional view chopped through its fin. The gate length ($L_g$) considered for the device is fixed at 20nm where silicon-doped hafnium oxide ($HfO_2FE$) is used as the ferroelectric material with thickness ($t_{fe}$) fixed at 6 nm with total gate oxide dielectric thickness ($t_{ox}$) fixed at 1nm. The total gate oxide is composed of a heterogeneous dielectric stack consisting of equal proportion of $Al_2O_3$ and

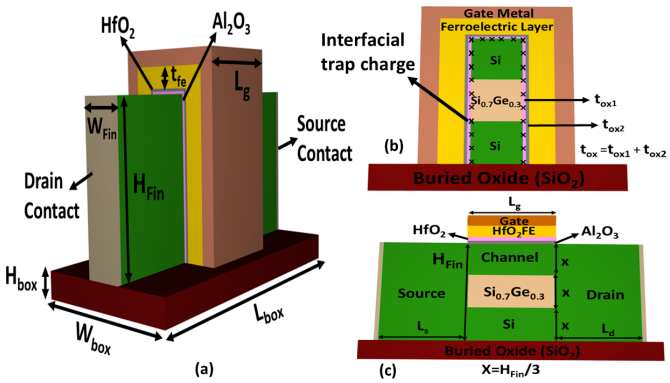



Fig. 1: (a) 3D schematic view of HD-VS-FeFinFET (b) vertical cross-sectional view of HD-VS-FeFinFET (c) lateral cross-sectional view of HD-VS-FeFinFET

$HfO_2$, each with a thickness of 0.5nm. So, the combined thickness of ferroelectric layer and gate oxide dielectric is 7nm. The height and width of the fin are taken as 30nm and 8 nm respectively. A tri-layered strained-silicon channel system is formed with the help of SiGe sandwiched between two silicon layers as Si-$Si_{0.7}Ge_{0.3}$-Si (where $Si_{0.7}Ge_{0.3}$ represents that the mole fraction of Si and Ge is taken as 0.7 and 0.3 respectively). The doping concentration of source/drain is fixed as $1\times10^{20}$ $cm^{-3}$ with n-type and that of channel is $1\times10^{16}$ $cm^{-3}$ with p-type with uniform profile. The length of the source/drain electrodes ($L_{s/d}$) is fixed at 20nm. In addition, a metal gate with a work function of 4.65 eV, which has excellent thermal stability and is compatible with CMOS processing, is utilized in this device to prevent any poly-depletion effect that is common in polysilicon gates. Further, to exploit the advantages of SOI technology, the entire heterostructure FeFinFET is built on an insulator $SiO_2$ box whose width, height, and length are kept 32nm, 5nm, and 62nm respectively. To unveil the intricacies of ITCs on the device reliability and performance, ITCs with different polarity have been introduced at the S/O interface with trap charge density to be fixed at $10^{12}$ $cm^{-2}$ based on formerly reported literature[7], [8], [11].

## III. Simulation Framework and Fabrication Feasibility

All the simulations have been carried out using Genius 3D TCAD simulator by Cogenda. For the validation of the models used in the simulator tool, the TCAD simulation result for Silicon on Insulator (SOI) FinFET structure has been calibrated with experimental data[13], and a reasonably good agreement between the two results validates the various models used in the study as displayed by Fig. 2(a). The simulation setup includes the drift-diffusion model level 1 (DDML1) which keeps the lattice temperature constant throughout the solving procedure, coupled with consistent Poisson and continuity equations to govern the carrier transport phenomena for the holes and the electrons. The temperature and doping-dependent carrier recombination-generation phenomenon are accurately captured by Shockley-Read-Hall (SRH) recombination and Auger models[14]. Further, Lombardi mobility model for silicon and Philips mobility model for SiGe are used to

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:48:23 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Nanotechnology. 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/TNANO.2025.3531937



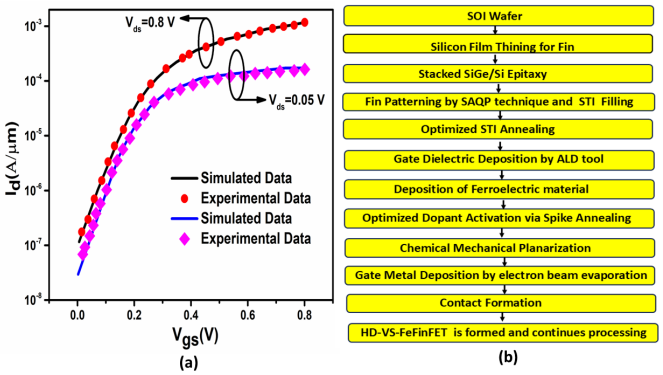



Fig. 2: (a) Experimental and simulated transfer characteristic for Silicon on Insulator (SOI) FinFET at $V_{ds}$=0.05V and $V_{ds}$=0.8V (b) Fabrication process flow chart for HD-VS-FeFinFET, where SiGe is silicon-germanium, SAQP is self-aligned quadruple patterning, STI is shallow trench isolation, and ALD is atomic layer deposition technique

consider the mobility degradation at the semiconductor-oxide interface[6]. Lucent model is used to consider the high-field mobility effects along with hot carrier models for capturing the effect of hot carriers. To check the velocity saturation and mobility of carriers, the velocity saturation model, and Esurface models are activated. The density gradient method is also used to incorporate quantum confinement phenomena along with the Fermi-Dirac carrier statistics model[15]. Further, the position and polarity of ITCs in the device are mentioned using the INTERFACE statement along with Trap models to capture the associated mechanisms[9][15]. A few abbreviations are used in this paper such as PITC - positive interfacial trap charge, NITC - negative interfacial trap charge, WITC - without (neutral) interfacial trap charge.

Further, Fig. 2(b) displays the fabrication feasibility of the HD-VS-FeFinFET using a step-by-step fabrication process flowchart. As the starting material SOI wafer is used with a thick buried oxide layer and thick silicon film followed by silicon film thinning for fin. Afterwards stacked SiGe/Si tri-layer were epitaxially grown by reduced-pressure chemical vapor epitaxial deposition. Thereafter, fin patterning is done using the self-aligned quadruple patterning (SAQP) technique followed by shallow trench isolation (STI) filling and an optimized STI densification annealing which is performed with a rapid thermal annealing (RTA) at an optimized temperature. The gate dielectric is deposited on the silicon interfacial layer by atomic layer deposition (ALD) followed by ferroelectric layer deposition over the gate oxide using ALD technique. The drain and source regions are implanted, and the dopants of these regions are activated using spike annealing followed by chemical mechanical planarization for smoothening the surface. Metal gate is deposited using electron beam evaporation at room temperature on the top of the ferroelectric layer. The source/drain metal contacts are deposited by electron beam evaporation followed by lift-off process. This completes the process of manufacturing the HD-VS-FeFinFETs[16][17].

## IV. Result and Discussion

### A. Influence of ITCs on static and analog performance

In this section, the effect of ITCs polarity at the S/O interface is analysed to study the static and analog performance for both the devices (HD-VS-FeFinFET and VS-FeFinFET) at supply voltage $V_{ds}$=0.4V and at constant temperature 300K. Fig. 3 displays the transfer characteristics of both the devices with different polarity. The presence of ITCs leads to altercation in flat band voltage ($V_{fb}$) as given by equation 1, where q is the electronic charge, $N_{it}$ is the ITCs density and $C_{ox}$ is gate oxide capacitance.

$$V_{\mathrm{fb}} = \frac{qN_{\mathrm{it}}}{C_{\mathrm{ox}}} \tag{1}$$

The changes in $V_{fb}$ caused by PITC/NITC leads to modification in threshold voltage ($V_{th}$) as related by equation 2, where $x_t$ is maximum space charge width and $\psi_{fp}$ is built in potential.

$$V_{\mathrm{th}} = V_{\mathrm{fb}} + \psi_{\mathrm{fp}} + \frac{qN_{\mathrm{a}}x_{\mathrm{t}}}{C_{\mathrm{ox}}} \tag{2}$$

This modification in $V_{th}$ leads to change in device characteristics thus change the ON current ($I_{on}$) and leakage current ($I_{off}$) of the device as shown in Fig. 3(a) [18]. The use of high-k dielectric led to reduced interface states and traps at the S/O interface which improves the carrier mobility within the channel region as can be very well observed in elevated $I_{on}$ of HD-VS-FeFinFET. The PITC (NITC) increases (decreases) the ON current by 0.59% (0.62%) and 3.35% (3.64%) in HD-VS-FeFinFET and VS-FeFinFET respectively. $I_{off}$ with NITC (PITC) decreases (increases) by 11.01% (10.02%) in HD-VS-FeFinFET as comparison to deviation in VS-FeFinFET in presence of PITC (NITC) by 111% (53.1%), as depicted in Fig. 3(b) which confirms that the transfer characteristics of HD-VS-FeFinFET are less affected by ITCs, proving it more reliable in contrast to VS-FeFinFET. Fig. 4(a) shows the variation of output characteristics with drain bias for different ITC polarity. Owing to the suppression of channel hot carrier injection with the incorporation of hetero dielectric gate engineering in HD-VS-FeFinFET, there is negligible impact on output conductance ($g_d$) due to ITCs as showcased in Fig. 4(b), thereby confirming its improved reliability.

Fig. 5(a) and Fig. 5(b) respectively shows the variation of transconductance ($g_m$) and device efficiency, also known as Transconductance Generation Factor (TGF) of both the devices for different ITC polarity. These are the important parameters in the design of analog circuits and Fig. 5 shows higher value of $g_m$ and TGF for HD-VS-FeFinFET due to improved carrier mobility and better electrostatic control along with minimum variations in presence of ITCs thus depicts its enhanced efficiency and reliability of converting DC parameter ($I_d$) into AC parameters ($g_m$) as compared to its counterpart[19].

PITC(NITC) acts to decrease (increase) the threshold voltage ($V_{th}$) but the variation in $V_{th}$ is minute in case of HD-VS-FeFinFET whereas the $V_{th}$ for PITC(NITC) is decreased (increased) by 8.9% (9.3%) in VS-FeFinFET as displayed in Fig. 6(a). ITCs introduce additional energy states within the bandgap of semiconductor material which can capture and

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:48:23 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Nanotechnology. 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/TNANO.2025.3531937



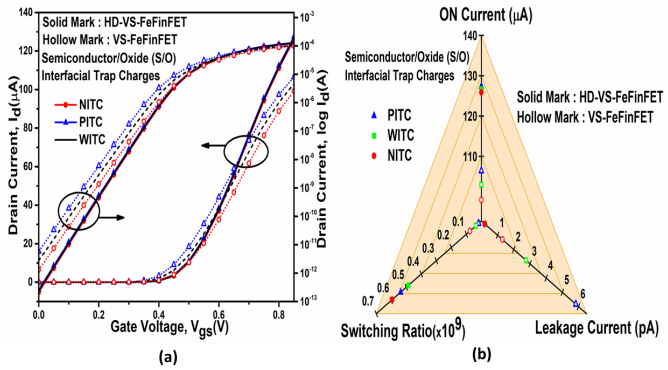



Fig. 3: Transfer characteristics of HD-VS-FeFinFET and VS-FeFinFET in (a) linear and log scale at $V_{ds}$=0.4V (b) Comparison of different device characteristics such as ON current ($I_{on}$), leakage current ($I_{off}$) and switching ratio ($I_{on}/I_{off}$) in presence of different ITCs at $V_{ds}$=0.4V

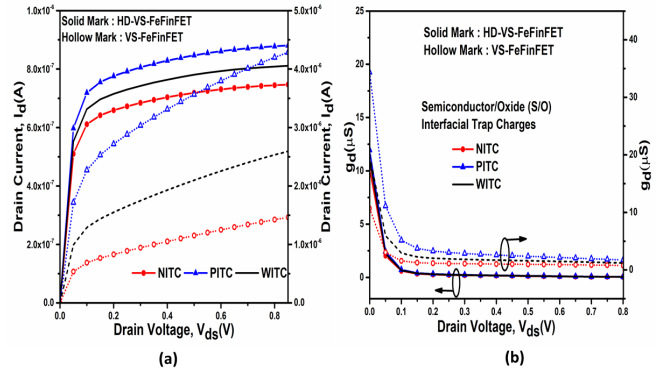



Fig. 4: Effect of ITCs on (a) drain current ($I_d$) and (b) output conductance ($g_d$) for HD-VS-FeFinFET and VS-FeFinFET at $V_{ds}$=0.4V

release charge carriers leading to energy dissipation[18][20]. This energy loss affects the overall Quality factor (QF) of the device as depicted in Fig. 6(b), as energy stored in the system is dissipated due to trap-assisted recombination and scattering mechanisms.

Fig. 7(a) shows the output resistance ($R_{out}$) for both the devices in presence of ITCs and higher $R_{out}$ ensures better signal fidelity by suppressing voltage variations for changes in load resistance as shown by HD-VS-FeFinFET. High-k dielectric materials help in minimizing the variation in

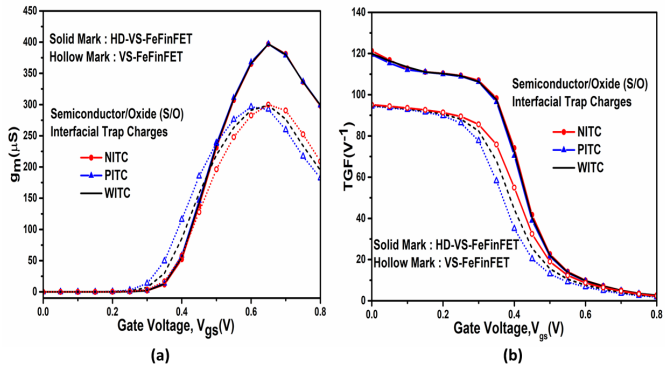



Fig. 5: Variation in (a) transconductance ($g_m$) and (b) device efficiency (TGF) in presence of different ITCs at $V_{ds}$=0.4V

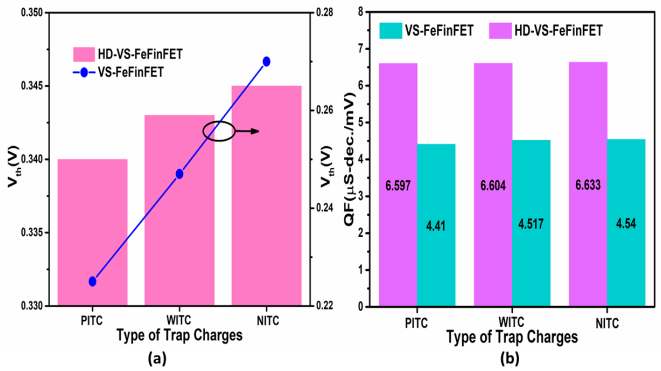



Fig. 6: Effect of ITCs on (a) threshold voltage ($V_{th}$) (b) quality factor (QF) of HD-VS-FeFinFET and VS-FeFinFET at $V_{ds}$=0.4V

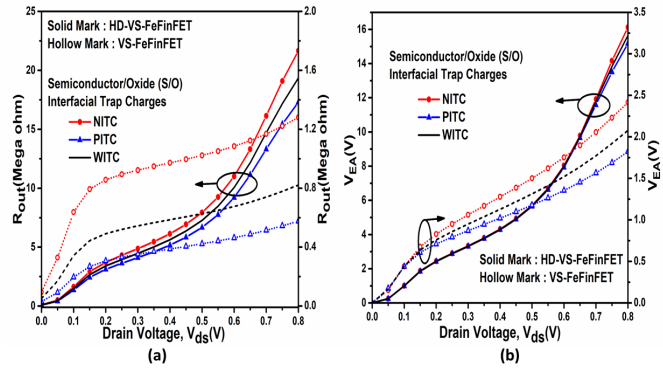



Fig. 7: Impact of ITCs on (a) output resistance ($R_{out}$) (b) early voltage ($V_{EA}$) of HD-VS-FeFinFET and VS-FeFinFET at $V_{ds}$=0.4V

$R_{out}$ caused by ITCs due to their ability to lower interface trap density and provide stronger electrostatic control over the channel[21][22]. The peak value of $R_{out}$ is increased (decreased) by 56% (29.9%) on introduction of NITC(PITC) at S/O interface of VS-FeFinFET as compared to increase (decrease) by 12% (10.6%) showing more immunity of HD-VS-FeFinFET towards ITCs.

Further, early voltage ($V_{EA}$) measures how quickly the output current increases in response to $V_{ds}$ and trap charges in

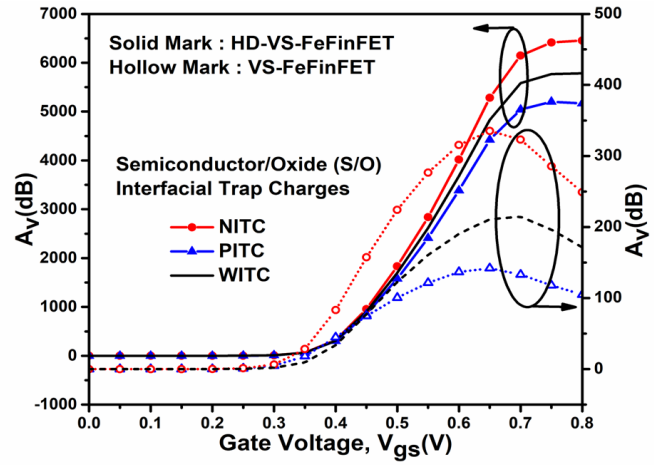



Fig. 8: Variation of intrinsic gain ($A_v$) for HD-VS-FeFinFET and VS-FeFinFET in presence of different ITCs at $V_{ds}$=0.4V.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:48:23 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Nanotechnology. 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/TNANO.2025.3531937



TABLE I: Summary of different Analog parameters for HD-VS-FeFinFET and VS-FeFinFET under the influence of ITCs

| Parameter | VS-FeFinFET | | | HD-VS-FeFinFET | | |
|---|---|---|---|---|---|---|
| | PITC | WITC | NITC | PITC | WITC | NITC |
| $I_{on}$ (A) | 1.06× $10^{-4}$ | 1.03× $10^{-4}$ | 9.92× $10^{-4}$ | 1.27× $10^{-4}$ | 1.26× $10^{-4}$ | 1.25× $10^{-4}$ |
| $I_{off}$ (A) | 5.83× $10^{-12}$ | 2.77× $10^{-12}$ | 1.30× $10^{-12}$ | 2.62× $10^{-13}$ | 2.36× $10^{-13}$ | 2.12× $10^{-13}$ |
| $I_{on}/I_{off}$ | 1.83× $10^{+7}$ | 3.72× $10^{+7}$ | 7.65× $10^{+7}$ | 5.37× $10^{+8}$ | 4.86× $10^{+8}$ | 5.94× $10^{+8}$ |
| $V_{th}(V)$ | 0.225 | 0.247 | 0.270 | 0.340 | 0.343 | 0.345 |
| QF ($\mu S - dec/mV$) | 4.41 | 4.52 | 4.54 | 6.60 | 6.60 | 6.63 |
| $g_m$ (S) | 2.92× $10^{-4}$ | 2.99× $10^{-4}$ | 3.0× $10^{-4}$ | 3.97× $10^{-4}$ | 3.97× $10^{-4}$ | 3.97× $10^{-4}$ |
| TGF ($V^{-1}$) | 94.43 | 94.76 | 95.10 | 119.46 | 119.91 | 121.23 |
| $A_v$ (dB) | 1.42× $10^{+2}$ | 2.11× $10^{+2}$ | 3.35× $10^{+2}$ | 5.16× $10^{+3}$ | 5.78 × $10^{+3}$ | 6.45× $10^{+3}$ |
| $V_{EA}(V)$ | 1.82 | 2.07 | 2.41 | 15.17 | 15.63 | 16.12 |
| $R_{out}$ (ohm) | 5.74× $10^{+5}$ | 8.19× $10^{+5}$ | 1.27× $10^{+6}$ | 1.72× $10^{+7}$ | 1.93× $10^{+7}$ | 2.16× $10^{+7}$ |

localized states within the semiconductor leading to deviations from ideal transistor behaviour. But owing to the implementation of heterogeneous dielectric in HD-VS-FeFinFET, there is only 3.09% (2.98%) increase (decrease) in presence of NITC (PITC) as compared to 12.3% (16.3%) increase(decrease) in its counterpart as shown in Fig. 7(b), making it more apt for high-speed analog and mixed-signal circuits[23]. Fig. 8 displays the comparison plots of intrinsic gain ($A_v$) for both the devices in presence of ITCs and it is found that with higher $A_v$ along with less variability due to ITCs, HD-VS-FeFinFET is more immune and reliable for analog performance as compared to VS-FeFinFET that shows huge deviations of 45.4% (38.8%) increment (decrement) for NITC(PITC).

### *B. Influence of ITCs on linearity, harmonic and intermodulation distortion performance*

Designing modern circuits and communication systems requires improved linearity and minimal distortion of signals to ascertain desired output at the receiver end. To take into account the non linearity issue, the influence of ITCs at S/O is investigated on various linearity and distortion parameters such as $g_{m2}$, $g_{m3}$, VIP2, VIP3, 1 dB compression point, IMD3, IIP3, etc for both the devices, HD-VS-FeFinFET and VS-FeFinFET. The importance of these higher-order derivatives lies in their direct influence on circuit gain, distortion, linearity, wireless performance, and power consumption. Designers must precisely analyze and optimize these parameters to meet the particular requirements of the application, thereby guaranteeing the circuit's reliability and overall success[24].

The higher-order coefficients of transconductance provide an estimate about the non-linearity in a device. $g_{m2}$ ($g_{m3}$) is the second (third) order transconductance coefficient that quantifies the second (third) order nonlinearity in a MOSFET, affecting even (odd) order distortion in circuits[25]. Fig. 9(a) displays that the incorporation of heterogeneous dielectric in HD-VS-FeFinFET results into negligible deviation in $g_{m2}$ and

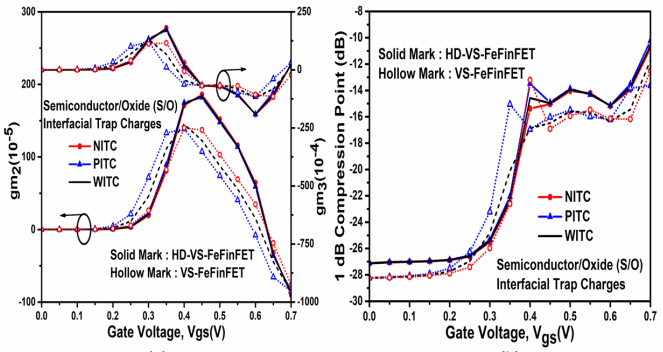



Fig. 9: Impact of ITCs on (a) higher order transconductance coefficient ($g_{m2}$ and $g_{m3}$) (b) 1-dB Compression Point of HD-VS-FeFinFET and VS-FeFinFET at $V_{ds}$=0.4V

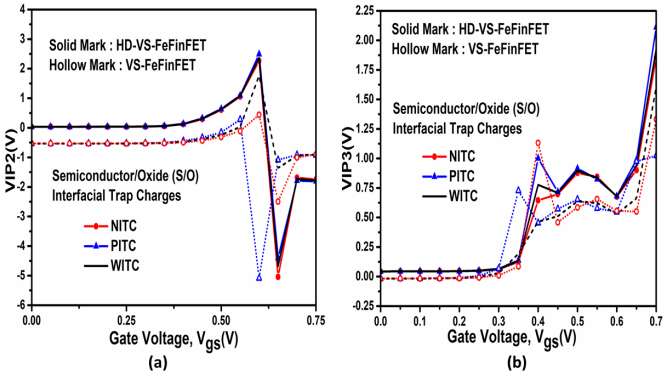



Fig. 10: Effect of ITCs on voltage intercept point (a) VIP2 (b) VIP3 of HD-VS-FeFinFET and VS-FeFinFET at $V_{ds}$=0.4V

$g_{m3}$ due to ITCs as compared to increase (decrease) in peak value in VS-FeFinFET in presence of NITC (PITC). Fig. 9(b) shows another important parameter, 1 dB compression point which denotes the level of input power at which shifting of output power from linearity by 1 dB occurs[26]. The figure displays higher value of 1 dB compression point for HD-VS-FeFinFET due to enhanced $g_m$ thus ensures reduced signal distortion. Also, there is negligible increase (decrease) of 7.57% (5.48%) in 1dB compression point on existence of PITC(NITC) in HD-VS-FeFinFET as compared to 27% (23.6%) variation in VS-FeFinFET for PITC(NITC).

Voltage intercept point (VIP) is an important parameter used

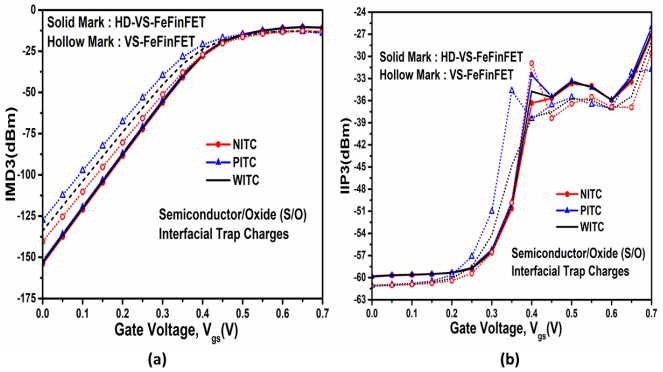



Fig. 11: Impact of ITCs on (a) third order intermodulation distortion power (IMD3) (b) third order intercept input power (IIP3) of HD-VS-FeFinFET and VS-FeFinFET at $V_{ds}$=0.4V

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:48:23 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Nanotechnology. 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/TNANO.2025.3531937



to describe the linearity of the device. VIP2 (VIP3) is the second (third) order voltage intercept point that depicts input voltage level at which the second (third) order harmonics equals the fundamental signal[23]. Higher VIP2 and VIP3 indicates better linearity and reduced distortion. Fig. 10(a) displays that along with higher value of VIP2, the influence of ITCs is negligible on HD-VS-FeFinFET due to hetero dielectric engineering. The presence of PITC(NITC) increases (decreases) the amplitude of VIP2 by 5.31% (4.14%) in case of HD-VS-FeFinFET. However, for VS-FeFinFET, the decrement in VIP2 is 57.3% (297%) in presence of NITC(PITC) in terms of WITC. Similar pattern is noticed in case of VIP3, its peak increases (decreases) only by 2.1%(2.03%) in presence of PITC(NITC) with respect to WITC whereas PITC(NITC) varies the VIP3 by 2.15% (7.57%) in VS-FeFinFET.

$$HD2 = \frac{1}{2} V_{\mathrm{a}} \frac{(\partial g_{\mathrm{m}}/\partial V_{\mathrm{gs}})}{2 g_{\mathrm{m}}} \tag{3}$$

$$HD3 = \frac{1}{4} {V_{\mathrm{a}}}^2 \frac{(\partial^2 g_{\mathrm{m}}/\partial {V_{\mathrm{gs}}}^2)}{6 g_{\mathrm{m}}} \tag{4}$$

$$THD = \sqrt{HD2^2 + HD3^2 + .....} \tag{5}$$

Third order intermodulation distortion power (IMD3) and third order intercept input power (IIP3) are other critical parameters of linearity in wireless applications. IMD3 refers to the distortion caused by third-order non linearities when two signals mix, producing undesired spectral components near the fundamental frequencies[25]. For better linearity, minimum value of IMD3 is desired as depicted by HD-VS-FeFinFET in Fig. 11(a). Further, the influence of ITCs on carrier mobility is less pronounced with high-k dielectrics which results in negligible influence on IMD3 in comparison to rise(fall) by 4.81% (4.89%) for PITC(NITC) in VS-FeFinFET. IIP3 is that extrapolated input power point where $1^{st}$ and $3^{rd}$ order harmonic power become equal. Higher IIP3 indicates better linearity[26]. Fig. 11(b) shows that IIP3 is nearly unaffected in presence of ITC for HD-VS-FeFinFET thus making it more reliable as compared to its counterpart. Furthermore, at low $V_{gs}$, IIP3 of HD-VS-FeFinFET is higher than the IMD3, which implies improved device power and performance by minimizing the hot carrier effect. In Fig. 11(b), at high $V_{gs}$, IIP3 of HD-VS-FeFinFET is lower than the IMD3, which is opposite to the trend at low $V_{gs}$. This occurs due to voltage-dependent non-linearities. Gate voltage significantly affects the channel's electrostatic potential, carrier mobility, and the non-linear terms in the device's transconductance. At low $V_{gs}$, the device operates in the subthreshold or near-threshold region, where second-order nonlinearities dominate due to weak inversion effects. These distortions influence IMD3 more strongly, making IIP3 appear relatively higher. On the other hand, at high $V_{gs}$, the device is in strong inversion, where third-order nonlinearities become more prominent. This increases the impact on IIP3, reducing its relative value compared to IMD3[26], [27].

To evaluate the intricacies of ITCs on the harmonics distortion characteristics, $2^{nd}$ order harmonics (HD2), $3^{rd}$ order harmonics (HD3), and total harmonic distortion (THD) of both

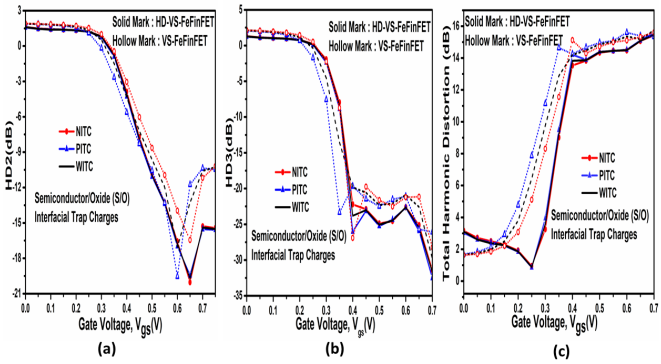



Fig. 12: Effect of ITCs on harmonics distortion (a) HD2 (b) HD3 (c) THD for HD-VS-FeFinFET and VS-FeFinFET at $V_{ds}$=0.4V

the devices are analyzed in presence of ITCs with different polarity as given by equation 3, 4, 5 respectively where amplitude of $V_a$ is taken as 50mV. HD2 (HD3) measures of the second (third) order nonlinearities producing signals at twice (thrice) the input frequency. And the measure of all harmonic distortion components relative to fundamental frequency is accounted as THD. The values of these parameters should be as low as possible for lesser distortion in device operation[27]. Fig. 12displays the impact of ITCs on HD2, HD3, and THD with change in $V_{gs}$ and it is found that HD-VS-FeFinFET shows negligible variations due to ITCs as compared to increment(decrement) in harmonic distortions with rise in $V_{gs}$ for VS-FeFinFET due to PITC(NITC). High-k materials tend to reduce the density of interface traps as compared to conventional dielectrics due to reduced dangling bonds at the interface. This reduction minimizes charge-trapping and de-trapping, leading to lower variations in harmonic distortions caused by PITC or NITC. Also, the high-k dielectric in HD-VS-FeFinFETs increases the gate capacitance, strengthening the gate's control over the channel potential. As a result, the channel becomes less sensitive to perturbations caused by ITCs. Thus, HD-VS-FeFinFET is proved to be more immune against any nonlinearity and distortion that arise from ITCs.

### C. Circuit level Analysis

To analyze the impact of ITCs of different polarity on HD-VS-FeFinFET based CMOS inverter, a required setup is designed on Cogenda Visual TCAD simulator with n-channel and p-channel HD-VS-FeFinFET isolated electrically with the help of 60nm $SiO_2$ spacer and later its performance has been compared with VS-FeFinFET based CMOS inverter. Fig 13(a) shows the schematic of the CMOS inverter circuit where $V_{dd}$, $V_{out}$ and $V_{in}$ are the supply voltage, output voltage, and input voltage respectively. Initially threshold matching of the n-channel and p-channel configuration is done for both the devices as displayed in Fig. 13(b) in presence of different ITCs. The $V_{gs}$ is varied from -0.85 to 0.85V with dual work function metal (DWFM) integration scheme.

It is very much required to examine the crucial parameters to estimate the performance of inverter and for this, voltage transfer characteristics (VTC) for both the inverters in presence of ITCs of different polarity is analyzed as shown

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:48:23 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Nanotechnology. 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/TNANO.2025.3531937



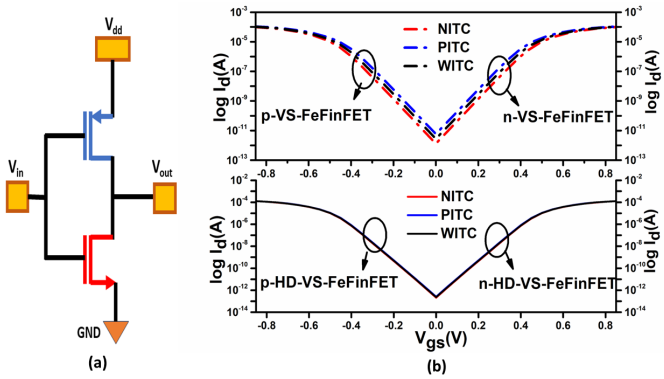



Fig. 13: (a) Schematic diagram of CMOS Inverter (b) Threshold Voltage Matching Curve for HD-VS-FeFinFET and VS-FeFinFET under the influence of ITCs

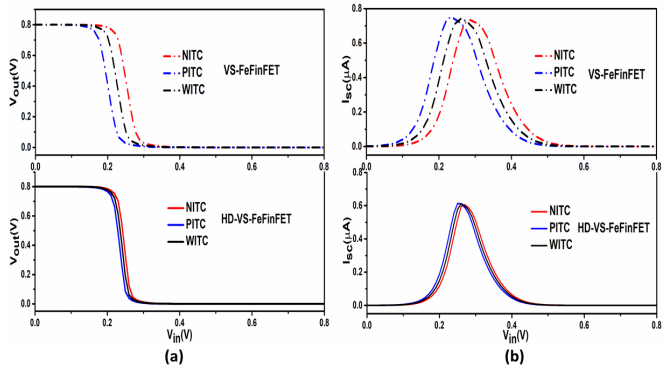



Fig. 14: Effect of ITCs on (a) Voltage Transfer Characteristics and (b) Switching Current of HD-VS-FeFinFET and VS-FeFinFET based inverter

in Fig. 14(a). It is visible that the transition range (TR) of HD-VS-FeFinFET based inverter is more sharper than the VS-FeFinFET based inverter due to improved electrostatic control and reduced SCEs, which ensures high speed switching[28], [29]. Further, the switching threshold voltage ($V_{TH}$) or the transition voltage, at which $V_{in}$=$V_{out}$ on the VTC curve, is examined and it is found to be 0.247V for HD-VS-FeFinFET with minimal variation due to ITC's polarity as compared to 0.238V for VS-FeFinFET in case of WITC.

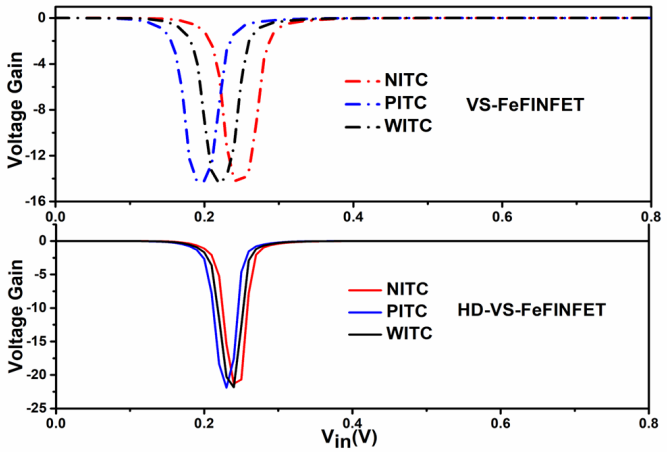



Fig. 15: Impact of ITCs polarity on Voltage Gain of HD-VS-FeFinFET and VS-FeFinFET based inverter.

TABLE II: Performance Comparison between HD-VS-FeFinFET and VS-FeFinFET based inverter under the influence of ITCs with different polarity

| Parameter | VS-FeFinFET | | | HD-VS-FeFinFET | | |
|---|---|---|---|---|---|---|
| | PITC | WITC | NITC | PITC | WITC | NITC |
| $V_{IL}$ (mV) | 143.473 | 169.841 | 196.792 | 183.994 | 191.341 | 198.031 |
| $V_{IH}$ (mV) | 270.794 | 285.762 | 297.713 | 267.452 | 274.142 | 280.174 |
| $V_{OL}$ (mV) | $5.77\times 10^{-8}$ | $1.39\times 10^{-7}$ | $3.41\times 10^{-7}$ | $3.60\times 10^{-8}$ | $4.58\times 10^{-8}$ | $3.35\times 10^{-8}$ |
| $V_{OH}$ (mV) | 799.99 | 800 | 800 | 800 | 800 | 800 |
| $NM_L$ (mV) | $1.43\times 10^{2}$ | $1.70\times 10^{2}$ | $1.97\times 10^{2}$ | $1.84\times 10^{2}$ | $1.91\times 10^{2}$ | $1.98\times 10^{2}$ |
| $NM_H$ (mV) | 529.196 | 514.238 | 502.287 | 532.548 | 525.858 | 519.826 |
| TR (mV) | 102.321 | 100.921 | 100.921 | 83.458 | 82.801 | 82.143 |

ITCs can generate additional spurious signals that can interfere with the proper functioning of the inverter. Noise Margin (NM) is a crucial parameter to depict the noise immunity of the circuit and its value should be higher for better tolerance to undesirable variations with more reliable operation. The noise margin for low signal levels ($NM_L$) and the noise margin for high signal levels ($NM_H$) is defined by equation 6 and 7. Here, $V_{OL}$ and $V_{OH}$ are minimum and maximum output voltage when corresponding output levels are logic 0 and logic 1 respectively. $V_{IL}$ and $V_{IH}$ are maximum and minimum input voltage which can be interpreted as logic 0 and logic 1 respectively[30]. It is found as shown in table II that with higher NM due to sharper switching characteristics along with negligible deviation under the influence of ITCs, HD-VS-FeFinFET is more reliable for digital circuits and applications.

$$NM_{\mathrm{L}} = V_{\mathrm{IL}} - V_{\mathrm{OL}} \tag{6}$$

$$NM_{\mathrm{H}} = V_{\mathrm{OH}} - V_{\mathrm{IH}} \tag{7}$$

Fig. 14(b) displays the switching current ($I_{sc}$) of both the logic circuits in presence of different ITCs. The static current (at $V_{in}$ = 0V, $V_{in}$ = $V_{dd}$) is less than 1 pA, confirming significantly lower static power dissipation and improved energy efficiency for HD-VS-FeFinFET along with negligible variations under the influence of ITCs with different polarity[28]. The effect of ITCs on inverter gain for both the devices is shown in Fig. 15. An increase of 51.674% in gain due to higher $g_m$ and $R_{out}$ along with minimum deviations due to ITCs is observed in HD-VS-FeFinFET based inverter as compared to VS-FeFinFET based inverter thus proved to be more desirable in various applications where signal fidelity and amplification are crucial.

## V. Conclusion

This paper presented a comprehensive assessment from device to circuit level showing the influence of ITCs on various analog, linearity, and distortion parameters of hetero dielectric gate engineered Si/SiGe strained vertically stacked

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:48:23 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. 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 Nanotechnology. 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/TNANO.2025.3531937



ferroelectric based FinFET (HD-VS-FeFinFET). The strained channel system and engineering modify the band structure and increase the carrier mobility thus improving the transistor's performance and making it more desirable for many electronic applications. The process improves the device performance in terms of $I_{on}$, $I_{off}$, $V_{th}$, $g_m$, VIP2, VIP3, IIP3, IMD3, 1-dB compression point, and harmonic distortion parameters. At the device level, HD-VS-FeFinFET outperforms VS-FeFinFET with improvement in various analog parameters like $A_v$ by 27 times, $V_{EA}$ by 7.5 times, and $R_{out}$ by 23 times along with minimal average variations of 11.15%, 3.03% and 11.39% respectively in contrast to 45.73%, 14.24% and 42.48% for VS-FeFinFET under the influence of ITCs. Further, circuit level comparison has also been made for CMOS inverter and it is found that HD-VS-FeFinFET based inverter offers improved performance in terms of noise margin, $I_{sc}$, and transition range along with more tolerance against ITCs. With the comprehensive analysis done on the intricacies of ITCs from device to circuit level, this work provides insights into the development of HD-VS-FeFinFET with enhanced functionality, reliability, and performance, poised to shape the landscape of modern electronics.

## ACKNOWLEDGMENT

The authors thank to Vinod Dham Centre of Excellence for Semiconductors and Microelectronics (VDCoE4SM), SERB POWER Grant, C2S Programme:MeiTY, and Microelectronics Research Lab at Delhi Technological University for providing essential facilities.

## REFERENCES

[1] N. Zagni, F. M. Puglisi, P. Pavan, and M. A. Alam, "Reliability of hfo2-based ferroelectric fets: A critical review of current and future challenges," *Proceedings of the IEEE*, 2023.

[2] W. M. Holt, "1.1 moore's law: A path going forward," in *2016 IEEE International Solid-State Circuits Conference (ISSCC)*. IEEE, 2016, pp. 8–13.

[3] A. Razavieh, P. Zeitzoff, and E. J. Nowak, "Challenges and limitations of cmos scaling for finfet and beyond architectures," *IEEE Transactions on Nanotechnology*, vol. 18, pp. 999–1004, 2019.

[4] K. Vanlalawmpuia and A. S. Medury, "Comparative analysis of ferroelectric quad-finfet with and without si3n4 spacer on analog/rf, linearity performance and digital inverter application with temperature variation," *Ferroelectrics*, vol. 613, no. 1, pp. 64–78, 2023.

[5] A. Dixit, D. P. Samajdar, N. Bagga, and D. S. Yadav, "Performance investigation of a novel gaas1-xsbx-on-insulator (gasoi) finfet: Role of interface trap charges and hetero dielectric," *Materials Today Communications*, vol. 26, p. 101964, 2021.

[6] K. Verma and R. Chaujar, "Analysis of vertically stacked ferroelectric based finfet for switching applications," in *2023 2nd International Conference on Futuristic Technologies (INCOFT)*. IEEE, 2023, pp. 1–4.

[7] T.-K. Chiang, "A compact model for threshold voltage of surrounding-gate mosfets with localized interface trapped charges," *IEEE Transactions on Electron Devices*, vol. 58, no. 2, pp. 567–571, 2011.

[8] S. Shabde, A. Bhattacharyya, R. S. Kao, and R. S. Muller, "Analysis of mosfet degradation due to hot-electron stress in terms of interface-state and fixed-charge generation," *Solid-state electronics*, vol. 31, no. 11, pp. 1603–1610, 1988.

[9] Y. Qiu, R. Wang, Q. Huang, and R. Huang, "A comparative study on the impacts of interface traps on tunneling fet and mosfet," *IEEE Transactions on Electron Devices*, vol. 61, no. 5, pp. 1284–1291, 2014.

[10] P. Zhao, A. Padovani, P. Bolshakov, A. Khosravi, L. Larcher, P. K. Hurley, C. L. Hinkle, R. M. Wallace, and C. D. Young, "Understanding the impact of annealing on interface and border traps in the cr/hfo2/al2o3/mos2 system," *ACS Applied Electronic Materials*, vol. 1, no. 8, pp. 1372–1377, 2019.

[11] Y. Qiu, R. Wang, Q. Huang, and R. Huang, "A comparative study on the impacts of interface traps on tunneling fet and mosfet," *IEEE Transactions on Electron Devices*, vol. 61, no. 5, pp. 1284–1291, 2014.

[12] S. Choudhary, D. Schwarz, H. S. Funk, R. Khosla, S. K. Sharma, and J. Schulze, "Impact of charge trapping on epitaxial p-ge-on-p-si and hfo 2 based al/hfo 2/p-ge-on-p-si/al structures using kelvin probe force microscopy and constant voltage stress," *IEEE Transactions on Nanotechnology*, vol. 20, pp. 346–355, 2021.

[13] C. Lin, B. Greene, S. Narasimha, J. Cai, A. Bryant, C. Radens, V. Narayanan, B. Linder, H. Ho, A. Aiyar, *et al.*, "High performance 14nm soi finfet cmos technology with 0.0174 $\mu$m 2 embedded dram and 15 levels of cu metallization," in *2014 IEEE International Electron Devices Meeting*. IEEE, 2014, pp. 3–8.

[14] Y. Pathak, B. D. Malhotra, and R. Chaujar, "Dft based atomic modeling and analog/rf analysis of ferroelectric hfo2 based improved fet device," *Physica Scripta*, vol. 98, no. 8, p. 085933, 2023.

[15] "Cogenda genius parallel 3d tcad simulation tool user's guide cogenda genius 2.2: User's guide."

[16] C. Auth, A. Aliyarukunju, M. Asoro, and B. et al., "A 10nm high performance and low-power cmos technology featuring 3rd generation finfet transistors, self-aligned quad patterning, contact over active gate and cobalt local interconnects," in *2017 IEEE International Electron Devices Meeting (IEDM)*, 2017, pp. 29.1.1–29.1.4.

[17] S.-N. Choi, S.-E. Moon, and S.-M. Yoon, "Impact of oxide gate electrode for ferroelectric field-effect transistors with metal-ferroelectric-metal-insulator-semiconductor gate stack using undoped hfo2 thin films prepared by atomic layer deposition," *Nanotechnology*, vol. 32, no. 8, p. 085709, 2020.

[18] B. Awadhiya, S. Yadav, and A. Acharya, "Interface trap charges analysis on dc and high frequency characteristics of utbb-fdsoi fet," *Silicon*, vol. 15, p. 937–942, 2023.

[19] M. H. Lin, C. C. Fan, H. H. Hsu, C. Liu, K. M. Chen, C. H. Cheng, and C. Y. Chang, "On the electrical characteristics of ferroelectric finfet using hafnium zirconium oxide with optimized gate stack," *ECS Journal of Solid State Science and Technology*, vol. 7, no. 11, p. P640, 2018.

[20] K. Verma and Chaujar, "Optimization and analysis of si/sige strained vertically stacked heterostructure on insulator fefinfet for high performance analog and rf applications," *Physica Scripta*, vol. 99, no. 11, p. 115960, 2024.

[21] V. R. Seshu, R. R. Shaik, and K. Pradhan, "Impact of unpreventable induced interface trapped charges on hzo based fdsoi ncfet," *Microelectronics Reliability*, vol. 139, p. 114748, 2022.

[22] E. Yurchuk, J. Müller, S. Müller, J. Paul, M. Pešić, R. van Bentum, U. Schroeder, and T. Mikolajick, "Charge-trapping phenomena in hfo2-based fefet-type nonvolatile memories," *IEEE Transactions on Electron Devices*, vol. 63, no. 9, pp. 3501–3507, 2016.

[23] Y. Pathak, K. Verma, B. D. Malhotra, and R. Chaujar, "Tcad analysis of linearity performance on modified ferroelectric layer in fet device with spacer," *Advanced Nanoscale MOSFET Architectures: Current Trends and Future Perspectives*, pp. 113–123, 2024.

[24] G.-B. Lee and Y.-K. Choi, "Extraction of interface trap density through synchronized optical charge pumping in gate-all-around mosfets," *IEEE Electron Device Letters*, vol. 41, no. 11, pp. 1629–1632, 2020.

[25] R. Das, S. Maity, and A. e. a. Chowdhury, "Effect of positive/negative interface trap charges on the performance of multi fin finfet (m-finfet)," *Silicon*, vol. 14, p. 8557–8566, 2022.

[26] R. Williams and D. Shaw, "Graded channel fet's: Improved linearity and noise figure," *IEEE Transactions on Electron Devices*, vol. 25, no. 6, pp. 600–605, 1978.

[27] A. Saxena, M. Kumar, R. Sharma, and R. Gupta, "Cylindrical soi schottky barrier mosfet with high linearity and low static power for digital and analog circuits application," *International Journal of High Speed Electronics and Systems*, vol. 30, no. 01n02, p. 2140003, 2021.

[28] A. Kumari N, B. Sreenivasulu V, and J. Singh, "Analog and mixed circuit analysis of nanosheet fet at elevated temperatures," *Physica Scripta*, vol. 98, no. 10, p. 105409, 2023.

[29] O. Prakash, A. Gupta, G. Pahwa, J. Henkel, Y. S. Chauhan, and H. Amrouch, "Impact of interface traps on negative capacitance transistor: Device and circuit reliability," *IEEE Journal of the Electron Devices Society*, vol. 8, pp. 1193–1201, 2020.

[30] V. B. Sreenivasulu, A. K. Neelam, A. K. Panigrahy, L. Vakkalakula, J. Singh, and S. G. Singh, "Benchmarking of multi-bridge-channel fets toward analog and mixed-mode circuit applications," *IEEE Access*, vol. 12, pp. 7531–7539, 2024.

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on January 27,2025 at 06:48:23 UTC from IEEE Xplore. Restrictions apply.
© 2025 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/377857473

# Validation of a Parameter Design methodology for Air-core two-coil WPT system with multi-resonant LC compensation

**Conference Paper** · December 2023

DOI: 10.1109/PESGRE58662.2023.10404443

CITATIONS
5

READS
12

**3 authors**, including:



Gaurav Yadav
Delhi Technological University
**14** PUBLICATIONS **40** CITATIONS
SEE PROFILE



Mukhtiar Singh
Delhi Technological University
**77** PUBLICATIONS **2,178** CITATIONS
SEE PROFILE

All content following this page was uploaded by Gaurav Yadav on 16 November 2024.

The user has requested enhancement of the downloaded file.

# Validation of a Parameter Design methodology for Air-core two-coil WPT system with multi-resonant LC compensation

Akshay
Electrical Engineering Department
Delhi Technological University
Delhi,India
akshayjanagal@gmail.com

Gaurav Yadav
IEEE Graduate Student Member
Delhi Technological University
Delhi,India
gaurav812786@gmail.com

Mukhtiar Singh
IEEE Senior Member
Delhi Technological University
Delhi,India
smukhtiar_79@yahoo.co.in

2023 IEEE International Conference on Power Electronics, Smart Grid, and Renewable Energy (PESGRE) | 979-8-3503-1057-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/PESGRE58662.2023.10404443

***Abstract*—Wireless Power Transfer (WPT) possesses the revolutionary potential to replace traditional wired power transmission methods. WPT significantly enhances the usability of Electric Vehicle (EV) charging applications by facilitating the removal of obstacles within charging operations. This study demonstrates a wireless adapter that has been optimized for electric vehicle charging. This article offers a solution for optimizing the compensation network, with a particular emphasis on coil parameter specifications, its modelling, mutual inductance calculation air gap, coupling factor, and compensated basic multi-resonant LC compensation resonant network parameters calculations to reduce output voltage sensitivity and charging the EV. The proposed converter makes use of the energy from the AC grid to correct power factor while wirelessly charging the EV battery. The entire charging procedure consists of AC-DC conversion, DC-HFAC inversion, wireless power transfer, and HFAC-DC rectification. A prototype charger is constructed for a multi-resonant topology, and the MATLAB/Simulink model is used to evaluate the performance of the optimized resonant network.**

***Keywords—EV Charging, Wireless Power Transfer (WPT), HFAC, Coupling Factor, Mutual Inductance, Compensation.***

## I. INTRODUCTION

When it comes to reducing global greenhouse gas (GHG) emissions, the transportation industry is a key offender; however, battery electric vehicles (BEVs) provide a promising answer by drastically lowering carbon emissions using renewable energy sources. [10][11]. By doing away with the need for wires,(WPT) technology facilitates this shift by revolutionising power transmission for electric vehicle charging applications and enabling fully autonomous, hands-free recharging with the proper alignment of transmitting (Tx) and receiving (Rx) coils [12],[13].

An integral part of every WPT system, high-frequency inverters improve the quality factor and power transfer efficiency [1]. Increased switching losses at higher operating frequencies are a problem for traditional hard-switched and pulse-width modulated inverters when used for WPT applications. Coil quality attributes, alignment, and coupling coefficient all have a major impact on efficiency and must be taken into account in an optimised design [4], [7].

Multi-resonant LC compensation inductive WPT systems are used in the suggested solution, together with active rectifiers on the mains side and converters on the load side, to improve the overall efficiency of the system [2], [5]. Precise control systems at the roadside ensure the best possible waveform output and automate adjustments at the vehicle end, resulting in increased efficiency [3].

The study's overarching goal is to learn more about how WPT technology is progressing of creating a future in which electric vehicle (EV) charging infrastructures are more eco-friendly and sustainable[15][16]. By delving into the nuts and bolts of WPT for EV applications, we may get a better idea of its potential effects on both the transportation industry and the natural world [6],[8], [9].

## II. PROPOSED TECHNOLOGY

The wireless electric vehicle (EV) charger consists of four crucial components for efficient wireless power transmission: an AC-DC converter, a High-Frequency AC Inverter (HFAC), Wireless Power Transmission (WPT) via Inductive Power Transfer (IPT), and a Full Bridge Rectifier with Capacitive Filter as shown in Fig (1). The system transforms grid AC power into a refined DC output for an EV battery.

The utilization of Active Front End (AFE) LCL filters serves to mitigate the presence of harmonics and noise when receiving a 415V $V_{PP}$ input power supply from a three-phase Grid. The 6-switch Bridge rectifier is utilized to transform the Grid voltage into a pulsating DC voltage of 800V. This DC voltage is subsequently stabilized by the presence of the DC-link capacitor. The control algorithm employed in the SRFT Control theory enables the voltage control loops of the AFE converter to modulate the output voltage within the ranges of $V_{PP}$ as (700V to 900V or 800V to 600V), commonly referred to as maintaining the DC Link Voltage.

The HFAC inverter functions at a frequency of 30 kHz and employs an open-loop control system. Despite the lack of feedback, this inverter consistently produces a stable and constant output. The WPT technique known as Inductive Power Transfer (IPT) involves the utilization of magnetic induction to transfer energy from a primary coil to a secondary coil. In this procedure, the initial 800V square wave alternating current (AC) voltage undergoes conversion into an (800V) $V_{PP}$ high frequency sinusoidal AC voltage. This resultant high frequency AC voltage subsequently generates a (1000V) $V_{PP}$ high frequency AC voltage.

The utilization of the Full Bridge Rectifier with Capacitive Filter is employed for the purpose of converting AC power to DC power. This conversion process results in the generation of a stable output voltage that is suitable for recharging the 48V lead-acid battery commonly found in EVs. The current paper is structured as follows:

Section II examines a multi-resonant wireless power transfer (WPT) charger with a circular spiral coil with perfect alignment. LCR meter results in Section III validate the coil model. Section V result reveals validate and support the coupling factor. Section VI concludes this study by summarizing towards control of AFE which ensures UPF.

979-8-3503-1057-3/23/$31.00 ©2023 IEEE

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 16,2024 at 05:17:55 UTC from IEEE Xplore. Restrictions apply.

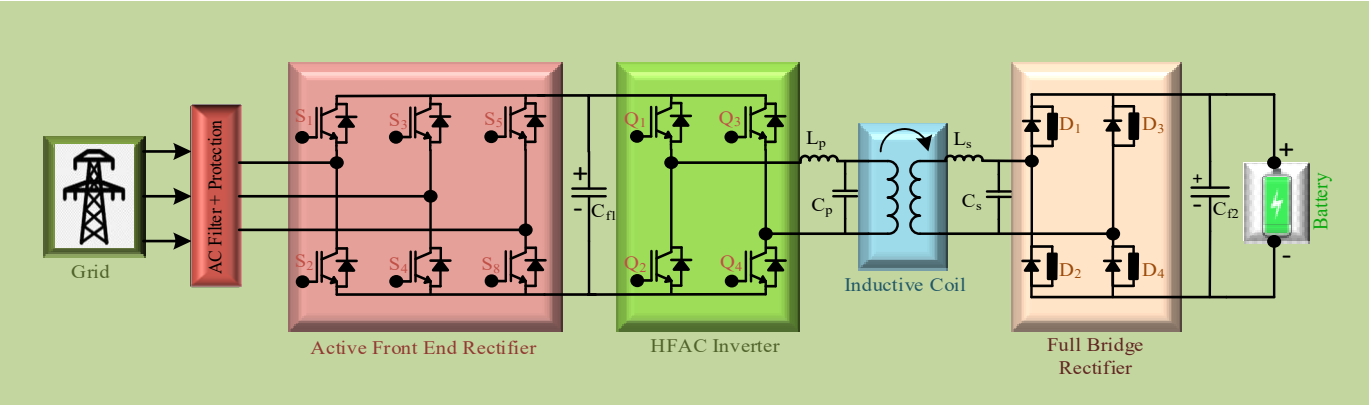



Fig.1. Charging topology for electric vehicles, showing primary and secondary charging stations

## III. MODELLING

The vast majority of WPT designs employ planar coils. There are two distinct possibilities: circular and rectangular coils. The parameters of the circular spiral coils used in this study are depicted in the figure below. The LCR meter can be used to measure the self-inductance, resistance, and diameter of a circular spiral coil with and turns, respectively by connecting the terminals to the LCR meter as shown in Fig (2).

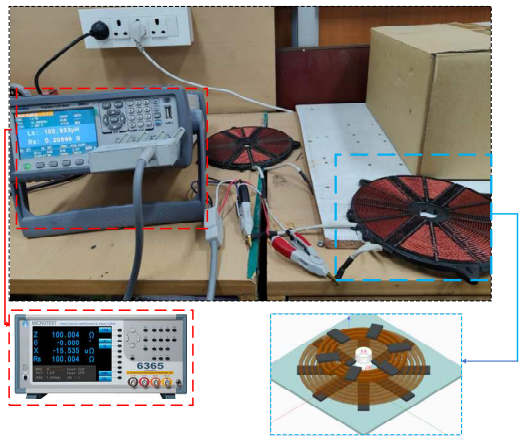


Fig.2 Measurement Setup by LCR Metre SM6023

***TABLE.1.-COIL PARAMETERS***

| COIL PARAMETERS | VALUE |
|---|---|
| Roadside Winding Turns | $N_1 = 28$ |
| Vehicle side Winding Turns | $N_2 = 28$ |
| Diameter of Roadside Coil | $D_1 = 19cm$ |
| Diameter of Vehicle side Coil | $D_2 = 19cm$ |
| Self-Inductance of Roadside Coil | $L_1 = 155.243\mu H$ |
| Self-Inductance of Vehicle side Coil | $L_2 = 110.313\mu H$ |
| Resistance of Roadside Coil | $R_1 = 0.23213\Omega$ |
| Resistance of Vehicle side Coil | $R_2 = 0.20755\Omega$ |
| Operating frequency | $f_s = 30kHz$ |

Table (1) presents an overview of the coil parameters, including the number of turns in the roadside coil, its diameter, self-inductance, resistance, as well as the number of turns in the vehicle side coil, its diameter, self-inductance, resistance, and the operating frequency also shown in Fig (3).

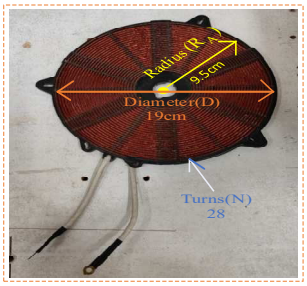



Fig.3 Front View and Specification of coil

Mono resonant compensation corrects system resonant with one frequency. Multi-resonant compensation systems with reactive components worked well. Our simulation tweaks basic LC topology. This work examines series-series compensation as shown in Fig (4) and LC topological parameters as shown in Fig (5). We study the development of these parameters by considering the roadside ($V_1$) and vehicle side ($V_2$) voltages and currents ($I_1$ and $I_2$).

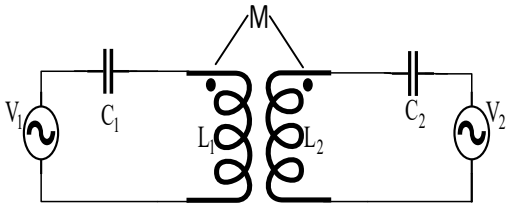



Fig.4 Series-Series mono-resonant Compensation

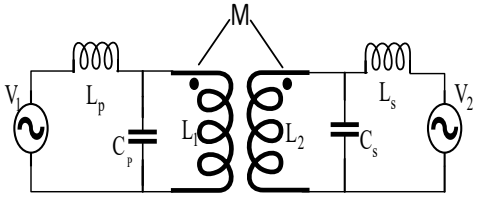



Fig.5 LC based multi-resonant Compensation

In order to determine parameters, a mesh-based analysis is first performed, after which the input impedance is inferred, and only components that are entirely resistive require their reactive component terms equal to zero.

$$V_1 = ((R_2 + j\omega L_1 + \frac{1}{j\omega C_1})I_1 - j\omega M I_2 \qquad (1)$$

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 16,2024 at 05:17:55 UTC from IEEE Xplore. Restrictions apply.

$$V_2 = ((R_2 + j\omega L_2)I_2 - j\omega M I_1 = (R_L + \frac{1}{j\omega C_2})I_2 \quad (2)$$

From Eq(1),(2), The $Z_{in}$ and ratio between $I_1$ and $I_2$ as follows:

$$\frac{I_2}{I_1} = \frac{j\omega M}{j\omega L_2 + R_2 + R_L + \frac{1}{j\omega C_2}} \quad (3)$$

$$Z_{in} = R_1 + j(\omega L_1 - \frac{1}{\omega C_1}) + \frac{\omega^2 M^2}{j\omega L_2 + R_2 + R_L - \frac{j}{\omega C_2}} \quad (4)$$

As a result, the series-series adjusted parameters are able to be determined using Equation(5)

$$\omega_0 = \frac{1}{\sqrt{L_1 C_1}} = \frac{1}{\sqrt{L_2 C_2}} \quad (5)$$

Multi-resonant compensation parameters as the process:

$$L_p = 1mH \quad (6)$$

$$C_p = 0.18\mu F \quad (7)$$

$$L_s = 1mH \quad (8)$$

$$C_s = 0.255\mu F \quad (9)$$

Parameters on both the vehicle and road sides of compensation are determined and shown by Equation (6), (7),(8)and(9)

### *A. MUTUAL INDUCTANCE CALCULATION*

Consider two circular spiral coils aligned in a way as shown in the Fig (6) with radius $R_p$ and $R_s$ and the distance between them from origin is h. As for analysis we are considering points A and B as its geometrical view is shown in Fig (7).

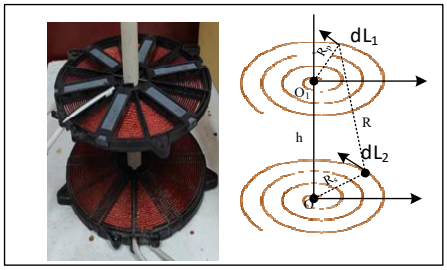



Fig.6 Planar spiral coil with perfect alignment

According to Neuman's formula, the mutual inductance of a planar spiral coil calculated using the magnetic vector potential method is as follows:

$$M = \frac{\mu_0}{4\pi} \oint_{L_1} \oint_{L_2} \frac{dL_1 \times dL_2}{R} \quad (10)$$

where $\mu_0$ is the air permeability, $dL_1$ and $dL_2$ are line elements and R is the spacing between them.

Also, Line tangential elements at an arbitrary point A, B using the cylindrical coordinates system can be described as:

$$dL_1 = -R_p \sin\theta_1 d\theta_1 \hat{x} + R_P \cos\theta_1 d\theta_2 \hat{y} \quad (11)$$

$$dL_2 = -R_s \sin\theta_1 d\theta_1 \hat{x} + R_s \cos\theta_1 d\theta_2 \hat{y} \quad (12)$$

where $R_p$ and $R_s$ represents the radius of primary and Secondary coil with respect to origin

$$dL_1 \cdot dL_2 = R_p R_B \cos(\theta_2 - \theta_1) d\theta_1 d\theta_2 \quad (13)$$

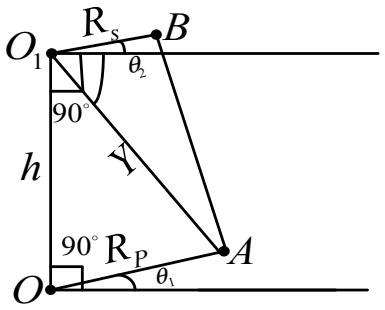



Fig.7 Geometry view for analysis

Using the cosine rule in triangle $OO_1A$, we can determine that Y is the distance to point $AO_1$

$$Y = h^2 + R_p{}^2 - 2hR_P \sin\theta_1 \quad (14)$$

By the triangle geometry rules

$$\angle AO_1B = \theta_2 - \theta_1 \quad (15)$$

Once more, we may use the cosine rule in triangle $O_1BA$ to get the distance R between $dL_1$ and $dL_2$

$$R^2 = Y^2 + R_s{}^2 - 2YR_S \cos(\theta_2 - \theta_1) \quad (16)$$

$$R^2 = h^2 + R_p{}^2 - 2hR_P \sin\theta_1 + R_s{}^2 - 2R_S \cos(\theta_2 - \theta_1)\sqrt{h^2 + R_p{}^2 - 2hR_P \sin\theta_1} \quad (17)$$

As $\theta_1$ and $\theta_2$ is very small ,so $\sin(\theta_1)$ and $\sin(\theta_2)$ is very small and neglected but cosine quantities at small angle is comparable ,Hence the above equation after this approximation becomes as shown in Eq(18)

$$R^2 = h^2 + R_p{}^2 + R_s{}^2 - 2R_sR_p \cos(\theta_2 - \theta_1) \quad (18)$$

Substituting Equation (13) in Eq(10),the mutual inductance of perfectly aligned spiral coil is obtained in:

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 16,2024 at 05:17:55 UTC from IEEE Xplore. Restrictions apply.

$$M = \frac{\mu_0}{4\pi} \iint_{2\pi N} \frac{R_p R_B \cos(\theta_2 - \theta_1) d\theta_1 d\theta_2}{R} \quad (19)$$

Substituting Equation (17) in Eq (19), the mutual inductance of perfectly aligned spiral coil is obtained in:

$$M = \frac{\mu_0}{4\pi} \int \int_{2\pi N} \frac{R_p R_B \cos(\theta_2 - \theta_1) d\theta_1 d\theta_2}{\sqrt{h^2 + R_p^{\,2} - 2hR_p \sin\theta_1 + R_s^{\,2} - 2R_S \cos(\theta_2 - \theta_1)} \sqrt{h^2 + R_p^{\,2} - 2hR_p \sin\theta_1}} \quad (20)$$

Substituting our coil parameter in Eq (20) as shown in Table (1) value of Mutual Inductance is obtained as shown in Eq (21).

$$M_{accurate} = 39.33 \mu H \quad (21)$$

Substituting our coil parameters in Eq (19) as shown in Table (1), the value of Mutual Inductance is obtained as shown in Eq (22).as mentioned in [14]

$$M = \frac{\mu_0}{4\pi} \iint_{2\pi N} \frac{R_p R_B \cos(\theta_2 - \theta_1) d\theta_1 d\theta_2}{\sqrt{h^2 + R_p^{\,2} + R_s^{\,2} - 2R_s R_p \cos(\theta_2 - \theta_1)}} \quad (22)$$

Evaluating double Integral for the above parameters and coil turns of the roadside and vehicle side and its calculation leads to the Eq (23).

$$M_{approx} = 40.323 \mu H \quad (23)$$

The evaluation of the double integral for the aforementioned parameters and its subsequent calculation results in Equation (24).

$$M = 39.33 \mu H \quad (24)$$

## IV. VERIFICATION

### A. MUTUAL INDUCTANCE MEASUREMENT METHOD

This mutual inductance measurement is performed using a Scientific Precision LCR Meter SM6023, and the results are compared to the calculated values. To measure mutual inductance, the following method is employed. First, the inductance is measured when the primary and secondary coils are connected in forwarding mode, as depicted in Fig (8), and then when they are connected in backward mode, as depicted in Fig (9) Their one-fourth subtraction results in the calculation of mutual inductance.

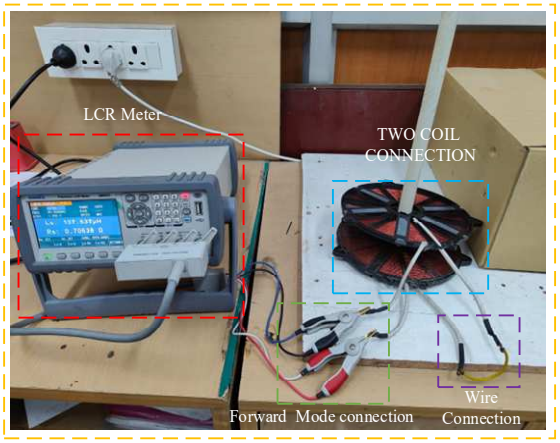



Fig.8 Forward mode Inductance test

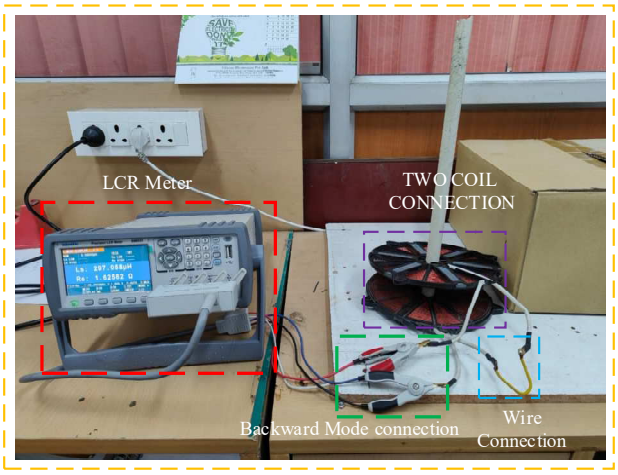



Fig.9 Backward Mode Inductance Test

The actual mutual inductance is derived by subtracting the forward connection inductance from backward connection inductance and dividing it by 4 leads to the equation as given in (20)

$$M = \left| \frac{L_f - L_b}{4} \right| \quad (25)$$

The forward mode Inductance Test, depicted in Figure 1, is accompanied by its corresponding readings as presented in Equation (26).

$$L_f - L_P + L_s + 2M = 137.637 \mu H \quad (26)$$

The Backward mode Inductance Test as shown in Fig.1.and its readings are shown in Equation (27).

$$L_b - L_P + L_s - 2M = 297.055 \mu H \quad (27)$$

Hence, utilizing this measurement technique, the value of mutual inductance as determined by Equation (29) closely approximates the value indicated in Equation (24).

$$M = \frac{297.055 - 137.637}{4} \quad (28)$$

$$M = 39.8545 \mu H \quad (29)$$

### B. COUPLING FACTOR CALCULATION

The verification of the coupling coefficient, which is an essential aspect of evaluating mutual inductance, plays a crucial role in ensuring the efficiency and safety of wireless power transfer systems. The optimization of coil designs, alignment, and positioning for quality power transfer can be achieved through the identification of anomalies in the coupling coefficient.

$$K = \frac{M}{\sqrt{L_1 L_2}} \quad (30)$$

Substituting Equation (24) and $L_1, L_2$ through Table (1)

$$K = 0.2861 \quad (31)$$

The coupling factor, as determined by Equation (31), falls within an acceptable range, as demonstrated.in [5]

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 16,2024 at 05:17:55 UTC from IEEE Xplore. Restrictions apply.

## V. RESULTS AND DISCUSSION

The Active Front End (AFE) system uses a 415(V)VPP three-phase AC power supply from a grid or generator. LCL filters minimise noise and harmonics. A 6-switch Bridge rectifier generates pulsing DC voltage from AC. IGBT bridges power this rectifier. The DC-link capacitor smooths 800V DC voltage. Voltage reference variations affect the AFE converter's voltage control loop. To maintain output voltage, this control loop minimises the difference between real and reference voltages. Figure 10 shows a reference voltage of 700V to 900V or 800V to 600V.

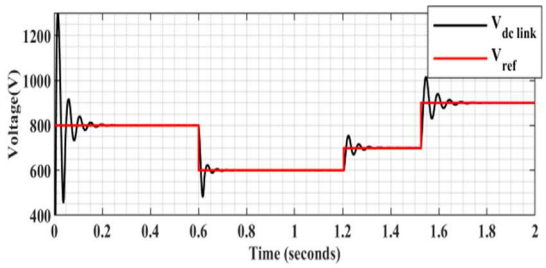



Fig.10 Output of AFE ( $V_{DC}$ ) at different Voltage reference

A 30kHz, 50% duty cycle high-frequency alternating current (HFAC) inverter with an open-loop control system switch. The 800V input voltage ( $V_{DC}$ ) inverter has reduced size and weight and faster response times. Due to the lack of input, an open-loop control system may have dynamic performance limits, although it can still be successful.

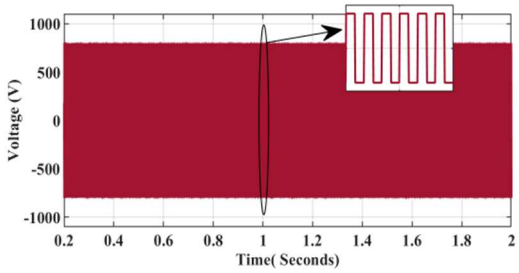



Fig.11 Output voltage of HFAC Inverter

Inductive Power Transfer (IPT), operates through the utilization of magnetic induction. An LC circuit is utilized to mitigate the effects of an (800V) $V_{pp}$ HFAC square wave as Fig (11) by generating a sinusoidal AC voltage. The power is received by the primary coil and subsequently after compensation as Fig (13). transferred to the secondary coil, resulting in the generation of a high frequency voltage of (1000V) $V_{pp}$ as Fig (14). Full Bridge Rectifier with Capacitive Filter power electrical circuit converts AC power to DC power. The rectifier's output is sent via a capacitive filter to charge a 48V lead-acid battery with a steady, high-quality DC output voltage.

The wireless EV charging topology shows a continuous sinusoidal pattern with a 50Hz frequency, indicating a stable and steady source of electricity. To improve Power Factor (PF) and reduce harmonic distortion, system monitoring is needed to reduce reactive power consumption in the current waveform. The AFE's well-regulated DC link voltage waveform reflects load or input grid supply changes as. as Fig (12) The battery voltage waveform shows rising and then stopping after it reaches 48Vas shown in Fig (12).

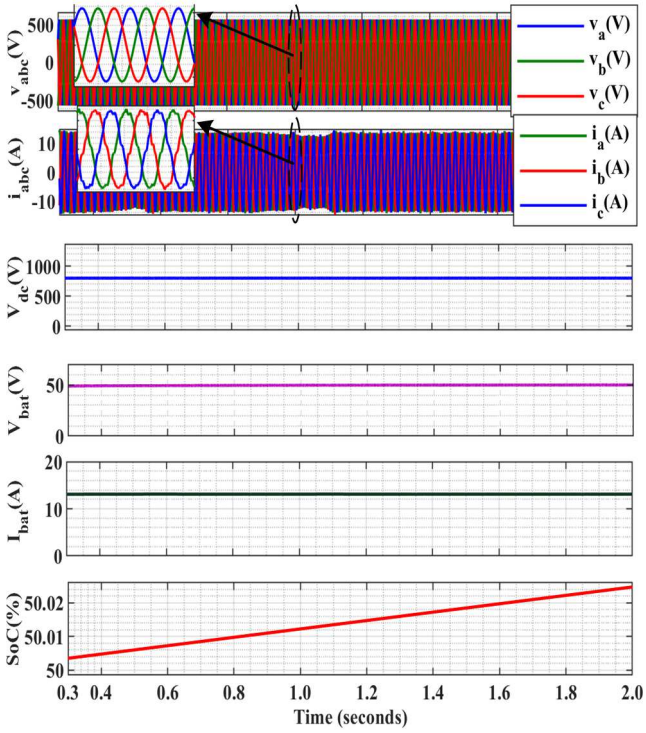



Fig.12 (a) Grid voltage(top), (b)grid current, (c) $V_{dc\ link}$ (AFE), (d) battery voltage, (e)battery current, (f) State of charge of battery(bottom)

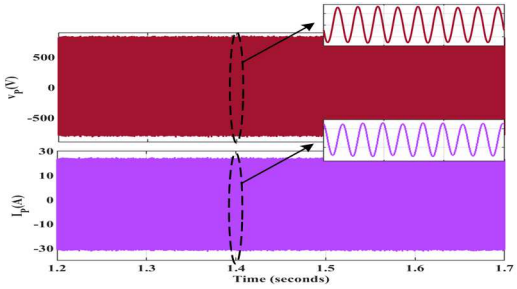



Fig.13 Voltage and current on the primary coil

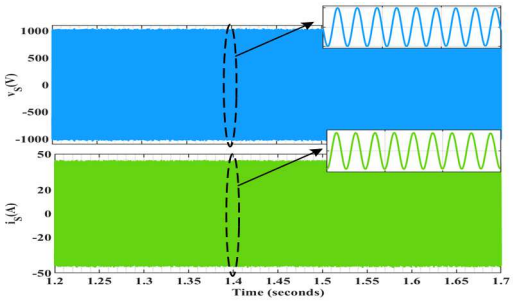



Fig.14 Voltage and current on the secondary coil

## VI. CONCLUSION

In conclusion, the analysis of an LC compensated (WPT) charger, implementing thorough calculations of mutual inductance and assuming a perfect alignment with a mutual inductance value of M = 39.33, has exhibited considerable promise for charging batteries operating at a standard voltage of 48 volts, while meeting the requirements of conventional

Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 16,2024 at 05:17:55 UTC from IEEE Xplore. Restrictions apply.

battery specifications. By employing a grid-side voltage with a peak-to-peak value $V_{pp}$ of 415 volts, the battery is effectively charged by the system, eliminating the need for a DC-DC converter on the load side. The charger is equipped with precise LC compensation, featuring primary and secondary inductance values of $L_p = 1mH$ and $L_s = 1mH$, respectively. Additionally, it incorporates capacitance values of $C_p = 0.18\mu F$ and $C_s = 0.25\mu F$, which effectively facilitate efficient power transfer while minimizing energy losses. Moreover, the utilization of an Active Front End (AFE) in the capacity of a grid-side AC-DC converter serves to accomplish genuine unity power factor correction (UPF), effectively aligning voltage and current waveforms.

## APPENDIX

### Control of AFE

*Voltage control*

$T_{C1} = 220\,ms$

$C_1 = 100\,mF$

$K_P = \dfrac{3 * C_1}{T_{C1}} = 1.5$

$K_I = \dfrac{3 * (20 * 10^{-3})}{T_{C1}} = 300$

I. *Current control*

$T_{C1} = 220ms$

$C_1 = 100mF$

$K_P = \dfrac{3 * C_1}{T_{C1}} = 1.5$

$K_I = \dfrac{3 * (20 * 10^{-3})}{T_{C1}} = 300$

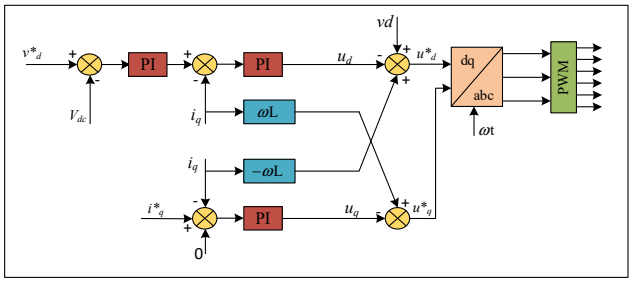



Fig.16. Voltage and Current loop Controlled Scheme

## REFERENCES


[1] A. A. S. Mohamed, D. Allen, T. Youssef and O. Mohammed, "Optimal design of high frequency H-bridge inverter for wireless power transfer systems in EV applications," *2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC)*, Florence, Italy, 2016, pp. 1-6, doi: 10.1109/EEEIC.2016.7555646.

[2] S. Li, F. Li, R. Zhang, C. Tao and L. Wang, "Accurate Modeling, Design and Load Estimation of LCC-S based WPT system with a wide range of load," in *IEEE Transactions on Power Electronics*, doi: 10.1109/TPEL.2023.3279659.

[3] F. P. Wijaya, T. Shimotsu, T. Saito and K. Kondo, "A Simple Active Power Control for a High-Power Wireless Power Transmission System Considering Coil Misalignment and Its Design Method," in *IEEE Transactions on Power Electronics*, vol. 33, no. 11, pp. 9989-10002, Nov. 2018, doi: 10.1109/TPEL.2018.2799679.

[4] K. Fotopoulou and B. W. Flynn, "Wireless Power Transfer in Loosely Coupled Links: Coil Misalignment Model," in *IEEE Transactions on Magnetics*, vol. 47, no. 2, pp. 416-430, Feb. 2011, doi: 10.1109/TMAG.2010.2093534.

[5] A. Ramezani and M. Narimani, "Optimized Electric Vehicle Wireless Chargers with Reduced Output Voltage Sensitivity to Misalignment," in *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 8, no. 4, pp. 3569-3581, Dec. 2020, doi: 10.1109/JESTPE.2019.2958932.

[6] .N. Fu, J. Deng, Z. Wang, W. Wang and S. Wang, "A Hybrid Mode Control Strategy for LCC–LCC- Compensated WPT System With Wide ZVS Operation," in *IEEE Transactions on Power Electronics*, vol. 37, no. 2, pp. 2449-2460, Feb. 2022, doi: 10.1109/TPEL.2021.3108637

[7] Y. Li, J. Zhao, Q. Yang, L. Liu, J. Ma and X. Zhang, "A Novel Coil with High Misalignment Tolerance for Wireless Power Transfer," in *IEEE Transactions on Magnetics*, vol. 55, no. 6, pp. 1-4, June 2019, Art no. 2800904, doi: 10.1109/TMAG.2019.2904086

[8] L. Li, H. Liu, H. Zhang and W. Xue, "Efficient Wireless Power Transfer System Integrating with Metasurface for Biological Applications," in *IEEE Transactions on Industrial Electronics*, vol. 65, no. 4, pp. 3230-3239, April 2018, doi: 10.1109/TIE.2017.2756580.

[9] Y. Zhang et al., "Integration of Onboard Charger and Wireless Charging System For Electric Vehicles With Shared Coupler, Compensation, and Rectifier," in *IEEE Transactions on Industrial Electronics*, vol. 70, no. 7, pp. 7511-7514, July 2023, doi: 10.1109/TIE.2022.3204857.

[10] . A. A. S. Mohamed, C. R. Lashway and O. Mohammed, "Modeling and Feasibility Analysis of Quasi-Dynamic WPT System for EV Applications," in *IEEE Transactions on Transportation Electrification*, vol. 3, no. 2, pp. 343-353, June 2017, doi: 10.1109/TTE.2017.2682111.

[11] Y. Jiang, L. Wang, Y. Wang, J. Liu, M. Wu and G. Ning, "Analysis, Design, and Implementation of WPT System for EV's Battery Charging Based on Optimal Operation Frequency Range," *in IEEE Transactions on Power Electronics*, vol. 34, no. 7, pp. 6890-6905, July 2019, doi: 10.1109/TPEL.2018.2873222.

[12] S. R. Meher, Y. Choudhary and R. K. Singh, "An Optimal Wireless Battery Charger for Electric Vehicle using EF2 Inverter at 6.78 MHz," 2022 *IEEE Energy Conversion Congress and Exposition (ECCE),* Detroit, MI, USA, 2022, pp. 1-5, doi: 10.1109/ECCE50734.2022.9947370.

[13] A. Mishra and M. Singh, "A 10-kW Active Front End Rectifier Fed Dual Active Bridge Converter For EV Charging," 2023 *4th International Conference for Emerging Technology (INCET),* Belgaum, India, 2023, pp. 1-6, doi: 10.1109/INCET57972.2023.10170584.

[14] Hussain, I.; Woo, D.-K. Simplified Mutual Inductance Calculation of Planar Spiral Coil for Wireless Power Applications. *Sensors* 2022, *22*, 1537. https://doi.org/10.3390/s22041537.

[15] M. Ansari, G. Yadav and M. Singh, "A 3.3kW Modified LLC Resonant Converter for Grid-Tied EV System Under Wide Voltage Range," 2023 3rd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2023, pp. 1-6, doi: 10.1109/CONIT59222.2023.10205690.

[16] M. Ansari, G. Yadav and M. Singh, "A 3.3kW Modified LLC Resonant Converter for Grid-Tied EV System Under Wide Voltage Range," 2023 3rd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2023, pp. 1-6, doi: 10.1109/CONIT59222.2023.10205690.


Authorized licensed use limited to: DELHI TECHNICAL UNIV. Downloaded on November 16,2024 at 05:17:55 UTC from IEEE Xplore. Restrictions apply.

View publication stats

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/388269468

# VSG-DC-Based Grid Forming Inverter Control for Standalone Microgrid

**Conference Paper** · January 2025

CITATIONS
0

READS
12

**3 authors**, including:



Deepak Kumar
Delhi Technological University
**11** PUBLICATIONS **33** CITATIONS

SEE PROFILE

All content following this page was uploaded by Deepak Kumar on 22 January 2025.

The user has requested enhancement of the downloaded file.

# VSG-DC-Based Grid Forming Inverter Control for Standalone Microgrid

Deepak Kumar
Department of Electrical Engineering
Delhi Technological University
New Delhi, India
Deepakkumar_23phdee503@dtu.ac.in

Jitendra Nath Rai
Department of Electrical Engineering
Delhi Technological University
New Delhi, India
jnrai@dce.ac.in

ST Nagarajan
Department of Electrical Engineering
Delhi Technological University
New Delhi, India
stnagarajan@dce.ac.in

***Abstract*— The modern power system integrated with inverter-based resources (IBRs), such as solar and wind utilizes complex control strategies to preserve grid stability. This paper introduces an improved control structure of a grid-forming inverter (GFMI) for a standalone (SA) microgrid system. The control strategy VSG-DC, emulates the inertial and damping characteristics and provides more virtual inertia to reduce the rate of change of frequency (RoCoF). A common cascade current and voltage controller is used to maintain and stabilize the voltage and current levels, by which improving the dynamic frequency and voltage response. simulation performed in MATLAB/Simulink, the proposed control structure maintains the voltage and frequency according to grid standards during transient or specific load changes. the result shows that the control structure is successful in improving microgrid stability.**

***Keywords*— *Islanded microgrid, Grid forming inverter (GFMI), Rate of change of frequency (RoCoF), VSG- control.***

## I. INTRODUCTION

The present power system is based on renewable energy sources (RES). RES plays an important role in power system as to the replacement of fossil-fuel-based synchronous generators and helps to reduce emissions by the increasing of IBRs, ensuring the grid stability and reliability [1],[2]. in this regard, IBR control technologies are required to allow the smooth integration of RES, which in turn has a major effect on grid stability.

Microgrids have become a platform for integrated IBRs, which can operate in both GC and IL mode [2]. In GC mode, IBRs function as GFLI because the grid controls the voltage and frequency. thus, GFLIs are widely used IBRs, but for the RES-dominated grids, the GFL control has some issues that can be resolved by using GFM controller to give grid support potential. GFMIs provide the auxiliary service to the grid like as synchronous generator [3]. because of their dominating act in a weak grid and the ability to perform in the islanded mode [4].

The two types of control methodologies are operated for GFM controllers: droop-based control and synchronous machine-based control [5]. droop control is simply a proportional controller. in GFMIs, a droop controller is used in both APCs and RPCs. In HVT networks, $P-\omega$ and $Q-V$ droop controller is used if the grid impedance (GI) is predominantly inductive in nature [6]. Due to the absence of virtual inertia in the droop controller, RoCoF in islanded mode after the load disturbance becomes high [7]. hence, synchronous machine modeling control strategies such as virtual synchronous machines are designed to produce low inertia. the most popular methodology to simulate synchronous machine inertia (SMI) and damping quantities (DQ) is the swing equation-based VSG control technique [8].

In many literatures, the VSG control methodologies are enhanced to meet different control goals. In [9], the authors proposed a PI controller-based APC that can monitor power correctly even for the duration of grid frequency changes recognized due to the PI-based active power loop controller's integral action. in [10], the authors represent the control approach that combines PSC and VSG controller design and also suggest order reduction by using a pole-zero cancellation-based method to find out the suitable inertia constant. In [11], an intelligent power oscillation damper (IPOD)is described and added with the conventional VSG, which is used to damp out the power oscillations from the electrotechnical inter-area. In [12], a synchronous power controller (SPC) is recommended to provide inertia, damping characteristics, and droop characteristics. In [13], suggests a second-order controller known as generalized droop control (GDC) to adhere to RoCoF limitation in the SA mode and provide step feedback with small exceeds and a brief resolve duration in the GC mode. for improving the damping, a cross-feed forward compensation-based topology is suggested in [14] for the generators with coupled active power and reactive power loops. in [15], a phase-amplitude cross-coupling method is proposed to improve damping with the help of APC and RPC. In [16], an advanced control technique that provides a power compensation mechanism to compensate for significant power grid voltage dips, and secure synchronization stability under difficult conditions.

this paper represents the VSG-DC controller, which has an APC and RPC loop controller that emulates the synchronous generator swing equation to elevate the inverter response in SA mode. in APC, the two parameters, virtual inertia constant(J) and damping coefficient (Dp), both play an effective role in impacting the VSG frequency response [20]. a common cascade voltage and current controller is used to maintain and stabilize the voltage and current levels, by improving the dynamic frequency and voltage response. The main contribution of this paper

1) a VSG-DC-based controller is proposed to enhance the performance of the controller in standalone (SA) mode.
2) RoCoF, or frequency stability, is a key indicator of the frequency transient behavior.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

## II. GFM INVERTER FOR VSG-DC CONTROL

### A. Modeling of VSG-DC-based GFMI

The control structure of GFMI is shown in Figure 1, a micro-grid consisting of a voltage source DC/AC inverter, cascade voltage and current control system (outer control), APC and RPC control system (inner control), and loads in standalone (SA) mode are investigated. The main focus of this paper is to enhance the stability by using a modified active power control loop proposed for GFMI. The voltage angle (θ) is controlled by the APC. GFMI was operated on standalone conditions with resistive load L1 and L2. Load L2 is connected with the microgrid by switch SW1, which is closed at t = 0.2s.

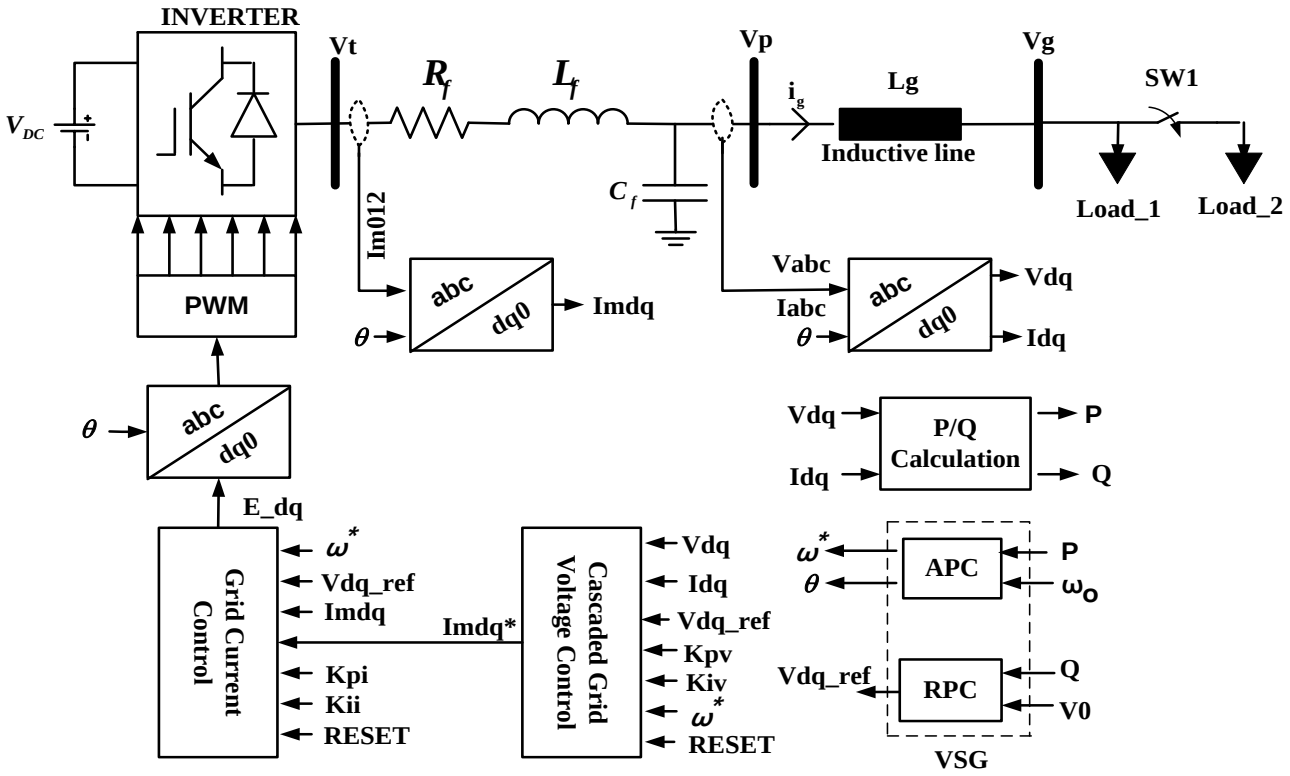



Fig. 1. Control Structure of GFMI

According to Figure 1, $R_f$, $L_f$, *and* $C_f$ represent the filter parameters, $L_g$ represents the grid line inductance and two loads L1, L2 connected to the inverter. L2 load connected with the help of switch SW1, which operates at the time t=0.2s. *Vabc* and *Iabc* represent the Voltage and current at PCC and $I_{m012}$ represents the inverter side current. with the help of equations (1) and (2), finds the AP and RP at PCC, which is fed into the VSG-DC controller, the VSG-DC provides the reference phase angle θ and reference voltage ($V_{dq_ref}$) for a common cascade voltage and current controller (CV-CC), thus CV-CC generates the voltage reference $E_{dq}$, both θ and $E_{dq}$ are essential input values for the PWM to operate the system inverter.

The AP and the RP of the inverter are defined as equations (1) and (2)[4].

$$P = \frac{3}{2}\left(V_{pd} I_{gd} + V_{pq} I_{gq}\right) \tag{1}$$

$$Q = \frac{3}{2}\left(V_{pq} I_{gd} - V_{pd} I_{gq}\right) \tag{2}$$

AP and RP both depend on, phase variation between the line voltage and variation in the peak voltage.

### B. VSG-based droop control for APC and RPC

By integrating electrical energy storage devices to simulate the inertia of a conventional synchronous generator, a VSG refines and improves durability in transient situations. The swing equation, which simulates the inertia of a traditional generator, acts as the starting point for a VSG's control approach. Using swing equation to merge with conventional droop control improves the transient performance of the system. Although LPF is used to slow down control dynamics and satisfy the maximum RoCoF criteria, HPF is employed to obtain nominal grid frequency in steady-state.

The VSG-droop control implementing a swing equation can be defined in equation (3)[7]:

$$(P_{ref} - P) - \frac{1}{m_p} d\omega = J\omega_0 \frac{d\omega}{dt} \tag{3}$$

$$m_p = \frac{\Delta\omega}{P_{\max}} = \frac{2\pi(f_{\max} - f_{\min})}{2 * P_{rated}} \tag{4}$$

The damping coefficient can be expressed as: the inverse of the static droop coefficient in equation (5) [4]

$$D_p = \frac{1}{m_p} = \frac{2 * P_{rated}}{2\pi(f_{\max} - f_{\min})} \tag{5}$$

$$J = \frac{H_J}{m_p \cdot \omega_0} \tag{6}$$

Where $\omega_o$ is the nominal grid frequency, $m_p$ frequency static droop and $J$ denotes the virtual inertia coefficient defined by equations (4), (6) [7], $\omega^*$ is the reference frequency. VSG-based-droop controller, including the swing equation. figure

2 shows the control block diagram of VSG-DC for APC and RPC, where θ is the reference phase angle.

$G_{VSG}(s)$ represents the transfer function of VSG, expressed as in equation (7).

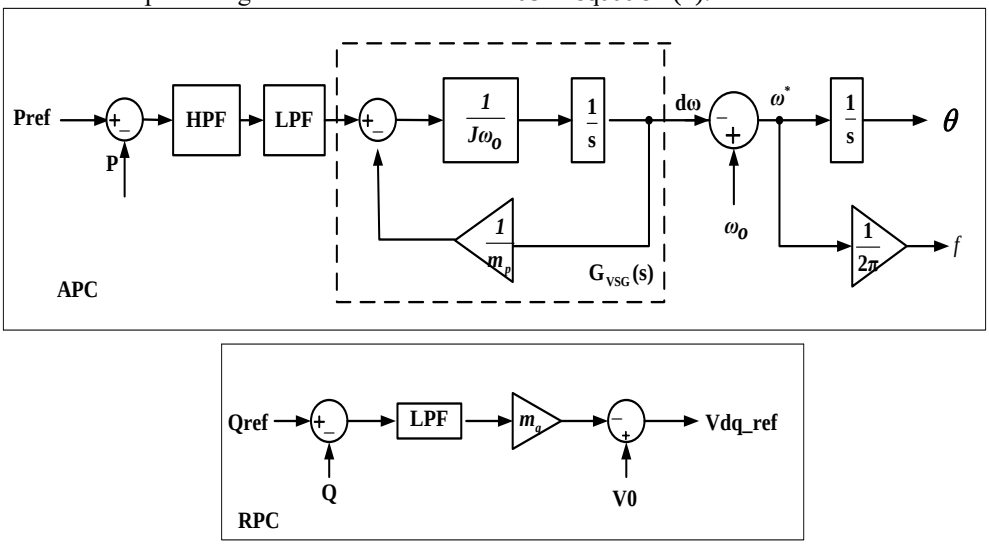



Fig. 2. Control block diagram VSG for droop control for APC and RPC

$$G_{VSG}(s) = \frac{m_p}{sH_J + 1} \quad (7)$$

Where $H_J = m_p J \omega_0$, denotes the inertial time constant, which allows for acting on the resulting RoCoF. Therefore, to ensure that the maximum RoCoF requirement is observed for a maximum active power step as follows, $H_j$ can be designed as follows.

$$H_J = \frac{m_p}{\rho} P_{\max} \quad (8)$$

where ρ is the maximum allowable RoCoF and $P_{max}$ is the maximum active power.

The reactive power control with a low pass filter, generates reference $V_{dq,}$ which can be expressed in equation (9).

$$V_{dq_ref} = -(Q_{ref} - Q)\mathrm{m}_q + V_0 \quad (9)$$

$$m_q = \frac{\Delta V}{Q_{\max}} \quad (10)$$

Where $m_q$ denotes the static droop coefficient, $V_0$ nominal grid RMS voltage. Δω and ΔV represent the maximum allowable frequency and voltage deviations, with $f_{max}$, $f_{min}$, $V_{max}$, and $V_{min}$ specifying the associated maximum and minimum allowable frequency and voltage deviation values in accordance with grid standards.

### *C. Current control (CC) loop*

The holdup impact of the converter pulse-width modulated (PWM) operation must be allowed while adjusting the settings of the inner loop CC in Figure 3. Assuming that the measured voltages are feed-forward and the d- and q-axes are perfectly decoupled, the open-loop transfer function (TF) for the CC can be obtained by Imprecise the result of the PWM using a first-order TF. Equation (11) defines the TF of CC [17-18] and the parameter value shown in the table1.

$$E_{cc,dq}(s) = G_{cc,pic}(s).G_{pwm}(s).G_f(s) \quad (11)$$

where, $\mathrm{G}_{cc,pic}(s) = K_{pi} + \frac{K_{ii}}{s}$ } PI current controller

$\mathrm{G}_{pwm}(s) = \frac{1}{1+T_{di}s}$ } PWM approximation

$\mathrm{G}_f(s) = \frac{1}{R_f + sL_f}$ } Filter inductor

$K_{pi} = \frac{L_f}{2T_{di}}, K_{ii} = \frac{R_f}{2T_{di}}, \ \mathrm{T}_{di}$ } *Delay approximation*

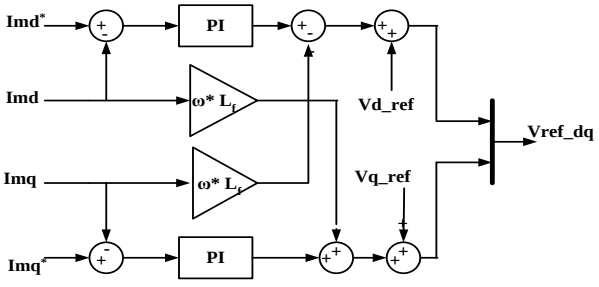



Fig. 3. Structure of current control loop

### D. Voltage control loop

The voltage control loop is designed to keep the inverter's output voltage within safe limits for the protection of system devices. Figure 4 shows that proportional-integral (PI) controllers separately regulate the voltages of the d-q axis. in order to mitigate the integral coupling between these axes, a decoupling term is comprised of the PI controllers' output that minimizes cross-axis interactions.

In order to achieve stable operation and efficient disturbance rejection, the Symmetrical Optimum (SO) approach is used to adjust the voltage controller's settings. design also allows current loop delays $T_{dq}$, to make sure that the voltage controller runs more slowly than the current controller to follow accurate voltage and $T_{di}$ defines the cumulative effect of all minor delays in the system, which is multiplied by a

factor of 10 to make certain adequate compensation in order to preserve stability and avoid overshooting.

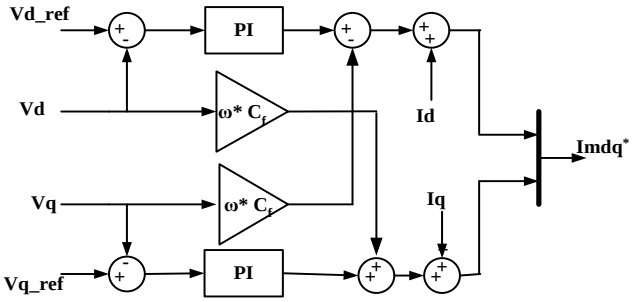



Fig. 4. Structure of Voltage control loop

The controller parameters are defined in equation (12) [17-18] as follows:

$$E_{vc,\mathrm{dq}}(s) = G_{vc,pic}(s).G_{fc}(s) \qquad (12)$$

where, $G_{vc,pic}(s) = K_{pv} + \frac{K_{iv}}{s}$ } PI voltage controller

$G_{fc}(s) = \frac{1}{sC_f}$ } Filter capacitor

$K_{pv} = C_f / (a * T_{deq})$ and $T_i = a^2 T_{deq}$

$K_{iv} = K_{pv} / T_i$ $T_{deq} = 10T_{di}$

$T_{di} = T_{ctrl} + T_{PWM} + T_{avg} = 1.5T_s,\ T_s = \frac{1}{f_s}$

TABLE 1. PARAMETER USED FOR THE SYSTEM

| S. N | Parameter | Value |
|---|---|---|
| 1 | LC Filter, Grid line inductance | $L_f$ = 950mH, $R_f$ = 54mΩ, $C_f$ = 12.5μF and $L_g$ = 2.2mH |
| 2 | Inverter rating, grid voltage | $V_{dc}$ = 800V, $S_{max}$ = 22KV, $V_0$ = 230V(RMS), $f_{sw}$ = 50KHz |
| 3 | APC & RPC | $J$ = 9.6255 Kg-m$^2$, $m_p$ = 1.6535x10$^{-4}$, $D_p$ = 6.050x10$^3$, $m_q$ = 0.0029, $\omega_0$ = 314.15 *rad/s*, $\rho$ =*1Hz/s* |
| 4 | Voltage and current controller | $K_{pv}$ = 0.0215, $K_{iv}$ = 17.916, $K_{pi}$ = 15.833, $K_{ii}$ = 900.00 |
| 5 | Loads | L1= 4.5Ω, L2=7.5Ω |

## III. Simulation Result and Discussion

Simulations are performed in MATLAB/Simulink to analyze the response of GFMI, using the parameters in Table 1. Figure 5 shows the change in frequency; the maximum RoCoF is set as 1 Hz/s. With the use of HPF and LPF in the APC loop system, there is no effect on frequency fluctuation when distortion occurs in active power. Both the RoCoF and the lowest frequency during the switching transient, recorded at 0.82 Hz/s and 49.954 Hz respectively, remain within the specified limits.

Additionally, the frequency reverts to its nominal state in approximately 6 seconds. For a given maximum RoCoF, the value of the corresponding virtual inertia $J$ is 9.6255 Kg.m$^2$. this result shows how the virtual inertia directly influences the RoCoF and, thus, how the inertia can be tuned accordingly to meet the specific RoCoF requirement within the imposed limitations. For this system, Δω and ΔV were taken as 1% and ±10% respectively.

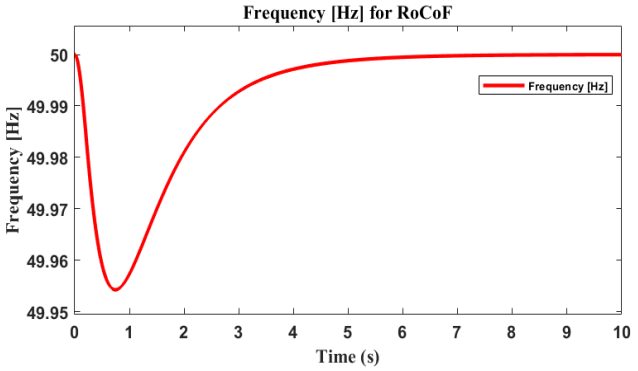



Fig. 5. APC Change in frequency

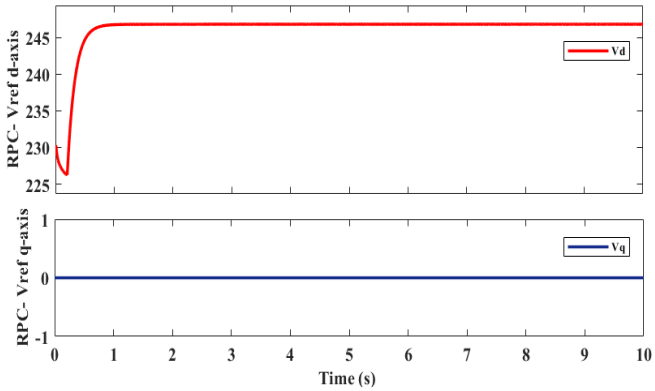



Fig. 6. RPC voltage reference for d-q-axis

by observing the d-q axis voltage and analyzing the voltage regulation capability of RPC. figure 6 shows the 5.5% voltage variation within the allowable variation of 10%, which shows the system voltage stability.

The simulation results show that the APC and RPC work in tandem to manage both AP and RP effectively. the values of load L1 is 4.5Ω and load L2 is 7.5Ω, load L2 connected with the microgrid by switch SW1 which is switched at time 0.2s. The inverter is connected to the DC supply, and the rating of the inverter is 22KVA, as shown in Figure 1. At t=0s to t=0.2s load L1 is connected, the AP and RP delivered to the load is 10KW and 5.7KVAR, and at t=0.2s the load L2 is connected. Thus, the maximum allowable AP and RP delivered to the Loads are 19KW and 11.0KVAR, respectively, which is displayed in Figure 7 (c), (d).

Figure 7(a) shows the phase voltage and angle variation when connecting the load L2 at t=0.2s, while Fig. 7(b) shows the rise of the grid current caused by the inherent rise of power.

The overall performance of the system confirms that the control structure is adequate in enhancing the stability of SA microgrids. It handles significant load changes without causing undesirable power oscillations or exceeding RoCoF constraints, which is vital for the reliable effort of future microgrids dominated by renewable energy sources.

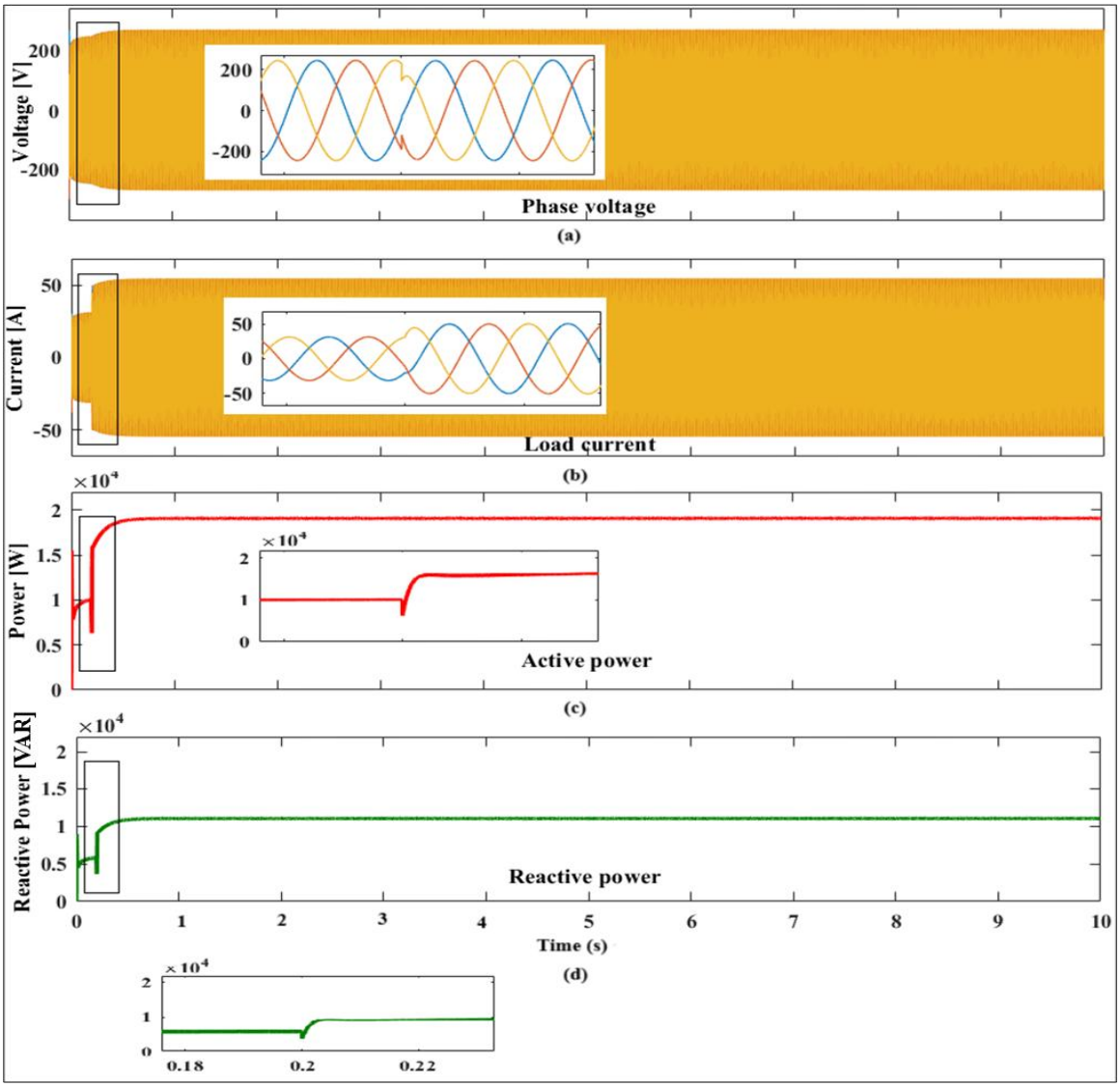



Fig. 7. Simulation result of GFMI (a) microgrid phase voltage (b) load current (c) output Active power (d) output reactive power

## IV. Conclusion

This paper introduced the VSG-DC control structure of GFMI, to maintain grid stability and minimize power oscillation when sudden load changes, The use of HPC and LPC in the APC enables smooth frequency recovery without exceeding RoCoF constraints. The results validate the model's effectiveness, showcasing that the control system is capable of handling varying load conditions while ensuring that both changes in frequency and voltage deviations remain within grid standards. This research contributes to advancing microgrid stability by providing a scalable solution for integrating inverter-based RES in standalone microgrids.

## References


[1] J. Matevo syan et al., "A future with inverter-based resources: Finding strength from traditional weakness," *IEEE Power Energy M ag.*, vol. 19, no. 6, pp. 18–28, Nov./Dec. 2021.

[2] M. S. Mastoi, S. Zhuang, M. Haris, M. Hassan, and A. Ali, "Large-scale wind power grid integration challenges and their solution: A detailed review," *Environ. Sci. Pollut. Res.*, vol. 30, no. 47, pp. 103424–103462, Oct. 2023.

[3] D. Sharma, F. Sadeque, and B. Mirafzal, "Synchronization of inverters in grid forming mode," *IEEE Access*, vol. 10, pp. 41341–41351, 2022.

[4] D. B. Rathnayake, R. Razzaghi and B. Bahrani, "Generalized Virtual Synchronous Generator Control Design for Renewable Power Systems," in *IEEE Transactions on Sustainable Energy*, vol. 13, no. 2, pp. 1021-1036, April 2022.

[5] R. H. Lasseter, Z. Chen and D. Pattabiraman, "Grid-Forming Inverters: A Critical Asset for the Power Grid," in *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 8, no. 2, pp. 925-935, June 2020.

[6] R. Rosso, X. Wang, M. Liserre, X. Lu and S. Engelken, "Grid-Forming Converters: Control Approaches, Grid-Synchronization, and Future Trends—A Review," in *IEEE Open Journal of Industry Applications*, vol. 2, pp. 93-109, 2021.

[7] N. Mohammed, H. Udawatte, W. Zhou, D. J. Hill and B. Bahrani, "Grid-Forming Inverters: A Comparative Study of Different Control Strategies in Frequency and Time Domains," in *IEEE Open Journal of the Industrial Electronics Society*, vol. 5, pp. 185-214, 2024

[8] X. Quan, A. Q. Huang, and H. Yu, “A novel order reduced synchronous power control for grid-forming inverters,” *IEEE Trans. Ind. Electron.*, vol. 67, no. 12, pp. 10 989–10 995, Dec. 2020.

[9] W. Zhang, A. Tarraso, J. Rocabert, A. Luna, J. I. Candela, and P. Rodriguez, “Frequency support properties of the synchronous power control for grid-connected converters,” *IEEE Trans. Ind. Appl*., vol. 55, no. 5, pp. 5178–5189, Sep. 2019.

[10] X. Quan, A. Q. Huang, and H. Yu, “A novel order reduced synchronous power control for grid-forming inverters,” *IEEE Trans. Ind. Electron*., vol. 67, no. 12, pp. 10 989–10 995, Dec. 2020.

[11] G. N. Baltas, N. B. Lai, L. Marin, A. Tarraso, and P. Rodriguez, “Grid forming power converters tuned through artificial intelligence to damp sub synchronous interactions in electrical grids,” *IEEE Access*, vol. 8, pp. 93369–93379, 2020

[12] W. Zhang, A. M. Cantarellas, J. Rocabert, A. Luna, and P. Rodriguez, “Synchronous power controller with flexible droop characteristics for renewable power generation systems,” *IEEE Trans. Sustain. Energy*, vol. 7, no. 4, pp. 1572–1582, Oct. 2016.

[13] X. Meng, J. Liu, and Z. Liu, “A generalized droop control for grid supporting inverter based on comparison between traditional droop control and virtual synchronous generator control,” *IEEE Trans. Power Electron*., vol. 34, no. 6, pp. 5416–5438, Jun. 2019.

[14] X. Li, Y. Hu, Y. Shao, and G. Chen, “Mechanism analysis and suppression strategies of power oscillation for virtual synchronous generator,” in Proc. IECON - 43rd Annu. Conf. IEEE Ind. Electron. Soc., 2017, pp. 4955–4960

[15] M. Ebrahimi, S. A. Khajehoddin, and M. Karimi-Ghartemani, “An improved damping method for virtual synchronous machines,” *IEEE Trans. Sustain. Energy*, vol. 10, no. 3, pp. 1491–1500, Jul. 2019.

[16] H. Zhang, W. Xiang, W. Lin and J. Wen, "Grid Forming Converters in Renewable Energy Sources Dominated Power Grid: Control Strategy, Stability, Application, and Challenges," in *Journal of Modern Power Systems and Clean Energy*, vol. 9, no. 6, pp. 1239-1256, November 2021.

[17] A. Hebing, X. Xiao, A. Pfendler and J. Hanson, "Tuning of Cascaded Droop Control Systems for Power Converters Based on the Eigenvalue Analysis," *2023 IEEE Belgrade PowerTech*, Belgrade, Serbia, pp. 1-6,2023.

[18] F. J. Matas-Díaz, M. Barragán-Villarejo, J. C. Olives-Camps, J. M. Mauricio and J. M. Maza-Ortega, "Virtual Conductance Based Cascade Voltage Controller for VSCs in Islanded Operation Mode," in *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 6, pp. 1704-1713, November 2022

[19] C. Li, Y. Yang, N. Mijatovic and T. Dragicevic, "Frequency Stability Assessment of Grid-Forming VSG in Framework of MPME With Feedforward Decoupling Control Strategy," in *IEEE Transactions on Industrial Electronics*, vol. 69, no. 7, pp. 6903-6913, July 2022

[20] Zhidong Sun, Shuai Zhang, Chenghui Zhang, Xiang Tao, Tongyin Dai, "A robust synchronous stability control for virtual synchronous generator of renewable energy adapted to large power grid voltage drop", *Electric Power Systems Research*, Volume 232, April 2024.

View publication stats

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 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/JSTARS.2025.3532829



# What Helps to Detect What? Explainable AI and Multi-Sensor Fusion for Semantic Segmentation of Simultaneous Crop and Land Cover Land Use Delineation

Saman Ebrahimi, Saurav Kumar *Affiliate Member, IEEE*

***Abstract*—This study introduces two novel explainable AI (XAI) frameworks, Interclass-Grad-CAM and Spectral-Grad-CAM, designed to enhance the interpretability of semantic segmentation models for Crop and Land Cover Land Use (CLCLU) mapping. Interclass-Grad-CAM provides insights into interactions between land cover classes, revealing complex spatial arrangements, while Spectral-Grad-CAM quantifies the contributions of individual spectral bands to model predictions, optimizing spectral data use. These XAI methods significantly advance understanding of model behavior, particularly in heterogeneous landscapes, and ensure enhanced transparency in CLCLU mapping. To demonstrate the effectiveness of these innovations, we developed a framework that addresses data asymmetry between the United States and Mexico in the transboundary Middle Rio Grande region. Our approach integrates pixel-level multi-sensor fusion, combining dual-month moderate-resolution optical imagery (July and December 2023), Synthetic Aperture Radar (SAR), and Digital Elevation Model (DEM) data, processed using a Multi-Attention Network (MANet) with a modified Mix Vision Transformer (MiT) encoder to process multiple spectral inputs. Results indicate a uniform improvement in class-specific Intersection over Union (IoU) by approximately 1% with multi-sensor integration compared to optical imagery alone. Optical bands proved most effective for crop classification, while SAR and DEM data enhanced predictions for non-agricultural types. This framework not only improves CLCLU mapping accuracy but also offers a robust tool for broader environmental monitoring and resource management applications.**

***Index Terms*—Optical data, Image/signal analysis (e.g., classification, segmentation, object detection), Deep learning, Vegetation, Others, Urban.**

## I. Introduction

CROP and Land cover and land use (CLCLU) represent a critical variable that encapsulates the complex interactions between anthropogenic activities and the terrestrial environment. It is a dynamic indicator of responses to hydrological and climatic fluctuations, anthropogenic modifications, and policy-driven interventions that impact resource availability and management strategies. In CLCLU, crops are not grouped into a single broad category, allowing for more detailed delineation of specific crops and land use types. In contrast, LCLU often integrates all crops into a broader class, depending on the research focus. [1], [2], [3], [4], [5], [6]. The significance of temporally relevant and detailed CLCLU maps, encompassing a diverse array of specific classes, is paramount for informed decision-making processes. Such high-resolution mapping is essential in accurately reflecting the ongoing temporal and spatial changes in land use patterns, providing crucial insights that can inform policy formulation and adaptive management strategies in response to both natural and anthropogenic perturbations [7], [8], [9], [10], [11], [12]. In transboundary agriculture-dominant regions, where shared resources necessitate coordinated management approaches, precise crop and land cover land use (CLCLU) mapping becomes increasingly salient [13], [14]. For instance, in river basins that traverse multiple national boundaries, such as the Middle Rio Grande, CLCLU data is indispensable for formulating evidence-based policies on water allocation, agricultural practices, and conservation efforts [10], [15]. Detailed CLCLU maps can elucidate areas susceptible to deforestation, land degradation, or urban expansion, facilitating targeted interventions that support sustainable development initiatives. Moreover, in regions characterized by common pool resources, such as shared groundwater aquifers, comprehending CLCLU dynamics is crucial for fostering international collaboration and mitigating potential conflicts over resource utilization [16], [17], [18].

Remote sensing technologies, particularly satellite-based earth observation missions such as Landsat and Sentinel, have revolutionized global-scale monitoring of CLCLU dynamics [19], [20], [21], [22], [23]. These missions provide continuous, open-access imagery, yielding valuable datasets for environmental analysis, including derived products like the Cropland Data Layer (CDL) and various vegetation indices [24], [25], [26]. Recent advancements in artificial intelligence (AI), specifically machine learning (ML) and deep learning (DL), have significantly enhanced the analysis of remotely sensed data for CLCLU classification. Convolutional neural networks (CNNs) have been effective in extracting both spectral and spatial features by using stacked convolution kernels, improving classification performance. Semantic segmentation, in particular, has emerged as a transformative approach, offering pixel-level precision that substantially improves land cover classification accuracy. [27], [28], [29], [30], [31], [32].

The efficacy of these models in LCLU delineation is at-

Saman Ebrahimi is a Graduate Research Assistant in the School of Sustainable Engineering and the Built Environment, Arizona State University, Tempe, AZ, USA.

Saurav Kumar (corresponding author) is with the Faculty of Sustainable Engineering and the Built Environment, Arizona State University, Tempe, AZ, USA.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 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/JSTARS.2025.3532829



tributed to their ability to leverage datasets like annotated land cover datasets such as CORINE Land Cover (CLC), the National Land Cover Database (NLCD), Gaofen-2 satellite images, Urban Segmentation – ISPRS, DeepSat (SAT-4) airborne dataset, EuroSAT dataset, and Munich dataset [22], [23], [33], [34], [35], [36]. Numerous studies have demonstrated the successful application of semantic segmentation models to LCLU classification using satellite imagery. For instance, U-Net architecture with Sentinel-2 data has been used for agricultural field mapping [37], while DeepLabv3+ with Worldview-3 imagery has been used to monitor urban expansion [38]. Furthermore, Vision Transformers (ViT) have been incorporated into feature extraction networks, combining convolution with attention mechanisms to enhance land cover classification accuracy. For example, a recent study introduced a multi-scale fusion module integrated with ViT for CLCLU classification, effectively capturing global and local features to improve model performance in complex land types [39].

While some research showed the effectiveness of utilizing multi-temporal satellite imagery for crop delineation [47],[46]. These applications highlight the potential of deep learning techniques to provide detailed insights into CLCLU dynamics, thereby becoming vital tools for environmental monitoring, urban planning, and resource management.

Data fusion, the integration of multiple data sources to produce more consistent, accurate, and useful information, plays a crucial role in enhancing CLCLU. This approach leverages the complementary strengths of different remote sensing datasets, such as optical and radar data, with different strategies, early, middle, and late fusion, overcome limitations of single-source data and achieve more precise classifications [7], [40], [41]. Optical sensors, such as Landsat-8 and Sentinel-2, provide spectral information across visible and near-infrared wavelengths, enabling distinction between land cover types based on spectral signatures. However, these sensors are limited by atmospheric conditions. Radar-based sensors like Sentinel-1 capture data independent of weather and daylight, offering valuable complementary information [7], [42], [43].

Several studies have demonstrated the efficacy of data fusion in improving CLCLU accuracy. Fusion of Sentinel-2 and Sentinel-1 data was shown to improve urban land cover classification accuracy, particularly by reducing confusion between built-up areas and water bodies [44]. The combination of Sentinel-2 and Landsat-8 data for national-scale crop and land cover mapping in Germany improved classification accuracy by leveraging the temporal and spatial complementarity of the two datasets [45]. Similarly, combining Sentinel-1 and Sentinel-2 data for predicting crop phenological stages was found to be advantageous [48], demonstrating that the fusion of both data types provides more accurate phenological stage predictions than using either data type alone. These studies underscore the potential of data fusion techniques to enhance CLCLU precision, contributing to more effective decision-making in fields such as agriculture, urban planning, and conservation [49].

Despite significant advancements in CLCLU mapping through deep learning, several research gaps persist. These include the reliance on broad land cover annotations that neglect agricultural diversity [50], limited utilization of multispectral data [27], neglect of temporal variations [51], and challenges in model generalizability across diverse regions [52]. Additionally, the inherent opacity of deep learning models raises concerns about their interpretability and transparency, particularly in critical applications. Various interpretability metrics have been developed to address the inherent ”black box” nature of deep learning models and raised limitations [53], [54], [55]. Techniques such as Gradient-weighted Class Activation Mapping (Grad-CAM), Integrated Gradients (IG), and Occlusion Sensitivity have been applied, enhancing model transparency by revealing the spatial regions and spectral bands most influential in classification decisions. However, these approaches offer instance-based insights, leaving a comprehensive understanding of model behavior across diverse circumstances still elusive.

To address the aforementioned gaps, we introduced two analytical frameworks, Interclass-Grad-CAM and Spectral-Grad-CAM, to provide deeper insights into model performance and transparency by:

- Examining the influence of different class types on target class predictions and assess the contribution of specific spectral information to delineating each annotation class.
- Systematically evaluating the importance of each spectral band, this approach transcends instance-based insights to provide a holistic understanding of model behavior across the entire dataset, significantly enhancing explainability and offering valuable perspectives for CLCLU mapping and resource management applications.

To this end, we integrate pixel-level multi-sensor fusion, combining Synthetic Aperture Radar (SAR) Sentinel-1 imagery with optical composites and Digital Elevation Model (DEM) data to enhance model performance. A semantic segmentation model, Multi Attention Network (MANet) Architecture [56], that contains a Mix Vision Transformer (MiT) encoder [57] for processing the multispectral input was developed in this study.

## II. Study Area

This study investigates CLCLU in the Middle Rio Grande Basin (34° 03’ 36” N to 29° 22’ 54” N, 107° 51’ 25” W to 104° 12’ 56” W), encompassing 36,988 km² (Fig.1). The basin extends 592 km from San Antonio, NM, to Presidio, TX, and Ojinaga, Chihuahua, with varying width (37 to 63 km). The region’s heterogeneous landscape comprises urban areas (e.g., El Paso, Ciudad Juárez), agricultural lands, and desert ecosystems across alluvial plains and mountainous terrain [58]. The Rio Grande River, originating in Colorado’s San Juan Mountains, serves as a critical hydrological resource, governed by international treaties: 1906, and 1944 [59], [60]. This research aims to enhance understanding of CLCLU dynamics to inform adaptive water resource management strategies in this ecologically and socio-economically vital transboundary watershed required to address the facing challenges, i.e., climate change-induced alterations in surface water availability, escalating anthropogenic demand, and diminishing river discharge [59], [61].

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 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/JSTARS.2025.3532829



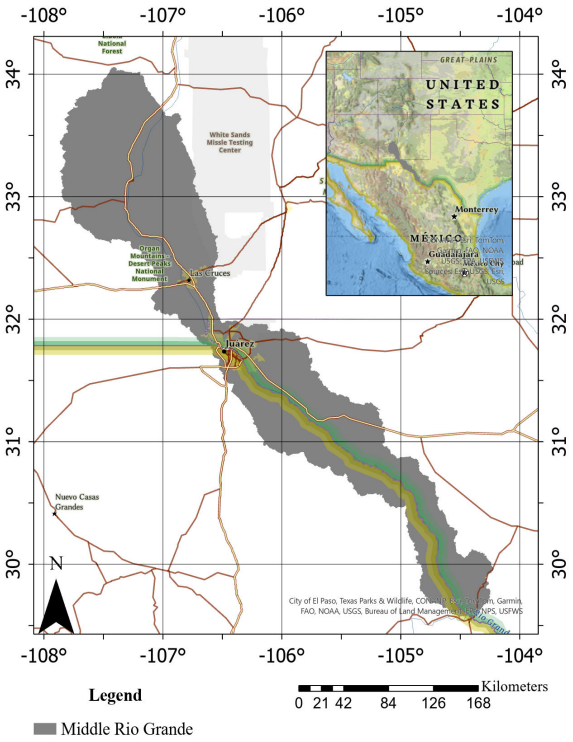



Fig. 1. Middle Rio Grande watershed

## III. Methodology

### A. Data Acquisition

*1) Optical Imagery:* This study utilized Collection 2, Level 2, Tier 1 Landsat 8 OLI/TIRS sensor imagery for CLCLU analysis, employing corrected surface reflectance and land surface temperature data [20], [62] (Table I). The dataset comprises four VNIR, two SWIR, and one TIR band processed to orthorectified surface temperature with 30-meter spatial resolution, corresponding to July and December 2023. The utilized collection incorporates significant improvements, including enhanced geometric and radiometric corrections, removal of extreme latitude DEM constraints, improved precision correction algorithms, and adjustments for the Landsat 8 TIRS anomaly. These enhancements result in increased accuracy, particularly in high-latitude areas, and improved overall data quality [62].

*2) Sentinel-1 Synthetic Aperture Radar (SAR):* The Sentinel-1 mission provides data from a dual-polarization C-band Synthetic Aperture Radar (SAR) instrument operating at 5.405 GHz. The dataset includes Ground Range Detected (GRD) products, which are processed to produce calibrated and ortho-corrected scenes [63]. Sentinel-1 operates in multiple imaging modes, and the available polarizations depend on the acquisition settings. In dual-polarization acquisitions, such as VV+VH (vertical transmit, vertical/horizontal receive) and HH+HV (horizontal transmit, horizontal/vertical receive), the data are stored as individual bands (e.g., VV, VH, HH, HV) in the final GRD product. Sentinel-1 GRD products are provided at a spatial resolution of 10 meters, after pre-processing steps such as thermal noise removal, radiometric calibration, and terrain correction using the SRTM 30 Digital Elevation Model (DEM) or ASTER DEM for areas above 60 degrees latitude where SRTM data is unavailable [19], [64].

*3) Copernicus Digital Elevation Model (DEM):* The Copernicus DEM is a Digital Surface Model (DSM) representing the Earth's surface, including buildings, infrastructure, and vegetation. This DEM is derived from the WorldDEM™ product, which incorporates edits for flattening water bodies and ensuring consistent river flow. The WorldDEM product is based on radar satellite data acquired during the TanDEM-X Mission, funded by a Public-Private Partnership between the German Aerospace Centre (DLR) and Airbus Defence and Space. The Copernicus DEM offers multiple resolutions, with slope computations typically requiring reprojection [65], [66].

*4) Fusion of Multi-Sensor Data:* We performed pixel-level fusion by combining multispectral imagery from Landsat 8 with synthetic aperture radar (SAR) data and digital elevation model (DEM) data with a homogenous spatial resolution of 30 m. Multispectral imagery can be defined as three-dimensional matrices, where the height and width correspond to the spatial dimensions (image height and width), and the depth (third dimension) represents the spectral resolution, encompassing multiple wavelength bands. The fusion of these datasets was conducted using the Google Earth Engine (GEE) Code Editor, leveraging its powerful cloud-based platform to efficiently process, fuse, and extract data across multiple sensor types. This approach allows us to investigate how the additional information provided by SAR and DEM data can enhance CLCLU mapping by segmentation models. The data fusion methodology commenced with the integration of Synthetic Aperture Radar (SAR) data, specifically vertical transmit-vertical receive (VV) and vertical transmit-horizontal receive (VH) polarizations, with bi-temporal Landsat 8 imagery corresponding to July and December 2023. SAR data underwent normalization to a dynamic range of -50 to 0 decibels (dB). For each temporal instance, median composites were generated for both optical and SAR datasets. These monthly composites were subsequently concatenated to form a comprehensive multi-temporal, multi-sensor dataset. The resultant dataset comprised the following components: monthly median composites of Landsat spectral bands (SR_B2, SR_B3, SR_B4, SR_B5, SR_B6, SR_B7, ST_B10), monthly Normalized Difference Vegetation Index (NDVI), monthly median composites of SAR polarimetric bands (VV and VH), Corresponding to July and December, 2023, and the inclusion of the Copernicus Digital Elevation Model (DEM) to provide topographical context. DEM was normalized (0-1) for an elevation range of 0 (min) to 5000 (max) meters (Fig. 2.A). This integrated approach facilitates a synergistic analysis of multispectral, radar, and topographic data. Table I provides a detailed description of the spectral architecture and characteristics of the dataset used for training, highlighting the key properties and configurations of each spectral band.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 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/JSTARS.2025.3532829



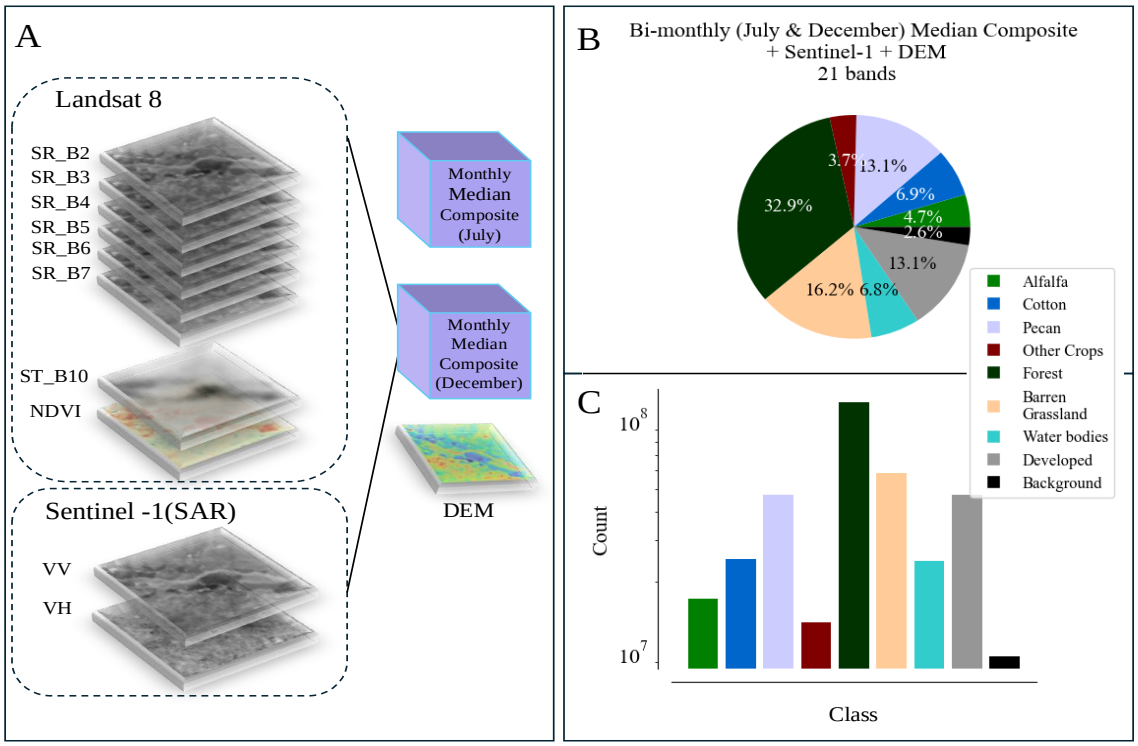



Fig. 2. Characteristics of the Multisensor Dataset (21 Bands); A: Spectral characteristics of the dataset composites, Landsat 8, Sentinel-1 SAR, and DEM data. B: Proportion of each class with respect to the total number of pixels, highlighting the distribution of land cover types within the dataset. C: Pixel count for each class, presented on a logarithmic scale to emphasize the relative abundance of different land cover types.

In addition, this study incorporated the USDA NASS Cropland Data Layer (CDL) for 2023 as ground truth for training and testing datasets. The CDL, derived from Landsat 8/9 and Sentinel-2A/B imagery, provides 30-meter resolution coverage only of the continental United States. It utilizes agricultural training and validation data from the FSA Common Land Unit Program, supplemented by crop-specific ground reference data from various state and federal agencies. This comprehensive dataset offers a reliable basis for land use classification and validation in agricultural areas within the study region [26]. The Cropland Data Layer (CDL) raster analysis for the U.S. portion of the Middle Rio Grande Basin revealed 41 distinct land cover types between 2008 and 2023. The agricultural landscape was dominated by three primary crops: alfalfa, cotton, and pecan, with the remaining crop types consolidated into an "other crops" category. Non-agricultural land cover was classified into four general categories: Forest/Shrubland, Barren/Grassland, Water bodies, Developed areas, and background. Table 1 provides a detailed breakdown of the land cover classification scheme. This comprehensive categorization enables a nuanced analysis of land use patterns and changes within the study area over the specified time period. This multi-sensor dataset enables a comprehensive analysis of the influence of each data type on CLCLU mapping, enhancing the model's capability to capture various land cover types and improve classification accuracy. The tiling process involved generating 10,000 tiles (64x64 pixels) for each class for training and 1,000 or more tiles for validation and testing, ensuring no overlap in center pixels between tiles to maintain the independence of training samples (Fig. 2 B&C).

### B. Semantic segmentation

This study presents a novel approach to semantic segmentation for CLCLU detection, with potential applications in autonomous navigation, medical imaging, and object recognition. The proposed architecture integrates a modified fully transformer encoder based on Mix Vision Transformers (MiT), suited for multispectral imagery analysis, with a Multi Attention Network (MANet) decoder.

The experimental design involved training the model on a specified dataset under two conditions: utilizing solely optical imagery and incorporating fused additional features. The training process consisted of 200 epochs with a batch size of 128 and an initial learning rate of 0.001. To mitigate class imbalance in the training data, a weighted cross-entropy loss function was implemented, defined as:

$$\text{Weighted Cross Entropy Loss} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} w_c \cdot y_{i,c} \cdot \log\left(\hat{y}_{i,c}\right) \tag{1}$$

where $y_{i,c}$ denotes the binary indicator if class $c$ is the correct classification for the sample $i$, $\hat{y}_{i,c}$ is the predicted probability that sample $i$ belongs to class $c$, and $N$ is the dataset size.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 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/JSTARS.2025.3532829



TABLE I
SPECTRAL CHARACTERISTICS OF THE MULTI-SENSOR DATASET

| Band Name | Description | Wavelength (µm) | Min | Max | Scale/Units |
|---|---|---|---|---|---|
| Band 1 (July) | Blue | 0.452–0.512 | 1 | 65455 | 2.75E-05 |
| Band 2 (July) | Green | 0.533–0.590 | 1 | 65455 | 2.75E-05 |
| Band 3 (July) | Red | 0.636–0.673 | 1 | 65455 | 2.75E-05 |
| Band 4 (July) | Near Infrared | 0.851–0.879 | 1 | 65455 | 2.75E-05 |
| Band 5 (July) | Shortwave Infrared 1 | 1.566–1.651 | 1 | 65455 | 2.75E-05 |
| Band 6 (July) | Shortwave Infrared 2 | 2.107–2.294 | 1 | 65455 | 2.75E-05 |
| Band 7 (July) | Orthorectified Surface Temperature | 10.60–11.19 | 0 | 65535 | 3.42E-03 |
| Band 8 (July) | Normalized Difference Vegetation Index (NDVI) | - | -1 | 1 | - |
| Band 9 (July) | Vertical-Vertical (VV) polarization-SAR | - | -50 | 1 | dB |
| Band 10 (July) | Vertical-Horizontal (VH) polarization-SAR | - | -50 | 1 | dB |
| Band 11 (December) | Blue | 0.452–0.512 | 1 | 65455 | 2.75E-05 |
| Band 12 (December) | Green | 0.533–0.590 | 1 | 65455 | 2.75E-05 |
| Band 13 (December) | Red | 0.636–0.673 | 1 | 65455 | 2.75E-05 |
| Band 14 (December) | Near Infrared | 0.851–0.879 | 1 | 65455 | 2.75E-05 |
| Band 15 (December) | Shortwave Infrared 1 | 1.566–1.651 | 1 | 65455 | 2.75E-05 |
| Band 16 (December) | Shortwave Infrared 2 | 2.107–2.294 | 1 | 65455 | 2.75E-05 |
| Band 17 (December) | Orthorectified Surface Temperature | 10.60–11.19 | 0 | 65535 | 3.42E-03 |
| Band 18 (December) | Normalized Difference Vegetation Index (NDVI) | - | -1 | 1 | - |
| Band 19 (December) | Vertical-Vertical (VV) polarization-SAR | - | -50 | 1 | dB |
| Band 20 (December) | Vertical-Horizontal (VH) polarization-SAR | - | -50 | 1 | dB |
| Band 21 | Digital Elevation Model | - | 0 | 5000 | m |

TABLE II
CDL RASTER VALUES FOR AUGMENTATIONS

| Class | CDL pixel value |
|---|---|
| Alfalfa/Hey | 36, 37 |
| Cotton | 2, 238 |
| Pecan | 74 |
| Other crops | 1, 4, 12, 22, 23, 24, 27, 28, 42, 47, 48, 49, 71, 74, 204, 205, 209, 216, 225, 226, 227, 236, 243 |
| Forest/Shrubland | 141, 142, 143, 152 |
| Barren/Grassland | 61, 131, 176 |
| Water Bodies | 111, 190, 195 |
| Developed | 121, 122, 123, 124 |
| Background | NaN / Others |

Class weights, $w_c$, are computed using the inverse frequency formula.

$$w_c = \frac{N}{N_c} \tag{2}$$

where $N$ signifies the total number of pixels in the dataset and $N_c$ is the number of pixels in class $c$. To prevent extreme values, weights are normalized by dividing each by the minimum weight, ensuring a minimum value of 1. The Adaptive Moment Estimation (Adam) optimizer was employed for model training. Additionally, a 'ReduceLROnPlateau' learning rate scheduler was implemented to dynamically adjust the learning rate based on validation set performance. This scheduler reduces the learning rate by a factor of 0.1 if no improvement in validation loss is observed over 5 consecutive epochs, facilitating more effective model convergence. All computational procedures were executed on an NVIDIA RTX A6000 GPU.

*1) MANet_B5:* The Multi-Attention-Network (MANet) is a novel semantic segmentation architecture specifically designed for high-resolution remote sensing imagery [56]. This architecture incorporates an innovative attention mechanism that efficiently captures intricate feature combinations extracted by the encoder module. The key innovation of MANet lies in its attention mechanism, which integrates kernel attention with linear complexity, significantly reducing the computational requirements of the attention module. The attention mechanism generalizes the dot-product attention (Equation 3) into a kernel-based formulation (Equation 4), referred to as the Kernel Attention Mechanism (KAM).

$$D(Q, K, V) = \text{softmaxrow}(QK^T)V \tag{3}$$

where:

- $Q$: Query matrix, representing the encoded feature map.
- $K$: Key matrix, which is used to compute the attention weights in conjunction with the Query.
- $V$: Value matrix, which provides the features weighted by the attention mechanism.
- $QK^T$: The dot-product between the Query and Key matrices, which calculates the attention scores.
- softmaxrow: Applies the softmax function to each row of the attention matrix.

The kernel-based formulation of attention (KAM) is expressed as:

$$D(Q, K, V)_i = \frac{\sum_{j=1}^{N} \phi(q_i)^T \phi(k_j)\nu_j}{\sum_{j=1}^{N} (\phi(q_i)^T \phi(k_j))} \tag{4}$$

where:

- $q_i$: The i-th query vector.
- $k_j$: The j-th key vector.
- $\nu_j$: The j-th value vector.
- $\phi(x)$: A kernel function applied to the input $x$, as described in Equation 5.

$$\phi(x) = \log(1 + e^x) \tag{5}$$

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 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/JSTARS.2025.3532829



This kernel-based approach enhances the flexibility of the attention mechanism, allowing it to capture more complex relationships between the query and key vectors. Additionally, MANet integrates the Kernel Attention Mechanism (KAM) with the Channel Attention Mechanism (CAM) to further improve segmentation performance. The CAM computes attention scores over channels as follows:

$$\text{CMP} = \text{MaxPool}(x), \quad \text{CAP} = \text{AvgPool}(x) \tag{6}$$

where:

- $\text{MaxPool}(x)$: The max-pooling operation applied over the feature map $x$, capturing the most prominent features across spatial dimensions.
- $\text{AvgPool}(x)$: The average pooling operation applied over the feature map $x$, summarizing the average feature responses across spatial dimensions.

The implemented MANet architecture in this study incorporates a modified version of Mix Transformer (MiT) encoders, initially proposed by [67], enabling the processing of multispectral input data (beyond the typical 3 bands) for deep, high-level feature extraction. The MiT encoder's structure progressively extracts features at multiple resolutions through a synergistic combination of convolutional layers, layer normalization, ReLU activation functions, and efficient self-attention mechanisms, significantly mitigating computational complexity while maintaining the capacity to capture intricate spatial and spectral relationships. The traditional multi-head self-attention mechanism operates on query (Q), key (K), and value (V) components for each channel, where $N = H \times W$ represents the total number of pixels in the input feature map. This architectural adaptation leverages the strengths of transformer-based models while optimizing for the spatial nature of remote sensing imagery, potentially enhancing the model's capability to discern subtle variations in land cover and land use through focused analysis of salient features across the multispectral input.

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_{\text{head}}}}\right) \tag{7}$$

Where the feature vector dimension per attention head, $d_{\text{head}}$, exhibits a computational complexity of $O(N^2)$, rendering it impractical for large-scale image processing. To mitigate this, an efficient self-attention mechanism is implemented, reducing the sequence length by a factor of $R$ (Equation 7), significantly decreasing computational demands while maintaining the ability to capture long-range dependencies. This modification enables more efficient processing of high-resolution imagery in remote sensing applications.

$$\hat{K} = \text{Reshape}(N/R, C \cdot R)(K) \tag{8}$$

$$K = \text{Linear}(C \cdot R, C)(\hat{K}) \tag{9}$$

The encoder architecture employs a multi-stage design incorporating Mix Feed Forward Networks (Mix-FNN) to integrate positional information. To accommodate multispectral data, the original MiT encoders, which were limited to three input bands, were modified to accept multiple spectral inputs. The decoder's structure is symmetrical to the encoder, utilizing convolutional layers in conjunction with Kernel Attention Mechanisms (KAM) and Channel Attention Mechanisms (CAM) for upsampling and refining feature maps. This process restores the spatial resolution to that of the original input, rendering the architecture particularly effective for high-resolution image segmentation tasks in remote sensing applications.

*2) Evaluation Metrics:* A suite of quantitative metrics was utilized to provide a comprehensive assessment of the model's efficacy. These metrics were selected to elucidate various aspects of model performance, including overall accuracy, class-specific precision, and the ability to generalize across diverse class distributions. This rigorous evaluation framework enables a nuanced understanding of the model's strengths and limitations in the context of multi-class segmentation tasks.
The $\text{IoU}_C$ quantifies the overlap between the predicted segmentation mask and the ground truth mask. The $\text{IoU}_C$ is mathematically defined as:

$$\text{IoU}_C = \frac{\text{TP}_C}{\text{TP}_C + \text{FP}_C + \text{FN}_C} \tag{10}$$

where $\text{TP}_C$ represents true positives, $\text{FP}_C$ represents false positives, and $\text{FN}_C$ represents false negatives for class $C$. Additionally, the mean $\text{IoU}_C$ value (mIoU) was computed to provide an aggregated performance measure across all classes.

The next metric, Accuracy, quantifies the proportion of correctly classified pixels over the total number of pixels. It is calculated as follows:

$$\text{Accuracy} = \frac{1}{C}\sum_{c=1}^{C}\frac{\text{TP}_C + \text{TN}_C}{\text{TP}_C + \text{TN}_C + \text{FP}_C + \text{FN}_C} \tag{11}$$

Where $C$ is the number of classes, and $\text{TN}_C$ represents true negatives for class $C$.

Recall, also termed sensitivity or true positive rate, measures the model's ability to correctly identify all relevant instances. It is calculated as:

$$\text{Recall} = \frac{1}{C}\sum_{c=1}^{C}\frac{\text{TP}_C}{\text{TP}_C + \text{FN}_C} \tag{12}$$

The F1 score represents the harmonic mean of precision and recall, providing a balance between model performance. It is particularly valuable in scenarios with class imbalance.

$$F_1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{13}$$

where Precision is the ratio of $\text{TP}_C$ to the sum of $\text{TP}_C$ and $\text{FP}_C$. The F2 score is similar to the F1 score, but it leans more towards recall:

$$F_2 = 5 \cdot \frac{\text{Precision} \cdot \text{Recall}}{4 \cdot \text{Precision} + \text{Recall}} \tag{14}$$

This suite of metrics enables a comprehensive and nuanced evaluation of the segmentation models' performance across various aspects of classification accuracy and generalization capability.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully edited content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2025.3532829



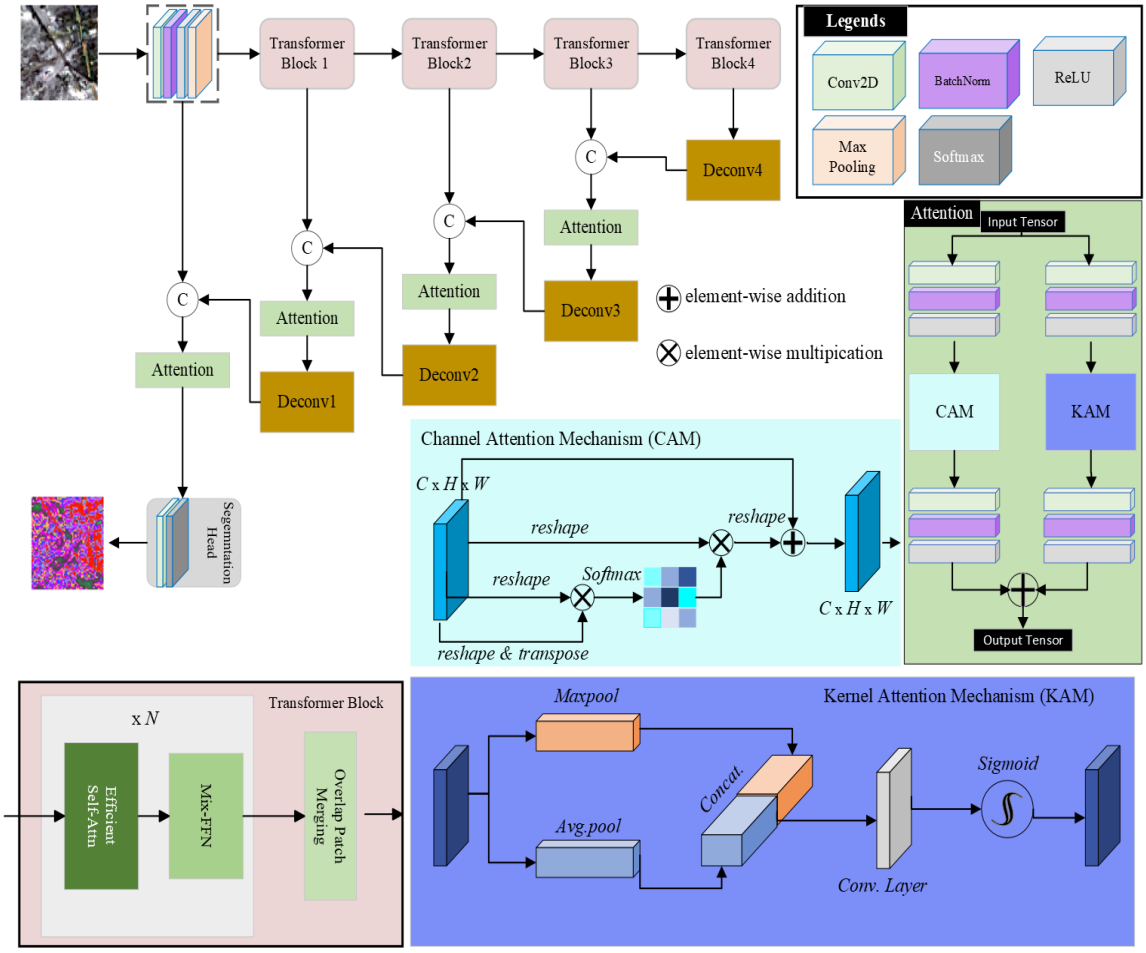



Fig. 3. The implemented MANet Architecture with MiT-B5 Encoder

### C. Explainable AI (XAI) Methods

Various techniques have been developed to generate visual explanations for Convolutional Neural Network (CNN)-based models, with the objective of enhancing their transparency and interpretability. The subsequent section elucidates the methods employed for this end, as well as the frameworks designed to provide a comprehensive analysis of the presence of each class within the input tile and the impact of individual spectral bands on the model's decision-making process. These techniques aim to elucidate the internal mechanisms of deep learning models in the context of remote sensing image analysis, facilitating a more thorough understanding of their detection and segmentation processes. To gain comprehensive insights into the model's prediction process, we developed two modules: 1. Interclass-Grad-CAM and 2. Spectral-Grad-CAM. These modules quantify the impact of each class presence on other class predictions and investigate individual band (feature) importance, respectively.

*1) Gradient-weighted Class Activation Mapping (Grad-CAM):* Grad-CAM utilizes gradient information propagating to the final convolutional layer of a Convolutional Neural Network (CNN) to assign importance values to individual neurons for a specific decision of interest [68]. To generate a class-discriminative localization map $L^c_{\text{Grad-CAM}}$, with respect to feature map activations $A^K$ of a convolutional layer, i.e., $\frac{\partial y^c}{\partial A^K}$, the method employs global average pooling of the backpropagated gradients across the spatial dimensions of the input tile. The neuron importance weights, $\alpha^c_k$, were computed as follows:

$$\alpha^c_k = \frac{1}{Z}\sum_i\sum_j \frac{\partial y^c}{\partial A^k_{ij}} \tag{15}$$

where $Z$ is the number of spatial locations, and $\frac{\partial y^c}{\partial A^k_{ij}}$ represents the gradients of the class score $y^c$ with respect to the feature map activations $A^K$ at location $(i, j)$. In the gradient backpropagation process for computing the weight $\alpha^c_k$ with respect to activations, the method involves successive matrix multiplications of weight matrices and gradients relative to activation functions, leading back to the target convolutional layer. Empirical evidence suggests that global average pooling is more effective than global max pooling in this context [33], [70]. The weight $\alpha^c_k$ serves as a partial linearization of the network downstream from the activation map $A$, indicating the importance of feature map $k$ for a target class $c$. The method subsequently combines forward activation maps using these weights and applies a ReLU function to extract class-specific feature importance, providing deeper insights into the

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 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/JSTARS.2025.3532829



network's decision-making process for image classification and segmentation tasks.

$$L^c_{\text{Grad-CAM}} = \text{ReLU}\left(\sum_k \alpha^c_k A^k\right) \tag{16}$$

Where positive $L^c_{\text{Grad-CAM}}$ values show pixels contributing to the detection of class $c$, while negative values indicate pixels that did not contribute to the classification as class $c$. Ideally, a robust model will have high $L^c_{\text{Grad-CAM}}$ values where the class under investigation is physically present. Contributions from other objects in the scene may be a result of overtraining and reflect suboptimal model behavior.

*2) Integrated Gradients(IG):* This method [71] offers a robust approach to attributing a deep learning model's prediction to its input features. This technique is grounded in the concept of evaluating gradients along the path from a baseline input (typically a constant image like black or blurred) to the actual input image. The IG method addresses the limitations of standard gradient-based methods, which often suffer from noise and local optima. Specifically, IG calculates the gradients along the linear interpolation between the baseline $\tilde{I}$ and the input $I$, summing these gradients to produce the final attribution. The formula for Integrated Gradients is expressed as:

$$\text{IG}_i(x) = (x_i - \tilde{x}_i) \times \int_{\alpha=0}^{1} \frac{\partial f(\tilde{x} + \alpha(x - \tilde{x}))}{\partial x_i} d\alpha \tag{17}$$

Where $f$ denotes the model prediction function, $x$ is the input, $\tilde{x}$ is the baseline, and $i$ represents the individual input feature.

*3) Occlusion :* The occlusion method [53] is a perturbation-based technique used to generate saliency maps that identify important regions in an image for a CNN's decision-making process. In this method, parts of the input image $I$ are systematically occluded by masking them with a baseline value (e.g., zero or the mean pixel value). The modified image $\tilde{I}$ is then passed through the CNN, and the impact on the prediction score for the target class $c$ is measured. The importance $S_{(i,j)}$ of a patch centered at pixel $(i, j)$ is calculated as:

$$S_{(i,j)} = f_c(I) - f_c(\tilde{I}_{(i,j)}) \tag{18}$$

where $f_c(I)$ is the original prediction score, and $f_c(\tilde{I}_{(i,j)})$ is the score with the patch occluded. By repeating this for all patches, a saliency map $S$ is generated, highlighting regions most influential to the network's output. Adjustments to patch size $p$ ($p = 15$) and step size $s$ ($s = 8$) can balance the resolution of the saliency map with computational efficiency. This method provides valuable insights into model interpretability, especially in critical applications like medical imaging.

The aforementioned approaches provide instance-based saliency maps, offering an overview of the model's logic in the prediction process for each input instance. However, a comprehensive understanding of the model's behavior remains elusive, particularly concerning the impact of class combinations on predictions for a specific class and the relative influence of individual features on the model's decision-making process.

*4) Interclass-Grad-CAM:* Building upon the Grad-CAM metric, we introduce a novel method to quantify the influence of other class pixels on the prediction of a target class $c$. $\text{PositiveImpact}_{(c,i)}$ and $\text{NegativeImpact}_{(c,i)}$ were quantified by counting the number of pixels where the Grad-CAM attribution of the last convolutional layer, and the model's predicted class label $i$, matched the ground truth class label. This can be expressed as:

$$\text{PositiveImpact}_{(c,i)} = \sum_{n=1}^{N} \sum_{(x,y)} 1\left(L^c_{\text{Grad-CAM}}(x,y) > 0 \;\&\; \text{Prediction}(x,y) = i\right) \tag{19}$$

$$\text{NegativeImpact}_{(c,i)} = \sum_{n=1}^{N} \sum_{(x,y)} 1\left(L^c_{\text{Grad-CAM}}(x,y) < 0 \;\&\; \text{Prediction}(x,y) = i\right) \tag{20}$$

where $1(\cdot)$ is the indicator function, $n$ represents the number of tiles in the dataset, and $x, y$ are the tile dimensions. The attribution values were stored based on their ground truth labels for each target class. To enable a meaningful comparison of the impacts across different classes, the normalized positive and negative impacts were calculated as follows.

$$\text{NormalizedPositiveImpact}_{(c,i)} = \frac{\text{PositiveImpact}_{(c,i)}}{\max_{0 \le j \le J} \begin{pmatrix}\text{PositiveImpact}_{(c,j)}, \\ \text{NegativeImpact}_{(c,j)}\end{pmatrix}} \tag{21}$$

$$\text{NormalizedNegativeImpact}_{(c,i)} = \frac{\text{NegativeImpact}_{(c,i)}}{\max_{0 \le j \le J} \begin{pmatrix}\text{PositiveImpact}_{(c,j)}, \\ \text{NegativeImpact}_{(c,j)}\end{pmatrix}} \tag{22}$$

where $j$ represents the number of classes in the dataset. We also computed $\text{CumulativeImpact}_{(c,i)}$ as:

$$\text{Cum.Impact}_{(c,i)} = \text{NormalizedPositiveImpact}_{(c,i)} + \text{NormalizedNegativeImpact}_{(c,i)} \tag{23}$$

*5) Spectral-Grad-CAM:* To evaluate the contribution of individual spectral bands to the prediction of a target class $c$ within a multi-band segmentation framework, we employed the following procedure. For each input image $I$ containing the target class $c$ in its ground truth, we generated modified inputs $I_b$ by retaining the $b^{\text{th}}$ spectral band while zeroing out the remaining bands, defined as:

$$I_b = \begin{cases} b^{\text{th}} \text{ band of } I, & \text{if } b = \text{target band} \\ 0, & \text{otherwise} \end{cases} \tag{24}$$

Subsequently, the Grad-CAM attribution, $L^c_{\text{Grad-CAM}}$, was computed for each modified input $I_b$ with respect to the target class (Algorithm 1).

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 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/JSTARS.2025.3532829



**Algorithm 1** Spectral Impact on Model Prediction.

**Set** $n$ as the number of tiles in the validation dataset.
**Set** $c$ as the classes predicted by the model for the $n^{\text{th}}$ tile.
**Set** $b$ as the number of input bands.
**Define** $I_b$ as:

$$I_b = \begin{cases} \text{bth band of } I, & \text{if } b = \text{target band} \\ 0, & \text{otherwise} \end{cases}$$

**Compute** $L^c_{\text{Grad-CAM}} = \text{ReLU}\left(\sum_k \alpha^c_k A^k\right)$.
**Store** the attribution values.
$b \rightarrow b + 1$.
$c \rightarrow c + 1$.
$n \rightarrow n + 1$.

This framework facilitates the assessment of model performance sensitivity with respect to individual input features. By systematically isolating each spectral band and computing the corresponding Grad-CAM attributes, it enables a quantitative evaluation of the relative importance of each band in the model's decision-making process. This approach provides insights into the spectral dependencies of the model and allows for a more nuanced understanding of how different wavelengths contribute to classification accuracy across various land cover types.

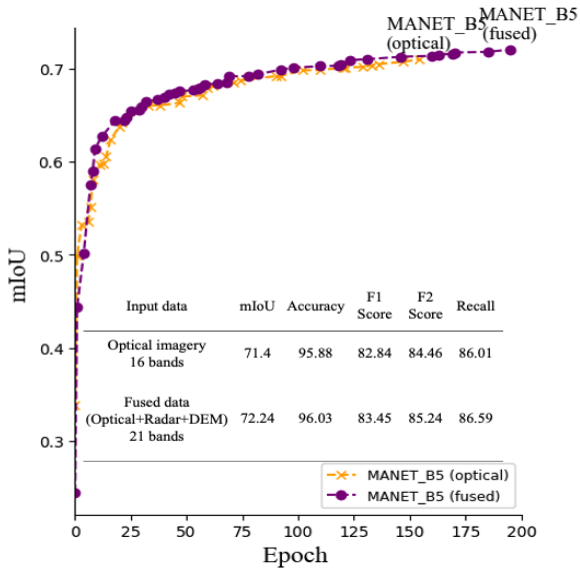
## IV. Results

### A. Model Evaluation

Figure 4 illustrates the evolution of the model's mIoU on the validation dataset as a function of training epochs. The results demonstrate an improvement in model performance upon integration of The mIoU of the model over 200 epochs of training across two different input configurations: (A) Dual-month (July & December) Median Optical Composite and (B) multisensory composite (Including Optical, SAR and DEM). Specifically, the mIoU increased from 71.4% using solely optical data to 72.24% with the inclusion of SAR and DEM data. This enhancement in overall accuracy is further corroborated by Figure 5, which depicts a uniform increase in class-wise IoU ($\text{IoU}_c$) across all land cover categories.

Normalized confusion matrices provide a quantitative representation of classification accuracy by illustrating the proportion of correctly predicted instances relative to the total instances for each class (Figure 6). This normalization facilitates a more robust interpretation of model performance across various land cover and land use categories, particularly in the presence of class imbalance. The analysis elucidates the models' efficacy in discriminating between distinct land cover types prevalent in the study region. The integration of multisensor data resulted in a statistically significant improvement in overall model accuracy across all classes (Figure 6.A). However, both models exhibited relatively lower classification accuracy for Forest/Shrubland and Barren/Grassland classes compared to other land cover types.

Figure 7 illustrates a comparative analysis of the predicted land cover masks generated by the model across various data




| Input data | mIoU | Accuracy | F1 Score | F2 Score | Recall |
|---|---|---|---|---|---|
| Optical imagery 16 bands | 71.4 | 95.88 | 82.84 | 84.46 | 86.01 |
| Fused data (Optical+Radar+DEM) 21 bands | 72.24 | 96.03 | 83.45 | 85.24 | 86.59 |

Fig. 4. The mean Intersection over Union (mIoU) of the segmentation mode (MANet-B5) over 200 epochs of training across two different input imagery configurations: (A) Dual-month (July & December) Landsat 8 Median Composite (16 bands), and (B) Dual-month optical imager + synthetic aperture radar (SAR) data and digital elevation model (DEM) (21 bands).

configurations alongside the corresponding ground truth masks derived from the CDL for each sample tile. This visualization enables a qualitative evaluation of the model's performance under different input spectral configurations.

### B. Explainable AI (XAI)

*1) Conventional XAI methods:* We employed three distinct explainability metrics—Gradient-weighted Class Activation Mapping (Grad-CAM), Integrated Gradients (IG), and Occlusion—to generate class-specific saliency maps for each input tile, as illustrated in Figures 8 and 9. These figures present a comparative analysis encompassing the input tile, the predicted segmentation mask, the ground truth mask, and the saliency maps generated by each explainability method. This analysis contributes to the broader understanding of model interpretability in remote sensing applications and highlights the potential for these techniques to enhance trust in and adoption of deep learning models for environmental monitoring and land use detection tasks.

The Grad-CAM saliency maps, normalized to the range [-1, 1], delineate pixels exhibiting both positive and negative contributions to the detection of each specific class. Results indicate that Grad-CAM effectively identifies regions within the input tile that significantly influence the model's predictions, demonstrating high spatial correspondence with both the predicted segmentation mask and the ground truth. This suggests that Grad-CAM provides valuable insights into the spatial features driving the model's classification decisions. Conversely, the saliency maps generated by the IG method, normalized to the range [0, 1], did not provide equivalent clarity regarding the model's prediction process for individual classes. While IG maps may be valuable in certain contexts, they appear less efficacious in this particular application for elucidating

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 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/JSTARS.2025.3532829



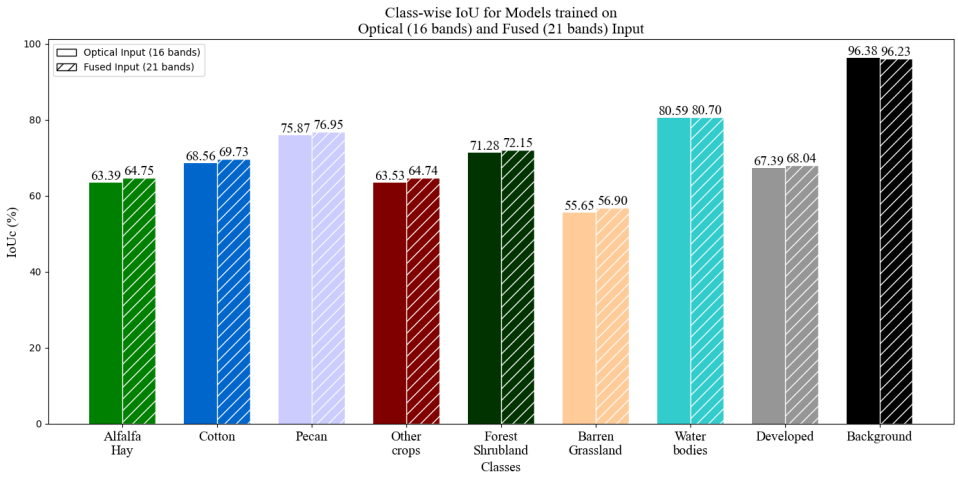



Fig. 5. Class-wise IoU ($IoU_c$) for two input data configurations. This figure shows the percentage change in IoUc using various data configurations:1. Optical Imagery (16 bands), 2. Optical Imagery fused with Synthetic Aperture Radar (SAR) data and Digital Elevation Model (DEM) (21 bands).

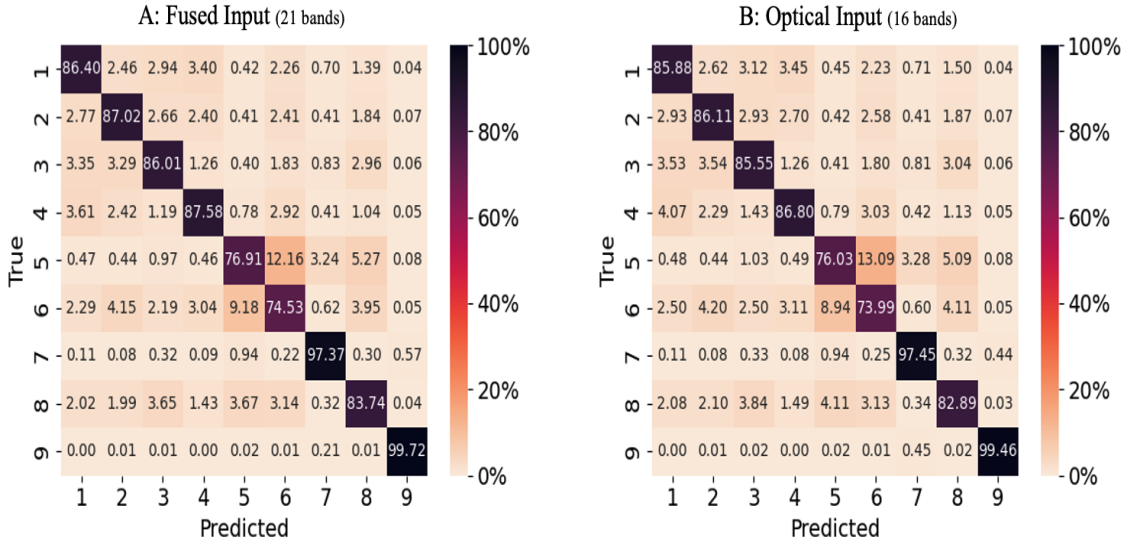



Fig. 6. Normalized Confusion Matrices for Two different input data configurations: A. Optical Imagery fused with Synthetic Aperture Radar (SAR) data and Digital Elevation Model (DEM) (21 bands), B. Optical Imagery (16 bands); Classes: 1- Alfalfa/Hay, 2- Cotton, 3- Pecan, 4. Other Crops, 5- Forrest/ Shrubland, 6- Barren/Grassland, 7- Water bodies, 8- Developed

detailed contributions of individual pixels or regions in the input tile. The Occlusion method, also normalized to the range [0, 1], presents a more distinct visualization of salient regions within the input tiles. By systematically occluding portions of the input and quantifying the resultant effect on the model's predictions, this metric provides a direct visual representation of the areas most critical to the model's decision-making process. However, it is noteworthy that the Occlusion method demonstrates limitations in revealing insightful details when a class is represented by only a few pixels. This constraint highlights the method's reduced effectiveness in capturing fine-grained spatial information for sparsely distributed or small-scale land cover features. Comparative analysis of these three explainability techniques reveals varying degrees of insight into the model's internal mechanisms. Grad-CAM and Occlusion methodologies yield more intuitive and interpretable results than IG in this specific context, with Grad-CAM potentially offering superior performance for classes with limited spatial extent. These findings underscore the importance of judicious selection of explainability tools when interpreting complex models, particularly in tasks involving spatially heterogeneous data such as land cover classification.

*2) Interclass-Grad-CAM:* Figure 10 illustrates the interaction of different classes within the model's predictions by depicting the $\text{NormalizedPositiveImpact}_{(c,i)}$, $\text{NormalizedNegativeImpact}_{(c,i)}$, and their cumulative impact for specific crop classes: Alfalfa/Hay, Cotton, Pecan, and Other Crops. For Class 0: Alfalfa/Hay, the metrics indicate that this class is generally distinguishable from all other classes, except for Pecan (Cum. Impact $< 0.20$), Other Crops (Cum. Impact $< 0.10$), and Background with a positive Cum. Impact value. In the case of Class 1: Cotton, the presence of Alfalfa/Hay

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully edited content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2025.3532829



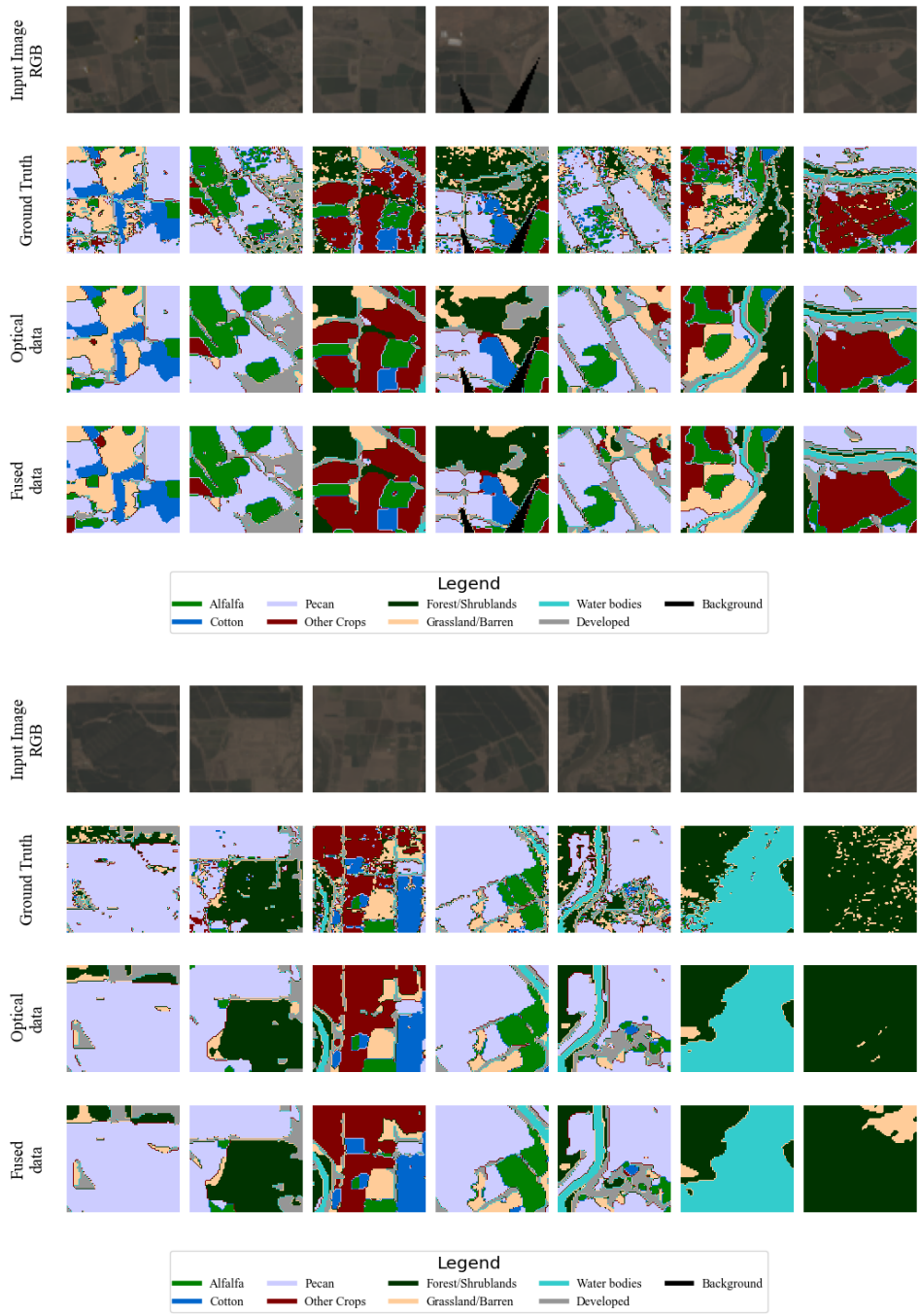



Fig. 7. Visual comparison of land cover predictions on two different input configurations. Optical Imagery (16 bands), and Optical Imagery fused with Synthetic Aperture Radar (SAR) data and Digital Elevation Model (DEM) (21 bands)

is identified as a potential source of confusion (Cum. Impact $< 0.05$), impairing the model's predictive accuracy. For Pecan (Class 2), the model demonstrates clear distinguishability from all classes except Alfalfa/Hay (Cum. Impact $< 0.25$), which again presents some challenge. Finally, for Class 3: Other Crops, the presence of Alfalfa/Hay is noted as having a negative impact on the model's prediction accuracy, with Cum. Impact equal to zero.

Figure 11 illustrates the interaction of different classes within the model's predictions for Class 4: Forest/Shrublands, Class 5: Barren/Grassland, Class 6: Water Bodies, and Class 7: Developed. The analysis reveals that Class 4 (Forest/Shrublands) is generally distinguishable from other classes, except for Other Crops, where the cumulative impact is less than 0.20. This indicates some level of confusion between these two classes. Moreover, Class 4 shows a lower cumulative impact in relation to classes such as Alfalfa/Hay, Cotton, Other Crops, and Water Bodies, making it the most sensitive class to the presence of other classes and the underlying pixel architecture within a tile. On the other hand, the Water Bodies

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully edited content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2025.3532829



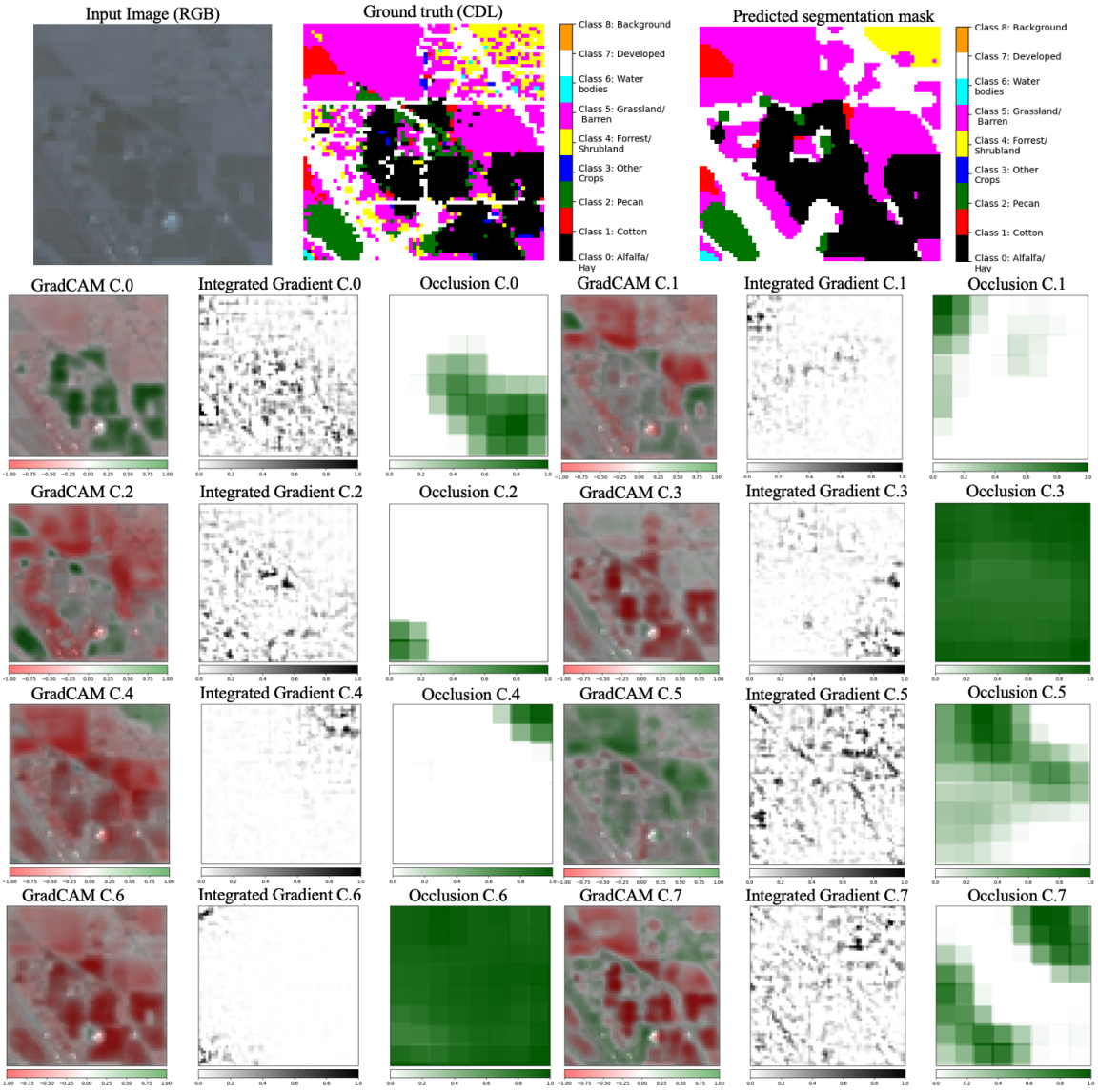



Fig. 8. Visualization of Model Explainability Techniques Applied to a Sample Input Tile. The figure displays the input tile in RGB bands, the corresponding ground truth mask derived from the Cropland Data Layer (CDL), and the predicted segmentation mask generated by the model. Saliency maps for each class present in the tile (C.0 to C.7), generated using three different explainability techniques: Grad-CAM, Integrated Gradients (IG), and Occlusion. Classes: 0- Alfalfa/Hay, 1- Cotton, 2- Pecan, 3- Other Crops, 4- Forrest/ Shrubland, 5- Barren/Grassland, 6- Water bodies, 7- Developed

(Class 6) and Developed (Class 7) classes exhibit greater robustness, showing less sensitivity to the influence of other classes. Their cumulative impact values exceed 0.20 for all other classes, indicating that these classes are less affected by the presence of other classes in the model's predictions. This suggests that Water Bodies and Developed areas are more consistently identified by the model, regardless of the presence of other land cover types.

*3) Spectral-Grad-CAM:* Figure 12 illustrates the contribution of each spectral band of the fused image to the prediction of each target class using Spectral-Grad-CAM. The analysis reveals that certain spectral bands exhibit significant contributions to detecting specific crop types. For Cotton, notable, influential bands include blue, green, shortwave infrared 2 (SWIR2), land surface temperature (LST), and Normalized Difference Vegetation Index (NDVI) corresponding to July and NDVI corresponding to December median imagery within the optical domain. Additionally, the Vertical-Horizontal (VH) polarization band of the Synthetic Aperture Radar (SAR) sensor for July demonstrates substantial influence, with Grad-CAM values ranging from -1 to 1, indicating its capability in classifying this certain class (Figure 12.A). Further analysis of Cotton detection reveals that the green band and, more prominently, SWIR2 corresponding to July, exhibit the most substantial impact on classification accuracy. These bands display a wide range of Grad-CAM values spanning from -1 to 1. Moreover, the Vertical-Vertical (VV) polarization of SAR for July also demonstrates significant influence, containing pixels with Grad-CAM values across the entire -0.75 to 1 spectrum (Figure 12.B). In contrast, for Pecan classification, the LST band for July emerges as the most critical predictor, accompanied by substantial contributions from the red band

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 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/JSTARS.2025.3532829



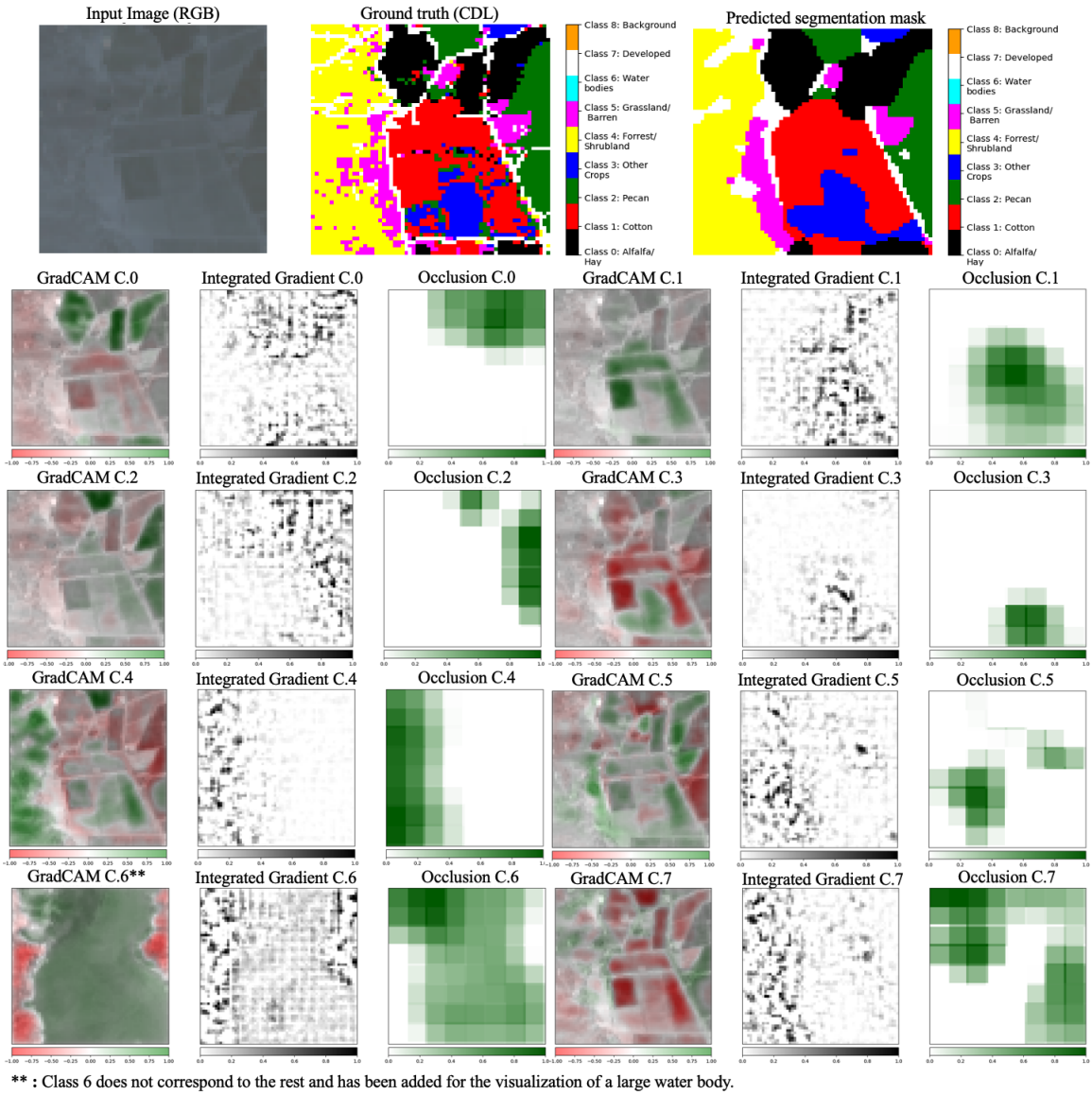



Fig. 9. Visualization of Model Explainability Techniques Applied to a Sample Input Tile. The figure displays the input tile in RGB bands, the corresponding ground truth mask derived from the Cropland Data Layer (CDL), and the predicted segmentation mask generated by the model. Saliency maps for each class present in the tile (C.0 to C.7), generated using three different explainability techniques: Grad-CAM, Integrated Gradients (IG), and Occlusion. Classes: 0- Alfalfa/Hay, 1- Cotton, 2- Pecan, 3- Other Crops, 4- Forrest/ Shrubland, 5- Barren/Grassland, 6- Water bodies, 7- Developed

in July and SWIR for December. This differential spectral importance underscores the crop-specific nature of spectral band contributions to accurate classification (Figure 12.C).

For the ”Other Crops” class, the blue and NIR bands, LST, VH polarization for July, along with SWIR2 and NDVI for December, emerge as the most influential spectral features. These bands exhibit a wide distribution of Grad-CAM values, ranging from -1 to 1, indicating their significant contribution to the model’s predictive capacity for this heterogeneous class (Figure 13. A). In the case of ”Forest/Shrubland,” key spectral indicators include the red band and NDVI for July, SWIR2 and NDVI for December, Vertical-Vertical (VV) polarization for July, and the Digital Elevation Model (DEM). The red band and NDVI for July demonstrate a substantial range of Grad-CAM values, underscoring their importance in discriminating this class. Similarly, the December SWIR2 and NDVI, July VV polarization, and DEM exhibit high impact, as evidenced by their wide Grad-CAM value distribution, rendering them crucial for accurate detection of forested and shrubland areas (Figure 13.B). For the ”Barren/Grassland” class, NDVI for July, blue and red bands for December, SWIR2 for December, VH polarization for July, and DEM are identified as the most influential spectral features. These bands are characterized by a substantial range of Grad-CAM values, spanning from -1 to 1, indicating their critical role in the model’s classification process for barren and grassland areas (Figure 13.C).

For the detection of ”Water Bodies,” the most salient spectral features include the DEM, VH polarization for December, NDVI and VV polarization for July, SWIR2, LST, and red bands for December. These spectral indicators exhibit an extensive distribution of Grad-CAM values, signifying their pivotal role in the model’s classification of hydrological fea-

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 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/JSTARS.2025.3532829



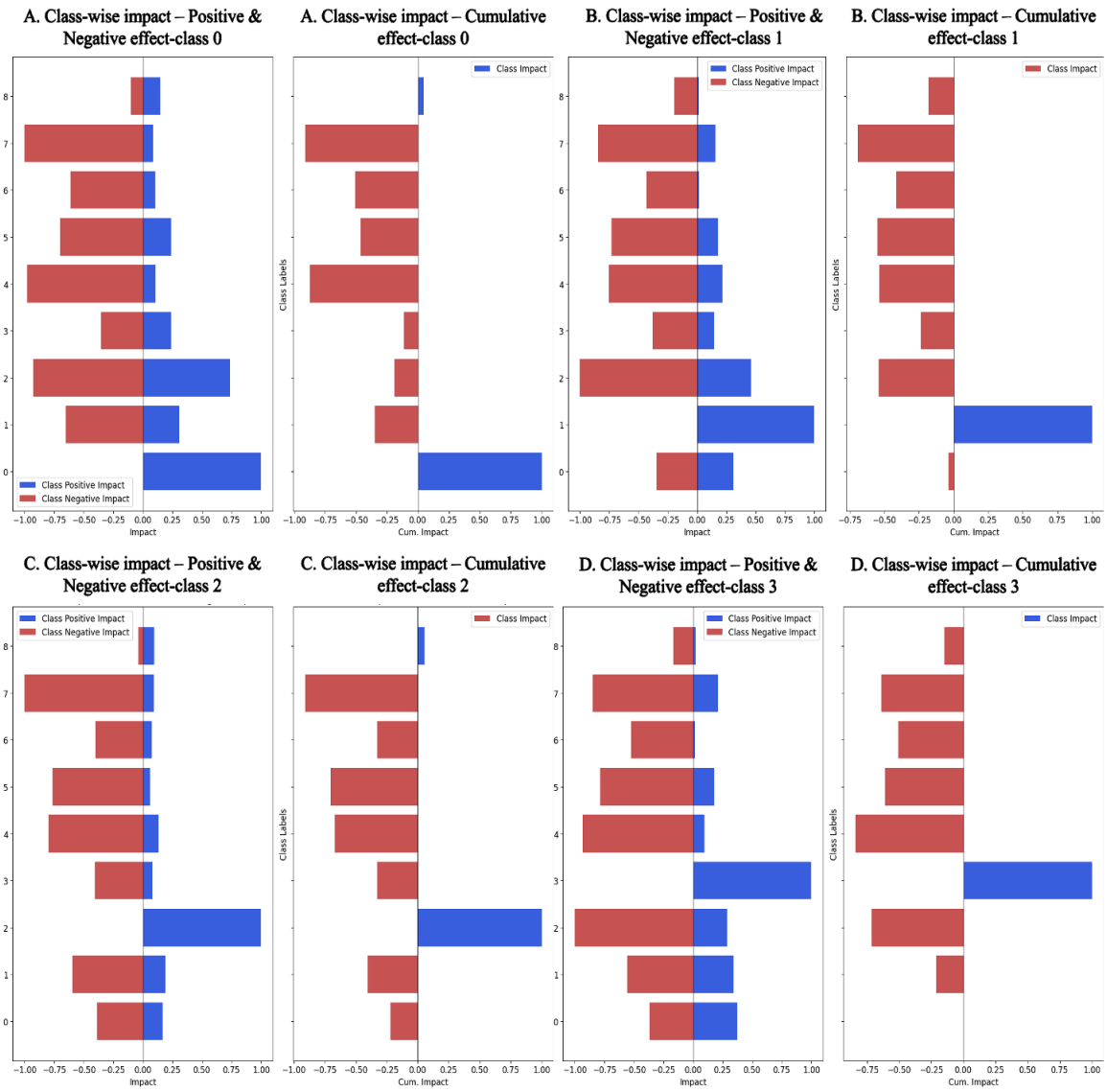



Fig. 10. Depicts the Normalized Positive Impact, Normalized Negative Impact, and Cumulative Impact of class interactions within a tile on the model's predictions. The classes represented include: 0 - Alfalfa/Hay, 1 - Cotton, 2 - Pecan, 3 - Other Crops 4 - Forest/Shrubland, 5 - Barren/Grassland, 6 - Water Bodies, and 7 – Developed.

tures (Figure 14.A). The broad range of Grad-CAM values observed for these bands reflects their substantial influence on the accurate delineation of water bodies within the study area. In contrast, the "Developed" class demonstrates a notably different pattern of spectral importance. No single spectral band exhibits a dominant range of Grad-CAM values for this class (Figure 14.B). This observation suggests that the model employs a more distributed approach, integrating information from multiple spectral features rather than relying on a single dominant band. The absence of pronounced Grad-CAM value ranges for individual bands indicates that the classification of developed areas involves a more complex, multi-dimensional spectral analysis. This differential spectral response between the "Water Bodies" and "Developed" classes underscores the variability in feature importance across land cover types. For water bodies, specific spectral bands and indices play a crucial role, likely due to the distinctive spectral signatures of water in these wavelengths (Figure 14.A). Conversely, the classification of developed areas appears to require a more nuanced integration of diverse spectral information, reflecting the heterogeneous nature of urban and built-up environments (Figure 14.B).

### C. Crop & Land Cover Land Use Prediction

The map depicted in Figure 15 illustrates the predicted land cover classification of the Middle Rio Grande (MRG) region for the year 2023. This map covers both the U.S. and Mexico portions of the MRG basin, providing a comprehensive overview of land use patterns across the transboundary region. This map was generated using MANet-B5 as the segmentation model and dual-monthly multi sensor imagery was the input.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully edited content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2025.3532829



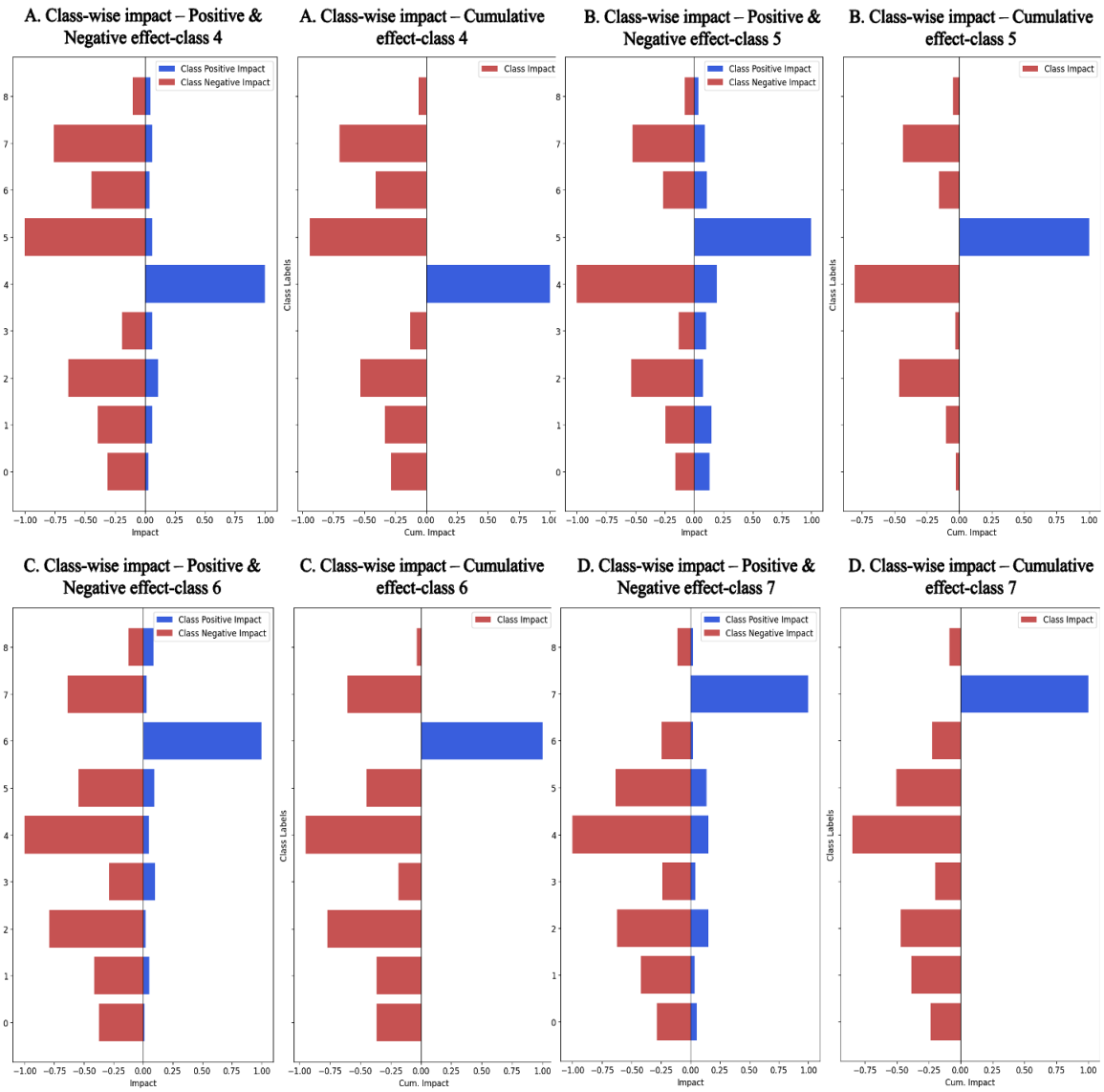



Fig. 11. Depicts the Normalized Positive Impact, Normalized Negative Impact, and Cumulative Impact of class interactions within a tile on the model's predictions. The classes represented include: 0 - Alfalfa/Hay, 1 - Cotton, 2 - Pecan, 3 - Other Crops 4 - Forest/Shrubland, 5 - Barren/Grassland, 6 - Water Bodies, and 7 – Developed.

## V. Discussion

This study introduces two novel XAI frameworks, leveraging Grad-CAM-based analyses of class interactions and spectral contributions to provide deeper insights into model performance and transparency in complex real-world applications, such as mapping heterogeneous landscapes. A semantic segmentation model utilizing the Multi-Attention Network (MANet) architecture, integrated with a modified Mix Vision Transformer (MiT) encoder, was developed to process multispectral, pixel-level, multi-sensor fused inputs. The inputs comprise median optical imagery (Landsat 8) composites for July and December, capturing temporal variations in crop-specific classes, along with Synthetic Aperture Radar (Sentinel-1) data and a Digital Elevation Model (DEM). This methodology addresses practical challenges in CLCLU mapping, including data asymmetry across the US-Mexico transboundary region. By systematically evaluating the influence of class types and specific spectral bands on model predictions, this approach enhances explainability while improving the effectiveness of CLCLU mapping applications.

### A. Impact of Fusion on CLCLU

The integration of SAR and DEM data with the optical median composites from Landsat 8 has led to a modest improvement in the mean Intersection over Union (mIoU) by approximately one percent (Figure 4). Although this increase may appear minor, potentially due to the minimal cloud cover in the optical imagery during the selected month, a notable outcome is a consistent improvement across all class IoU scores (Figure 5). Incorporating SAR and DEM data resulted in an approximately one percent increase in IoU for each land cover class. This observation is particularly significant,

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully e
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2025.3532829



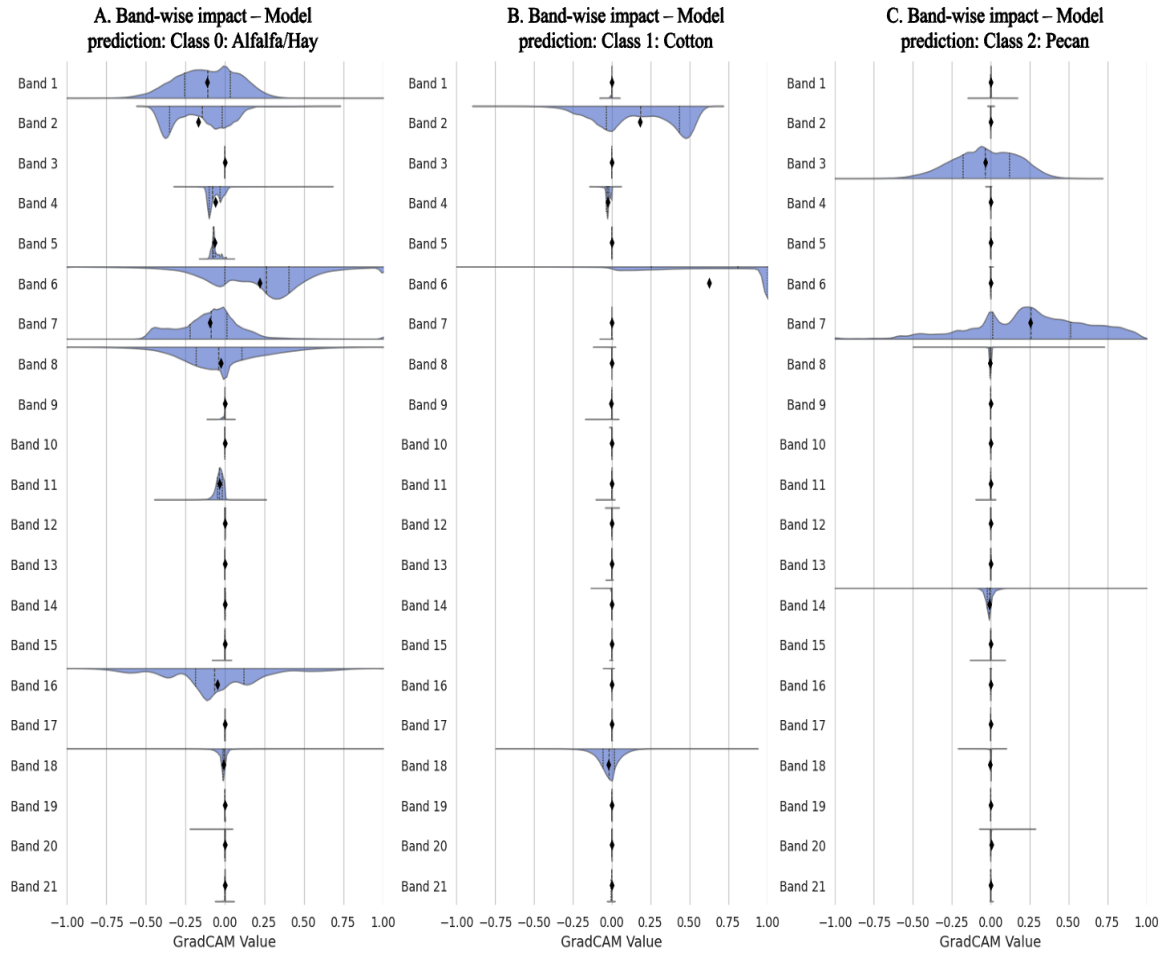



Fig. 12. Spectral impact on the prediction of A. alfalfa/Hay, B. Cotton, and C. Pecan classes and elucidates the model's capacity to discern specific classes using individual spectral bands. The graph show Grad-CAM attribution values for each band as input to the model for 500 random tiles with each specific target class, where the width of the graph shows the normalized abundance of Grad -CAM values over the -1 to 1 span.

highlighting that the fusion of diverse data sources can yield uniform benefits across different land cover categories. In contrast, increasing the number of optical bands through the inclusion of additional temporal composites, while potentially resulting in a more substantial overall mIoU improvement, only sometimes leads to uniform gains across all classes. This indicates that the added spectral information may disproportionately benefit certain classes while having a limited impact on others. The uniformity observed in the class-specific IoUc improvements when integrating multi-sensor data suggests that the complementary information provided by SAR and DEM enhances the model's generalization capability across diverse land cover types, a critical factor in CLCLU mapping.

### *B. Conventional XAI comparison*

The saliency maps generated using Grad-CAM, Integrated Gradients (IG), and Occlusion metrics demonstrate that Grad-CAM provides more interpretable insights into the model's decision-making logic compared to the other methods (Figures 8 and 9). The saliency maps derived from Grad-CAM effectively align with the model's prediction masks, even in cases where a class occupies a minimal number of pixels within the tile. This consistency underscores the ability of Grad-CAM to offer interpretable visualizations that accurately reflect the spatial distribution of predicted classes. In contrast, saliency maps produced by IG often fail to deliver interpretable outputs that adequately elucidate the model's prediction process. This limitation may be attributed to the low spatial resolution of the satellite imagery and the relatively small size of certain land cover types within the input tiles. These factors hinder IG from generating comprehensible saliency maps, particularly when compared to applications in more traditional computer vision tasks, such as classes present in the ImageNet dataset, where the target class typically occupies a substantial portion of the input image. However, in scenarios where the predicted class covers a larger area within the tile, the IG-derived saliency maps are slightly more interpretable (Figure 9 C.6). On the other hand, Occlusion-based saliency maps demonstrate a more precise delineation of regions that are critical to the model's decision-making process. This method effectively highlights the spatial areas contributing most significantly to the model's predictions. Nevertheless, the Occlusion metric encounters challenges in accurately reflecting class importance when the predicted class comprises only a small number

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully e
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2025.3532829



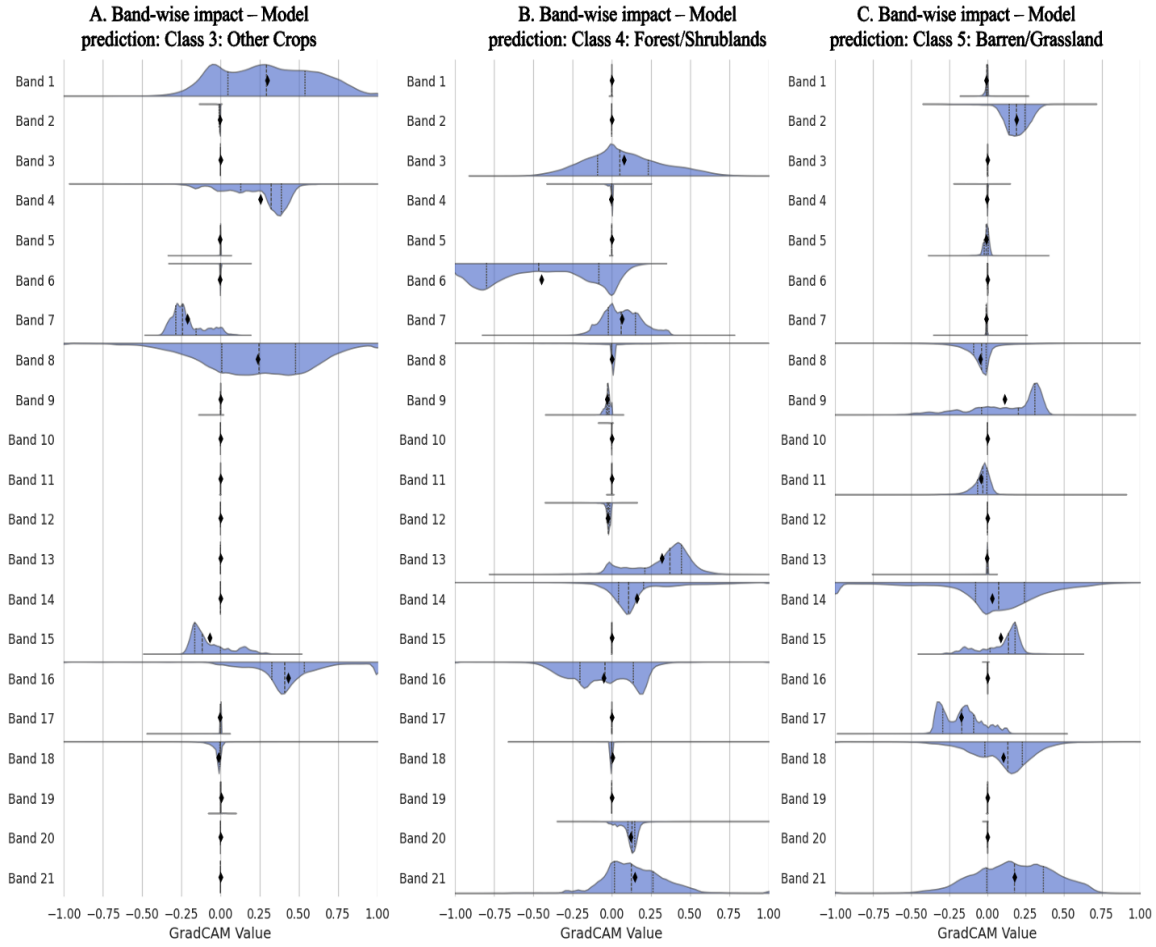



Fig. 13. Spectral impact on the prediction of A. Other Crops, B. Forest/Shrubland, and C. Barren/Grassland classes and elucidates the model's capacity to discern specific classes using individual spectral bands. The graphs show Grad -CAM attribution values for each band as input to the model for 500 random tiles with each specific target class, where the width of the graph shows the normalized abundance of Grad -CAM values over the -1 to 1 span.

of pixels (Figure 8 C.3 and C.6). This suggests that while Occlusion can be useful for identifying key regions, it may struggle to produce reliable interpretability in scenarios with sparse class representation.

### C. Interclass-Grad-CAM

The overall comparison of these saliency metrics highlights the strengths of Grad-CAM in providing a coherent understanding of model predictions in the context of CLCLU classification, particularly when dealing with complex, heterogeneous landscapes characterized by variable spatial extents of target classes. Therefore, this metric was used for further analysis of the models, where an investigation of model behavior over the entire dataset was performed instead of having an instance-oriented understanding. The analysis of class interactions revealed key insights regarding the model's ability to detect and differentiate between crop-specific classes in the presence of other land cover types. Notably, the model successfully distinguishes cotton from other classes, except in cases where pecan is present, which negatively impacts the model's ability to identify cotton accurately. Specifically, when cotton is the target class, the presence of alfalfa/hay pixels can introduce bias into the model's predictions, reducing the reliability of the prediction. Moreover, distinguishing between the "other crops" class and alfalfa/hay proves to be particularly challenging, leading to higher ambiguity in the model's predictions. In other words, the presence of "other crops" pixels has the most adverse effect on the model's reliability when predicting alfalfa/hay. For the remaining land cover types, the model tends to perform better when non-agricultural categories dominate the scene. However, the presence of crop-specific pixels often has a detrimental effect on the prediction accuracy of these classes. For example, the presence of "other crops" pixels negatively impacts the model's reliability in predicting forest/shrubland, while alfalfa/hay, cotton, and "other crops" classes reduce the model's reliability in detecting barren/grassland areas. These results highlight the complexity of class interactions and their influence on the model's performance.

Although the model is robust in detecting different landcover types, illustrated in Figure 6, the Interclass-Grad-CAM analysis suggests that crop-specific classes introduce considerable uncertainties in the prediction process, leading to reduced confidence across various land cover types. These findings may

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully e
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2025.3532829



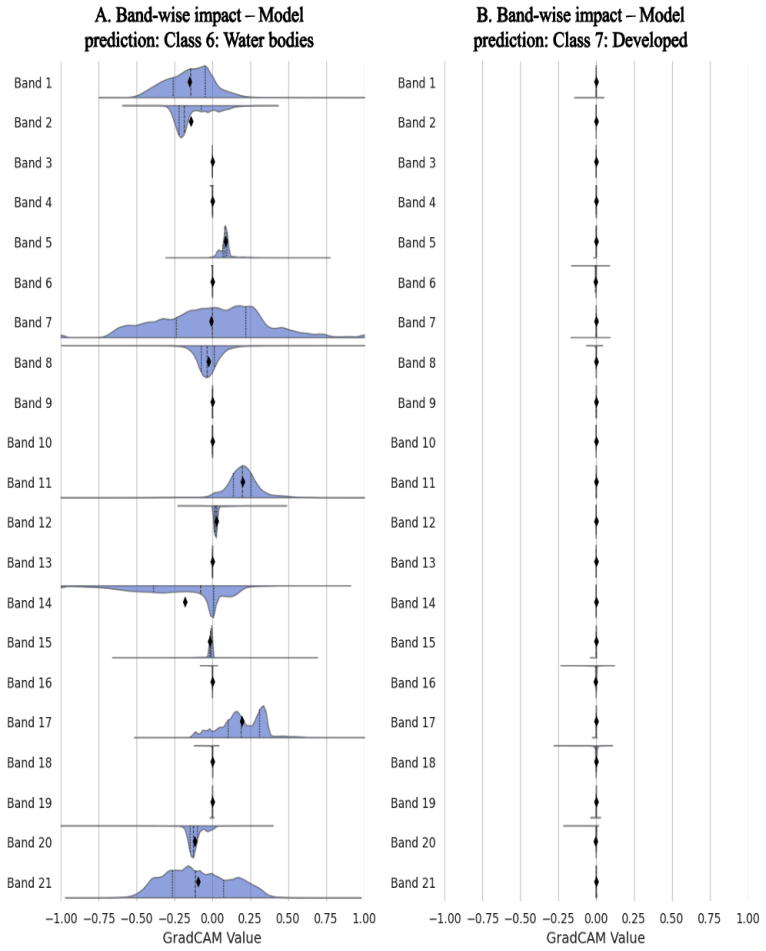



Fig. 14. Spectral impact on the prediction of A. Waterbodies, and B. Developed classes and elucidates the model's capacity to discern specific classes using individual spectral bands. The graphs show Grad -CAM attribution values for each band as input to the model for 500 random tiles with each specific target class, where the width of the graph shows the normalized abundance of Grad -CAM values over the -1 to 1 span.

be attributed to the limited spatial coverage of farmland classes and the challenges associated with delineating these specific classes using moderate-resolution imagery.

### *D. Spectral-Grad-CAM*

The spectral (feature) importance analysis demonstrated that, for each target class, specific bands possess the ability to distinctly influence the land cover classification, with Grad-CAM values ranging from -1 to 1. Conversely, certain bands exhibit minimal or no contribution to the model's prediction process. By normalizing the Grad-CAM values and dividing them into eight equal intervals of 0.25, we ranked each spectral band based on the span of its Grad-CAM value within each interval, as shown in Figure 16. In this ranking system, bands with a span from -1 to 1 were assigned a value of 8, indicating a significant contribution to classifying a particular class. In contrast, bands with a span less than 0.25 span received a value of 0, implying no independent contribution to the classification process. Figure 16 illustrates the contribution of each spectral band individually in predicting specific classes (Figure 16.A) and across all classes (Figure 16.B). The figure indicates that optical imagery provides more critical information for delineating crop-specific classes compared to SAR composites and DEM data. It also highlights the substantial impact of SAR and DEM bands in distinguishing non-agricultural land types. Notably, surface temperature and vertical-horizontal (VH) polarization from July imagery exert the most significant influence on the model's decision-making process overall. In contrast, the green band and vertical-vertical (VV) polarization show negligible impact on the model's predictions. Depending on the application's objectives, this insight can be leveraged to enhance data structuring and fine-tuning strategies. Additionally, it underscores the importance of considering domain adaptability and transfer learning approaches when applying the model to diverse or heterogeneous agricultural landscapes. However, it is critical to note that this analysis considers each band's sole impact on the model prediction and does not investigate any combination of spectra that can be explored in future versions.

## VI. CONCLUSION

This study introduces two novel analytical frameworks, Interclass-Grad-CAM and Spectral-Grad-CAM, that extend the Grad-CAM technique to provide a deeper and more

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 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/JSTARS.2025.3532829



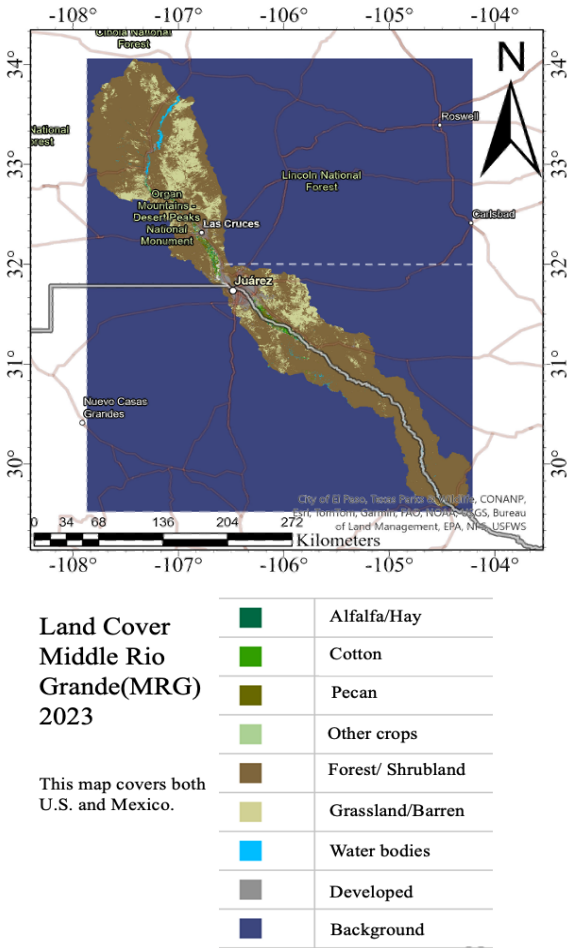



Fig. 15. Predicted Middle Rio Grande (MRG) land Cover Map 2023

comprehensive understanding of model behavior in CLCLU applications. These frameworks address critical challenges in model interpretability and reliability by offering unique insights into the interactions between different land cover types and the importance of specific spectral bands in the classification process.

Interclass-Grad-CAM enables a more nuanced understanding of how the presence of certain classes affects the prediction of target classes, which is particularly important in CLCLU tasks characterized by complex spatial arrangements and overlapping land cover types. By elucidating these interclass interactions, this method enhances the model's ability to differentiate between closely related classes, such as different crop types or combinations of crops and natural vegetation. This contributes to more accurate and reliable land cover mapping, especially in heterogeneous environments where traditional methods often struggle. Spectral-Grad-CAM quantifies the contribution of individual spectral bands to the model's decision-making process, providing valuable insights into which features are most informative for specific land cover types. The findings reveal that certain spectral features, such as VH polarization, are crucial for accurately identifying non-agricultural classes, while other bands, like the green and VV polarization band, have minimal impact. This detailed understanding of spectral importance can inform future data collection strategies, model refinement, and the selection of spectral inputs, optimizing performance for specific applications, such as detecting subtle variations in vegetation health or early signs of land degradation.

Beyond CLCLU tasks, the novel Grad-CAM extensions presented in this study have potential applications in other fields that require high interpretability and reliability in data-scarce and heterogeneous environments. The integration of multi-sensor data, combining SAR and DEM with optical imagery, demonstrated a consistent improvement in classification performance across all land cover types, as indicated by a uniform increase of approximately 1% in the mean IoU metric. While this improvement may seem modest, its uniformity across diverse classes highlights the enhanced generalization capability of the model when multiple data sources are fused. This is a crucial outcome for CLCLU applications, where diverse environmental conditions and mixed land cover types are common. Such balanced improvements suggest that integrating multi-sensor data is more effective for achieving reliable performance across various land cover types than simply adding more spectral bands from the same sensor.

## VII. Data and Code Availability

The source code for data preprocessing, automated training tile generation from satellite imagery, and PyTorch implementations of all model architectures utilized in this study, along with training and evaluation protocols, are publicly available in a GitHub repository. These resources are purportedly accessible at https://github.com/waterdmd/DataFusionForLC_with_Interpretability. The open dissemination of these materials promotes scientific transparency and facilitates study replication. Furthermore, it provides a valuable resource for researchers and practitioners in the domain of land cover and land use classification employing moderate resolution multispectral imagery, potentially catalyzing further advancements in the field.

## VIII. Acknowledgment

The authors acknowledge research funding support from NASA's LCLUC program. A research grant for the project *"Exploring the Nexus between LCLUC, Socio-Economic Factors, and Water for a Vulnerable Arid US-Mexico Transboundary Region"* (#80NSSC23K0507) supported this work.

## References

[1] K. Azizi, S. K. Diko, and C. I. Meier, "A Citizen Science Approach to the Characterisation and Modelling of Urban Pluvial Flooding," vol. 16, no. 1, 2023.

[2] S. Ebrahimi, K. Azizi, M. Khorram, A. R. Kashani, and A. S. A. Ali, "Evaluation of hydrological variabilities of water quality models considering event-based scenarios: A case study," *Stochastic Environmental Research and Risk Assessment*, vol. 38, no. 6, pp. 2425–2449, Jun. 2024, doi: 10.1007/s00477-024-02687-2.

[3] T. Huang et al., "An Earth System Digital Twin for Flood Prediction and Analysis," in *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, Kuala Lumpur, Malaysia: IEEE, Jul. 2022, pp. 4735–4738, doi: 10.1109/IGARSS46834.2022.9884830.

[4] C. Justice, G. Gutman, and K. P. Vadrevu, "NASA Land Cover and Land Use Change (LCLUC): An interdisciplinary research program," *Journal of Environmental Management*, vol. 148, pp. 4–9, Jan. 2015, doi: 10.1016/j.jenvman.2014.12.004.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 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/JSTARS.2025.3532829



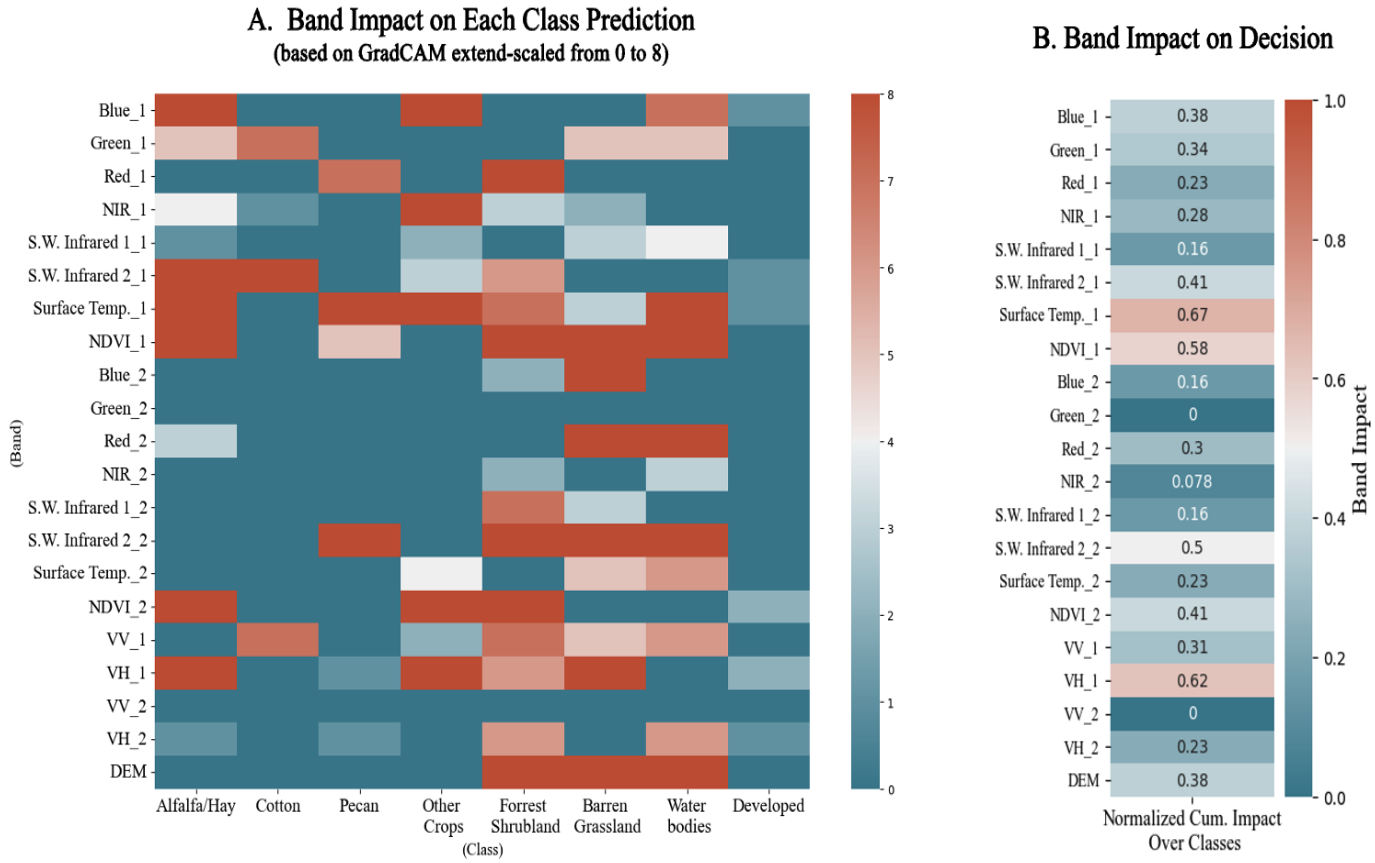



Fig. 16. Analysis of Spectral (feature) Contribution for Class Prediction and Overall Model Decision-Making. (A) illustrates the impact of each spectral band on predicting individual classes, scaled based on Grad-CAM values from 0 to 8. Panel (B) represents the normalized cumulative impact of each spectral band on overall model decision-making.

[5] A. Raad and J. D. White, "Socioeconomic and Climatic Drivers of Land-Cover and Land-Use Change: An Assessment with NASA's Sedac Datasets, GIS, and Remote Sensing Techniques," Jun. 28, 2023, Rochester, NY: SSRN, doi: 10.2139/ssrn.4494618.

[6] W. Tay and Y. He, "Detecting land cover and land use change and its impact on biomass carbon emission from 2001 to 2019 in Arkansas, U.S.A," *Biogeochemistry*, vol. 165, no. 1, pp. 29–42, Aug. 2023, doi: 10.1007/s10533-023-01042-x.

[7] B. Chen, B. Huang, and B. Xu, "Multi-source remotely sensed data fusion for improving land cover classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 124, pp. 27–39, Feb. 2017, doi: 10.1016/j.isprsjprs.2016.12.008.

[8] G. Chen et al., "Multifunctional Models in Digital and Physical Twinning of the Built Environment—A University Campus Case Study," *Smart Cities*, vol. 7, no. 2, Art. no. 2, Apr. 2024, doi: 10.3390/smartcities7020035.

[9] M. Dumitrașcu, Y. Himiyama, M. Gabrovec, M. Kopecká, L. Kupková, and I. Bicik, "Land Use and Land Cover Changes in a Global Environmental Change Context—The Contribution of Geography," in *Research Directions, Challenges and Achievements of Modern Geography*, J. Bański and M. Meadows, Eds., Singapore: Springer Nature, 2023, pp. 175–203, doi: 10.1007/978-981-99-6604-2_10.

[10] N. A. Houston, J. V. Thomas, L. K. Foster, D. E. Pedraza, and T. L. Welborn, "Hydrogeologic framework and groundwater characterization in selected alluvial basins in the upper Rio Grande basin, Colorado, New Mexico, and Texas, United States, and Chihuahua, Mexico, 1980 to 2015," *U.S. Geological Survey*, 2021–5035, 2021, doi: 10.3133/sir20215035.

[11] B. Wu, Z. Fu, B. Fu, C. Yan, H. Zeng, and W. Zhao, "Dynamics of land cover changes and driving forces in China's drylands since the 1970s," *Land Use Policy*, vol. 140, p. 107097, May 2024, doi: 10.1016/j.landusepol.2024.107097.

[12] G. Z. Xian, T. Loveland, S. M. Munson, J. E. Vogelmann, X. Zeng, and C. J. Homer, "Climate sensitivity to decadal land cover and land use change across the conterminous United States," *Global and Planetary Change*, vol. 192, p. 103262, Sep. 2020, doi: 10.1016/j.gloplacha.2020.103262.

[13] M. Elliott, Á. Borja, and R. Cormier, "Managing marine resources sustainably—Ecological, societal and governance connectivity, coherence and equivalence in complex marine transboundary regions," *Ocean and Coastal Management*, vol. 245, p. 106875, Nov. 2023, doi: 10.1016/j.ocecoaman.2023.106875.

[14] E. Golovina and O. Shchelkonogova, "Possibilities of Using the Unitization Model in the Development of Transboundary Groundwater Deposits," *Water*, vol. 15, no. 2, Art. no. 2, Jan. 2023, doi: 10.3390/w15020298.

[15] S. Namany, R. Govindan, and T. Al-Ansari, "Operationalising transboundary cooperation through game theory: An energy water food nexus approach for the Middle East and North Africa," *Futures*, vol. 152, p. 103198, Sep. 2023, doi: 10.1016/j.futures.2023.103198.

[16] J. Ngondo, J. Mango, R. Liu, J. Nobert, A. Dubi, and H. Cheng, "Land-Use and Land-Cover (LULC) Change Detection and the Implications for Coastal Water Resource Management in the Wami–Ruvu Basin, Tanzania," *Sustainability*, vol. 13, no. 8, Art. no. 8, Jan. 2021, doi: 10.3390/su13084092.

[17] P. Verma, P. Singh, and S. K. Srivastava, "Impact of land use change dynamics on sustainability of groundwater resources using earth observation data," *Environment, Development and Sustainability*, vol. 22, no. 6, pp. 5185–5198, 2020, Accessed: Aug. 14, 2024. [Online]. Available: https://ideas.repec.org//a/spr/endesu/v22y2020i6d10.1007_s10668-019-00420-6.html.

[18] R. Yonaba et al., "Spatial and transient modelling of land use/land cover (LULC) dynamics in a Sahelian landscape under semi-arid climate in northern Burkina Faso," *Land Use Policy*, vol. 103, p. 105305, Apr. 2021, doi: 10.1016/j.landusepol.2021.105305.

[19] "Copernicus Data Space Ecosystem — Europe's eyes

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully e
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2025.3532829



on Earth," Accessed: Jul. 18, 2024. [Online]. Available: https://dataspace.copernicus.eu/.

[20] "Landsat 8 — Landsat Science," Accessed: Jun. 25, 2024. [Online]. Available: https://landsat.gsfc.nasa.gov/satellites/landsat-8/.

[21] L. L. Coulter et al., "Classification and assessment of land cover and land use change in southern Ghana using dense stacks of Landsat 7 ETM+ imagery," *Remote Sensing of Environment*, vol. 184, pp. 396–409, Oct. 2016, doi: 10.1016/j.rse.2016.07.016.

[22] Y. Dong et al., "Recording Urban Land Dynamic and Its Effects during 2000–2019 at 15-m Resolution by Cloud Computing with Landsat Series," *Remote Sensing*, vol. 12, no. 15, Art. no. 15, Jan. 2020, doi: 10.3390/rs12152451.

[23] I. Gallo, R. La Grassa, N. Landro, and M. Boschetti, "Sentinel 2 Time Series Analysis with 3D Feature Pyramid Network and Time Domain Class Activation Intervals for Crop Mapping," *ISPRS International Journal of Geo-Information*, vol. 10, no. 7, p. 483, Jul. 2021, doi: 10.3390/ijgi10070483.

[24] S. Soni, S. Shiwani, and A. Arora, "Impact of Land Cover Land Use Changes (LCLUC) on Earth Resources & Human Life—A Literature Survey," Mar. 05, 2020, Rochester, NY: SSRN, doi: 10.2139/ssrn.3549156.

[25] J. Svoboda, P. Štych, J. Laštovička, D. Paluba, and N. Kobliuk, "Random Forest Classification of Land Use, Land-Use Change and Forestry (LULUCF) Using Sentinel-2 Data—A Case Study of Czechia," *Remote Sensing*, vol. 14, no. 5, Art. no. 5, Jan. 2022, doi: 10.3390/rs14051189.

[26] "USDA - National Agricultural Statistics Service - Research and Science - Cropland Data Layers," Accessed: Jun. 25, 2024. [Online]. Available: https://www.nass.usda.gov/Research_and_Science/Cropland/sarsfaqs2.php.

[27] W. Li et al., "Stacked Autoencoder-based deep learning for remote-sensing image classification: a case study of African land-cover mapping," *International Journal of Remote Sensing*, vol. 37, no. 23, pp. 5632–5646, Dec. 2016, doi: 10.1080/01431161.2016.1246775.

[28] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature Pyramid Networks for Object Detection," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI: IEEE, Jul. 2017, pp. 936–944, doi: 10.1109/CVPR.2017.106.

[29] W. Liu et al., "SSD: Single Shot MultiBox Detector," in *Computer Vision – ECCV 2016*, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds., Cham: Springer International Publishing, 2016, pp. 21–37, doi: 10.1007/978-3-319-46448-0_2.

[30] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," May 18, 2015, arXiv: arXiv:1505.04597. Accessed: May 30, 2024. [Online]. Available: http://arxiv.org/abs/1505.04597.

[31] M. J. Shafiee, B. Chywl, F. Li, and A. Wong, "Fast YOLO: A Fast You Only Look Once System for Real-time Embedded Object Detection in Video," Sep. 18, 2017, arXiv: arXiv:1709.05943, doi: 10.48550/arXiv.1709.05943.

[32] W. Yu, K. Yang, Y. Bai, T. Xiao, H. Yao, and Y. Rui, "Visualizing and Comparing AlexNet and VGG using Deconvolutional Layers," in *Proceedings of the 33rd International Conference on Machine Learning (ICML)*, vol. 48, 2016.

[33] S. Basu, S. Ganguly, S. Mukhopadhyay, R. DiBiano, M. Karki, and R. Nemani, "DeepSat: a learning framework for satellite imagery," in *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems*, in SIGSPATIAL '15. New York, NY, USA: Association for Computing Machinery, Nov. 2015, pp. 1–10, doi: 10.1145/2820783.2820816.

[34] B. Feizizadeh, K. Mohammadzade Alajujeh, T. Lakes, T. Blaschke, and D. Omarzadeh, "A comparison of the integrated fuzzy object-based deep learning approach and three machine learning techniques for land use/cover change monitoring and environmental impacts assessment," *GIScience and Remote Sensing*, vol. 58, no. 8, pp. 1543–1570, 2021, doi: 10.1080/15481603.2021.2000350.

[35] S. M. Mirmazloumi et al., "ELULC-10, a 10 m European Land Use and Land Cover Map Using Sentinel and Landsat Data in Google Earth Engine," *Remote Sensing*, vol. 14, no. 13, Art. no. 13, Jan. 2022, doi: 10.3390/rs14133041.

[36] M. Rußwurm and M. Körner, "Multi-Temporal Land Cover Classification with Sequential Recurrent Encoders," *ISPRS International Journal of Geo-Information*, vol. 7, no. 4, p. 129, Mar. 2018, doi: 10.3390/ijgi7040129.

[37] G. Singh, S. Singh, G. Sethi, and V. Sood, "Deep Learning in the Mapping of Agricultural Land Use Using Sentinel-2 Satellite Data," *Geographies*, vol. 2, no. 4, Art. no. 4, Dec. 2022, doi: 10.3390/geographies2040042.

[38] E. Sertel, B. Ekim, P. Ettehadi Osgouei, and M. E. Kabadayi, "Land Use and Land Cover Mapping Using Deep Learning Based Segmentation Approaches and VHR Worldview-3 Images," *Remote Sensing*, vol. 14, no. 18, Art. no. 18, Jan. 2022, doi: 10.3390/rs14184558.

[39] K. Hu, E. Zhang, X. Dai, M. Xia, F. Zhou, L. Weng, and H. Lin, "MCSGNet: An encoder–decoder architecture network for land cover classification," *Remote Sensing*, vol. 15, no. 11, p. 2810, 2023.

[40] R. Vohra and K. C. Tiwari, "Comparative Analysis of SVM and ANN Classifiers using Multilevel Fusion of Multi-Sensor Data in Urban Land Classification," *Sensing and Imaging*, vol. 21, no. 1, p. 17, Mar. 2020, doi: 10.1007/s11220-020-00280-9.

[41] J. Wang, C. Li, and P. Gong, "Adaptively weighted decision fusion in 30 m land-cover mapping with Landsat and MODIS data," *International Journal of Remote Sensing*, vol. 36, no. 14, pp. 3659–3674, Jul. 2015, doi: 10.1080/01431161.2015.1047049.

[42] B. Bigdeli, P. Pahlavani, and H. A. Amirkolaee, "An ensemble deep learning method as data fusion system for remote sensing multisensor classification," *Applied Soft Computing*, vol. 110, p. 107563, Oct. 2021, doi: 10.1016/j.asoc.2021.107563.

[43] N. Torbick, X. Huang, B. Ziniti, D. Johnson, J. Masek, and M. Reba, "Fusion of Moderate Resolution Earth Observations for Operational Crop Type Mapping," *Remote Sensing*, vol. 10, no. 7, Art. no. 7, Jul. 2018, doi: 10.3390/rs10071058.

[44] B. Hu et al., "Improving Urban Land Cover Classification with Combined Use of Sentinel-2 and Sentinel-1 Imagery," *ISPRS International Journal of Geo-Information*, vol. 10, no. 8, Art. no. 8, Aug. 2021, doi: 10.3390/ijgi10080533.

[45] P. Griffiths, C. Nendel, and P. Hostert, "Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping," *Remote Sensing of Environment*, vol. 220, pp. 135–151, Jan. 2019, doi: 10.1016/j.rse.2018.10.031.

[46] P. Dimitrov, E. Roumenina, D. Ganeva, A. Gikov, I. Kamenova, and V. Bozhanova, "Enhancing Pléiades-based crop mapping with multi-temporal and texture information," *Remote Sensing Applications: Society and Environment*, vol. 36, p. 101339, 2024, doi: 10.1016/j.rsase.2024.101339.

[47] B. Watkins and A. Van Niekerk, "Automating field boundary delineation with multi-temporal Sentinel-2 imagery," *Computers and Electronics in Agriculture*, vol. 167, p. 105078, 2019, doi: 10.1016/j.compag.2019.105078.

[48] A. Mercier et al., "Evaluation of Sentinel-1 & 2 time series for predicting wheat and rapeseed phenological stages," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 163, pp. 231–256, May 2020, doi: 10.1016/j.isprsjprs.2020.03.009.

[49] E. Cherif, M. Hell, and M. Brandmeier, "DeepForest: Novel Deep Learning Models for Land Use and Land Cover Classification Using Multi-Temporal and -Modal Sentinel Data of the Amazon Basin," *Remote Sensing*, vol. 14, no. 19, Art. no. 19, Jan. 2022, doi: 10.3390/rs14195000.

[50] A. Vali, S. Comai, and M. Matteucci, "Deep Learning for Land Use and Land Cover Classification Based on Hyperspectral and Multispectral Earth Observation Data: A Review," *Remote Sensing*, vol. 12, no. 15, Art. no. 15, Jan. 2020, doi: 10.3390/rs12152495.

[51] O. Sefrin, F. M. Riese, and S. Keller, "Deep Learning for Land Cover Change Detection," *Remote Sensing*, vol. 13, no. 1, Art. no. 1, Jan. 2021, doi: 10.3390/rs13010078.

[52] M. Martini, V. Mazzia, A. Khaliq, and M. Chiaberge, "Domain-Adversarial Training of Self-Attention-Based Networks for Land Cover Classification Using Multi-Temporal Sentinel-2 Satellite Imagery," *Remote Sensing*, vol. 13, no. 13, Art. no. 13, Jan. 2021, doi: 10.3390/rs13132564.

[53] O. Gorokhovatskyi, O. Peredrii, V. Gorokhovatskyi, and N. Vlasenko, "Chapter 7 - Explanation of CNN image classifiers with hiding parts," in *Explainable Deep Learning AI*, J. Benois-Pineau, R. Bourqui, D. Petkovic, and G. Quénot, Eds., Academic Press, 2023, pp. 125–146, doi: 10.1016/B978-0-32-396098-4.00013-2.

[54] A. Temenos, N. Temenos, M. Kaselimi, A. Doulamis, and N. Doulamis, "Interpretable Deep Learning Framework for Land Use and Land Cover Classification in Remote Sensing Using SHAP," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1–5, 2023, doi: 10.1109/LGRS.2023.3251652.

[55] P. Thunki, S. R. B. Reddy, M. Raparthi, S. Maruthi, S. B. Dodda, and P. Ravichandran, "Explainable AI in Data Science - Enhancing Model Interpretability and Transparency," *African Journal of Artificial Intelligence and Sustainable Development*, vol. 1, no. 1, Art. no. 1, Apr. 2021, Accessed: Aug. 14, 2024. [Online]. Available: https://africansciencegroup.com/index.php/AJAISD/article/view/12.

[56] R. Li et al., "Multiattention Network for Semantic Segmentation of Fine-Resolution Remote Sensing Images," *IEEE Transactions on*

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully e
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2025.3532829



*Geoscience and Remote Sensing*, vol. 60, pp. 1–13, 2022, doi: 10.1109/TGRS.2021.3093977.

[57] X. Yu, J. Wang, Y. Zhao, and Y. Gao, "Mix-ViT: Mixing attentive vision transformer for ultra-fine-grained visual categorization," *Pattern Recognition*, vol. 135, p. 109131, Mar. 2023, doi: 10.1016/j.patcog.2022.109131.

[58] O. Belhaj et al., "Determination of Change in Surface Waterbodies in The Middle Rio Grande Basin by Modified Normalized Difference Water Index (MNDWI) 1994-2020," vol. 4, pp. 67–80, Dec. 2022.

[59] M. Pascolini-Campbell, R. Seager, A. Pinson, and B. I. Cook, "Covariability of climate and streamflow in the Upper Rio Grande from interannual to interdecadal timescales," *Journal of Hydrology: Regional Studies*, vol. 13, pp. 58–71, Oct. 2017, doi: 10.1016/j.ejrh.2017.07.007.

[60] J. Schmandt, "Bi-national water issues in the Rio Grande/Río Bravo basin," *Water Policy*, vol. 4, no. 2, pp. 137–155, Jan. 2002, doi: 10.1016/S1366-7017(02)00007-7.

[61] J. T. Overpeck and B. Udall, "Climate change and the aridification of North America," *Proceedings of the National Academy of Sciences*, vol. 117, no. 22, pp. 11856–11858, Jun. 2020, doi: 10.1073/pnas.2006323117.

[62] Earth Resources Observation And Science (EROS) Center, "Collection-2 Landsat 8-9 OLI (Operational Land Imager) and TIRS (Thermal Infrared Sensor) Level-1 Data Products," *U.S. Geological Survey*, 2013, doi: 10.5066/P975CC9B.

[63] "Sentinels Toolboxes," Accessed: Aug. 25, 2024. [Online]. Available: https://sentiwiki.copernicus.eu/web/sentinels-toolboxes.

[64] "Sentinel-1 SAR GRD: C-band Synthetic Aperture Radar Ground Range Detected, log scaling — Earth Engine Data Catalog," Google for Developers. Accessed: Aug. 19, 2024. [Online]. Available: https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S1_GRD.

[65] "Copernicus DEM GLO-30: Global 30m Digital Elevation Model — Earth Engine Data Catalog," Google for Developers. Accessed: Aug. 19, 2024. [Online]. Available: https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_DEM_GLO30.

[66] "Copernicus Digital Elevation Model," Copernicus Contributing Missions Online. Accessed: Aug. 19, 2024. [Online]. Available: https://spacedata.copernicus.eu/collections/copernicus-digital-elevation-model.

[67] E. Xie, W. Wang, Z. Yu, A. Anandkumar, J. M. Alvarez, and P. Luo, "SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers," Oct. 28, 2021, arXiv: arXiv:2105.15203. Accessed: Jun. 03, 2024. [Online]. Available: http://arxiv.org/abs/2105.15203.

[68] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization," *International Journal of Computer Vision*, vol. 128, no. 2, pp. 336–359, Feb. 2020, doi: 10.1007/s11263-019-01228-7.

[69] M. Kolla and T. Venugopal, "Semantic image clustering with global average pooled deep convolutional autoencoder," *HELIX*, vol. 8, no. 4, pp. 3561–3566, 2018.

[70] V. Christlein, L. Spranger, M. Seuret, A. Nicolaou, P. Král, and A. Maier, "Deep Generalized Max Pooling," in *Proceedings of the 2019 International Conference on Document Analysis and Recognition (ICDAR)*, Sep. 2019, pp. 1090–1096, doi: 10.1109/ICDAR.2019.00177.

[71] M. Sundararajan, A. Taly, and Q. Yan, "Axiomatic Attribution for Deep Networks," Jun. 12, 2017, arXiv: arXiv:1703.01365. Accessed: Aug. 09, 2024. [Online]. Available: http://arxiv.org/abs/1703.01365.



**Saurav Kumar (Affiliate Member IEEE)** is an assistant professor in the School of Sustainable Engineering and Built Environment (SSEBE) at Arizona State University. He received his PhD from Virginia Tech, MS from the Singapore Stanford Partnership Program in 2005, and BE from Delhi College of Engineering in 2004. His current research focuses on developing models that use earth observations for decision-making and the application of aerial hyperspectral imaging for civil and environmental engineering applications.

**Saman Ebrahimi** is a Ph.D. student in Hydrosystems Engineering in the School of Sustainable Engineering and Built Environment (SSEBE) at Arizona State University. He holds a Bachelor's degree in Civil Engineering from K. N. Toosi University of Technology (2016) and a Master's degree from Tehran University (2019). Saman's research interests cover topics including Remote Sensing, Uncertainty Assessment, Data Fusion, Data Assimilation, Hydrology, and Digital twin for Earth observation.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/